

# Data structures

# Contents

## Articles

Introduction	<b>1</b>
Data structure	1
Linked data structure	3
Succinct data structure	6
Implicit data structure	8
Compressed data structure	9
Search data structure	10
Persistent data structure	11
Concurrent data structure	18
Abstract data types	<b>21</b>
Abstract data type	21
List	29
Stack	32
Queue	61
Deque	63
Priority queue	66
Map	70
Bidirectional map	73
Multimap	74
Set	75
Tree	80
Arrays	<b>85</b>
Array data structure	85
Row-major order	91
Dope vector	93
Iliffe vector	94
Dynamic array	95
Hashed array tree	98
Gap buffer	99
Circular buffer	101
Sparse array	111

Bit array	112
Bitboard	117
Parallel array	121
Lookup table	123
<b>Lists</b>	<b>129</b>
Linked list	129
XOR linked list	145
Unrolled linked list	147
VList	149
Skip list	151
Self-organizing list	157
<b>Binary trees</b>	<b>162</b>
Binary tree	162
Binary search tree	170
Self-balancing binary search tree	180
Tree rotation	182
Weight-balanced tree	185
Threaded binary tree	186
AVL tree	191
Red-black tree	195
AA tree	210
Scapegoat tree	215
Splay tree	219
T-tree	234
Rope	237
Top Trees	242
Tango Trees	246
Van Emde Boas tree	268
Cartesian tree	272
Treap	277
<b>B-trees</b>	<b>281</b>
B-tree	281
B+ tree	292
Dancing tree	297
2-3 tree	298

2-3-4 tree	299
Queaps	301
Fusion tree	305
Bx-tree	309
<b>Heaps</b>	<b>312</b>
Heap	312
Binary heap	315
Binomial heap	321
Fibonacci heap	326
2-3 heap	331
Pairing heap	331
Beap	334
Leftist tree	335
Skew heap	338
Soft heap	341
$d$ -ary heap	343
<b>Tries</b>	<b>346</b>
Trie	346
Radix tree	353
Suffix tree	358
Suffix array	363
Compressed suffix array	367
FM-index	368
Generalised suffix tree	371
B-trie	372
Judy array	372
Directed acyclic word graph	374
<b>Multiway trees</b>	<b>376</b>
Ternary search tree	376
And-or tree	379
(a,b)-tree	380
Link/cut tree	381
SPQR tree	381
Spaghetti stack	384
Disjoint-set data structure	385



<b>Space-partitioning trees</b>	<b>389</b>
Space partitioning	389
Binary space partitioning	390
Segment tree	395
Interval tree	399
Range tree	404
Bin	406
$k$ -d tree	408
Implicit $k$ -d tree	416
min/max kd-tree	419
Adaptive $k$ -d tree	420
Quadtree	421
Octree	427
Linear octrees	429
Z-order	429
UB-tree	434
R-tree	435
R+ tree	441
R* tree	442
Hilbert R-tree	445
X-tree	452
Metric tree	452
vP-tree	453
BK-tree	454
<b>Hashes</b>	<b>455</b>
Hash table	455
Hash function	468
Open addressing	476
Lazy deletion	479
Linear probing	479
Quadratic probing	480
Double hashing	484
Cuckoo hashing	486
Coalesced hashing	491
Perfect hash function	494
Universal hashing	496

Linear hashing	501
Extendible hashing	502
2-choice hashing	508
Pearson hashing	508
Fowler–Noll–Vo hash function	509
Bitstate hashing	511
Bloom filter	512
Locality preserving hashing	523
Morton number	524
Zobrist hashing	529
Rolling hash	530
Hash list	531
Hash tree	532
Prefix hash tree	534
Hash trie	535
Hash array mapped trie	535
Distributed hash table	536
Consistent hashing	542
Stable hashing	544
Koorde	544
<b>Graphs</b>	<b>547</b>
Graph	547
Adjacency list	549
Adjacency matrix	551
And-inverter graph	554
Binary decision diagram	556
Binary moment diagram	560
Zero-suppressed decision diagram	562
Propositional directed acyclic graph	563
Graph-structured stack	564
Scene graph	565
<b>Appendix</b>	<b>570</b>
Big O notation	570
Amortized analysis	581
Locality of reference	582
Standard Template Library	585

## References

Article Sources and Contributors	596
Image Sources, Licenses and Contributors	605

## Article Licenses

License	610
---------	-----

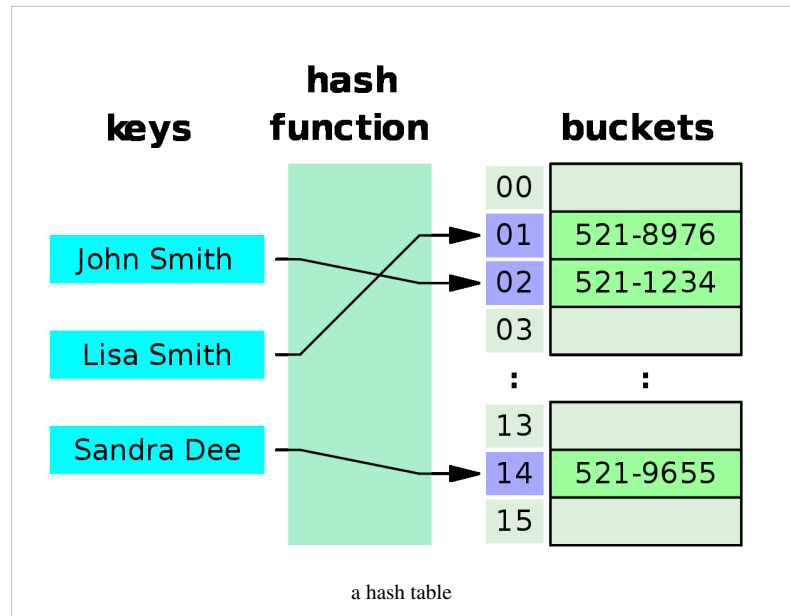
# Introduction

## Data structure

In computer science, a **data structure** is a particular way of storing and organizing data in a computer so that it can be used efficiently.<sup>[1][2]</sup>

Different kinds of data structures are suited to different kinds of applications, and some are highly specialized to specific tasks. For example, B-trees are particularly well-suited for implementation of databases, while compiler implementations usually use hash tables to look up identifiers.

Data structures provide a means to manage huge amounts of data efficiently, such as large databases and internet indexing services. Usually, efficient data structures are a key to designing efficient algorithms. Some formal design methods and programming languages emphasize data structures, rather than algorithms, as the key organizing factor in software design.



## Overview

- An array data structure stores a number of elements of the same type in a specific order. They are accessed using an integer to specify which element is required (although the elements may be of almost any type). Arrays may be fixed-length or expandable.
- Record (also called tuple or struct) Records are among the simplest data structures. A record is a value that contains other values, typically in fixed number and sequence and typically indexed by names. The elements of records are usually called *fields* or *members*.
- A hash or dictionary or map is a more flexible variation on a record, in which name-value pairs can be added and deleted freely.
- Union. A union type definition will specify which of a number of permitted primitive types may be stored in its instances, e.g. "float or long integer". Contrast with a record, which could be defined to contain a float *and* an integer; whereas, in a union, there is only one value at a time.
- A tagged union (also called a variant, variant record, discriminated union, or disjoint union) contains an additional field indicating its current type, for enhanced type safety.
- A set is an abstract data structure that can store specific values, without any particular order, and no repeated values. Values themselves are not retrieved from sets, rather one tests a value for membership to obtain a boolean "in" or "not in".
- An object contains a number of data fields, like a record, and also a number of program code fragments for accessing or modifying them. Data structures not containing code, like those above, are called plain old data

structure.

Many others are possible, but they tend to be further variations and compounds of the above.

## Basic principles

Data structures are generally based on the ability of a computer to fetch and store data at any place in its memory, specified by an address—a bit string that can be itself stored in memory and manipulated by the program. Thus the record and array data structures are based on computing the addresses of data items with arithmetic operations; while the linked data structures are based on storing addresses of data items within the structure itself. Many data structures use both principles, sometimes combined in non-trivial ways (as in XOR linking)

The implementation of a data structure usually requires writing a set of procedures that create and manipulate instances of that structure. The efficiency of a data structure cannot be analyzed separately from those operations. This observation motivates the theoretical concept of an abstract data type, a data structure that is defined indirectly by the operations that may be performed on it, and the mathematical properties of those operations (including their space and time cost).

## Language support

Most assembly languages and some low-level languages, such as BCPL(Basic Combined Programming Language), lack support for data structures. Many high-level programming languages, and some higher-level assembly languages, such as MASM, on the other hand, have special syntax or other built-in support for certain data structures, such as vectors (one-dimensional arrays) in the C language or multi-dimensional arrays in Pascal.

Most programming languages feature some sorts of library mechanism that allows data structure implementations to be reused by different programs. Modern languages usually come with standard libraries that implement the most common data structures. Examples are the C++ Standard Template Library, the Java Collections Framework, and Microsoft's .NET Framework.

Modern languages also generally support modular programming, the separation between the interface of a library module and its implementation. Some provide opaque data types that allow clients to hide implementation details. Object-oriented programming languages, such as C++, Java and .NET Framework may use classes for this purpose.

Many known data structures have concurrent versions that allow multiple computing threads to access the data structure simultaneously.

## References

- [1] Paul E. Black (ed.), entry for *data structure* in *Dictionary of Algorithms and Data Structures*. U.S. National Institute of Standards and Technology. 15 December 2004. Online version (<http://www.itl.nist.gov/div897/sqg/dads/HTML/datastructur.html>) Accessed May 21, 2009.
- [2] Entry *data structure* in the Encyclopædia Britannica (2009) Online entry (<http://www.britannica.com/EBchecked/topic/152190/data-structure>) accessed on May 21, 2009.

## Further reading

- Peter Brass, *Advanced Data Structures*, Cambridge University Press, 2008.
- Donald Knuth, *The Art of Computer Programming*, vol. 1. Addison-Wesley, 3rd edition, 1997.
- Dinesh Mehta and Sartaj Sahni *Handbook of Data Structures and Applications*, Chapman and Hall/CRC Press, 2007.
- Niklaus Wirth, *Algorithms and Data Structures*, Prentice Hall, 1985.

## External links

- UC Berkeley video course on data structures (<http://academicearth.org/courses/data-structures>)
- Descriptions (<http://nist.gov/dads/>) from the Dictionary of Algorithms and Data Structures
- CSE.unr.edu (<http://www.cse.unr.edu/~bebis/CS308/>)
- Data structures course with animations ([http://www.cs.auckland.ac.nz/software/AlgAnim/ds\\_ToC.html](http://www.cs.auckland.ac.nz/software/AlgAnim/ds_ToC.html))
- Data structure tutorials with animations (<http://courses.cs.vt.edu/~csonline/DataStructures/Lessons/index.html>)
- An Examination of Data Structures from .NET perspective ([http://msdn.microsoft.com/en-us/library/aa289148\(VS.71\).aspx](http://msdn.microsoft.com/en-us/library/aa289148(VS.71).aspx))
- Schaffer, C. *Data Structures and Algorithm Analysis* (<http://people.cs.vt.edu/~shaffer/Book/C++3e20110915.pdf>)

## Linked data structure

In computer science, a **linked data structure** is a data structure which consists of a set of data records (*nodes*) linked together and organized by references (*links* or *pointers*).

In linked data structures, the links are usually treated as special data types that can only be dereferenced or compared for equality. Linked data structures are thus contrasted with arrays and other data structures that require performing arithmetic operations on pointers. This distinction holds even when the nodes are actually implemented as elements of a single array, and the references are actually array indices: as long as no arithmetic is done on those indices, the data structure is essentially a linked one.

Linking can be done in two ways - Using dynamic allocation and using array index linking.

Linked data structures include linked lists, search trees, expression trees, and many other widely used data structures. They are also key building blocks for many efficient algorithms, such as topological sort<sup>[1]</sup> and set union-find.<sup>[2]</sup>

## Common types of linked data structures

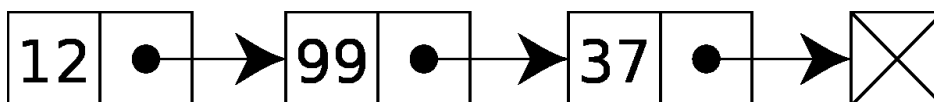
### Linked lists

A linked list is a collection of structures ordered not by their physical placement in memory but by logical links that are stored as part of the data in the structure itself. It is not necessary that it should be stored in the adjacent memory locations. Every structure has a data field and a address field. The Address field contains the address of its successor.

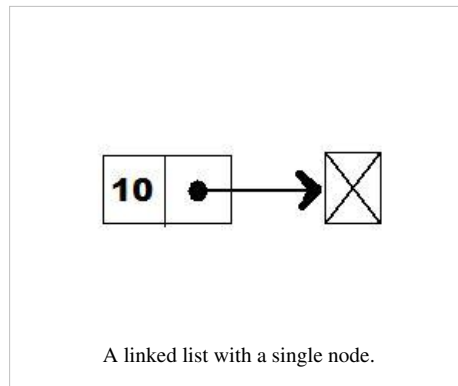
Linked list can be singly, doubly or multiply linked and can either be linear or circular.

#### Basic Properties

- Objects, called **nodes**, are linked in a linear sequence
- A reference to the first node of the list is always kept. This is called the 'head' or 'front'.<sup>[3]</sup>



A linked list with three nodes contain two fields each: an integer value and a link to the next node



### Example in Java

This is an example of the node class used to store integers in a Java implementation of a linked list.

```
public class IntNode {  
    public int value;  
    public IntNode link;  
    public IntNode(int v) { value = v; }  
}
```

### Example in C

This is an example of the node structure used for implementation of linked list in C.

```
struct node  
{  
    int val;  
    struct node *next;  
};
```

This is an example using typedefs.

```
typedef struct node_s node_t;  
  
struct node_s  
{  
    int val;  
    node_t *next;  
};
```

**Note:** A structure like this which contains a member that points to the same structure is called a self referential structure.

## Search trees

A search tree is a tree data structure in whose nodes data values can be stored from some ordered set, which is such that in an in-order traversal of the tree the nodes are visited in ascending order of the stored values.

### Basic Properties

- Objects, called nodes, are stored in an ordered set.
- In-order traversal provides an ascending readout of the data in the tree
- Sub trees of the tree are in themselves, trees.

## Advantages and disadvantages

### Advantages against arrays

Compared to arrays, linked data structures allow more flexibility in organizing the data and in allocating space for it. In arrays, the size of the array must be specified precisely at the beginning, this can be a potential waste of memory. A linked data structure is built dynamically and never needs to be bigger than the programmer requires. It also requires no guessing in terms of how much space you must allocate when using a linked data structure. This is a feature that is key in saving wasted memory.

In array, the array elements have to be in contiguous (connected and sequential) portion of memory. But in linked data structure, the reference to each node gives us the information where to find out the next one. The nodes of a linked data structure can also be moved individually to different locations without affecting the logical connections between them, unlike arrays. With due care, a process can add or delete nodes to one part of a data structure even while other processes are working on other parts.

On the other hand, access to any particular node in a linked data structure requires following a chain of references that stored in it. If the structure has  $n$  nodes, and each node contains at most  $b$  links, there will be some nodes that cannot be reached in less than  $\log_b n$  steps. For many structures, some nodes may require worst case up to  $n-1$  steps. In contrast, many array data structures allow access to any element with a constant number of operations, independent of the number of entries.

Broadly the implementation of these linked data structure is through dynamic data structures. It gives us the chance to use particular space again. Memory can be utilized more efficiently by using this data structures. Memory is allocated as per the need and when memory is not further needed, deallocation is done.

### General disadvantages

Linked data structures may also incur in substantial memory allocation overhead (if nodes are allocated individually) and frustrate memory paging and processor caching algorithms (since they generally have poor locality of reference). In some cases, linked data structures may also use more memory (for the link fields) than competing array structures. This is because linked data structures are not contiguous. Instances of data can be found all over in memory, unlike arrays.

In arrays,  $n$ th element can be accessed immediately, while in a linked data structure we have to follow multiple pointers so element access time varies according to where in the structure the element is.

In some theoretical models of computation that enforce the constraints of linked structures, such as the pointer machine, many problems require more steps than in the unconstrained random access machine model.



## References

- [1] Donald Knuth, The Art of Computer Programming
- [2] Bernard A. Galler and Michael J. Fischer. An improved equivalence algorithm. *Communications of the ACM*, Volume 7, Issue 5 (May 1964), pages 301-303. The paper originating disjoint-set forests. ACM Digital Library (<http://portal.acm.org/citation.cfm?doid=364099.364331>)
- [3] <http://www.cs.toronto.edu/~hojjat/148s07/lectures/week5/07linked.pdf>

## Succinct data structure

In computer science, a **succinct data structure** is data structure which uses an amount of space that is "close" to the information-theoretic lower bound, but (unlike other compressed representations) still allows for efficient query operations. The concept was originally introduced by Jacobson<sup>[1]</sup> to encode bit vectors, (unlabeled) trees, and planar graphs. Unlike general lossless data compression algorithms, succinct data structures retain the ability to use them in-place, without decompressing them first. A related notion is that of a compressed data structure, in which the size of the data structure depends upon the particular data being represented.

Suppose that  $Z$  is the information-theoretical optimal number of bits needed to store some data. A representation of this data is called

- *implicit* if it takes  $Z + O(1)$  bits of space,
- *succinct* if it takes  $Z + o(Z)$  bits of space, and
- *compact* if it takes  $O(Z)$  bits of space.

Implicit structures are thus usually reduced to storing information using some permutation of the input data; the most well-known example of this is the heap.

## Succinct dictionaries

Succinct indexable dictionaries, also called *rank/select* dictionaries, form the basis of a number of succinct representation techniques, including binary trees,  $k$ -ary trees and multisets,<sup>[2]</sup> as well as suffix trees and arrays.<sup>[3]</sup>

The basic problem is to store a subset  $S$  of a universe  $U = [0 \dots n] = \{0, 1, \dots, n-1\}$ , usually represented as a bit array  $B[0 \dots n]$  where  $B[i] = 1$  iff  $i \in S$ . An indexable dictionary supports the usual methods on dictionaries (queries, and insertions/deletions in the dynamic case) as well as the following operations:

- $\mathbf{rank}_q(x) = |\{k \in [0 \dots x] : B[k] = q\}|$
- $\mathbf{select}_q(x) = \min\{k \in [0 \dots n] : \mathbf{rank}_q(k) = x\}$

for  $q \in \{0, 1\}$ .

There is a simple representation<sup>[4]</sup> which uses  $n + o(n)$  bits of storage space (the original bit array and an  $o(n)$  auxiliary structure) and supports **rank** and **select** in constant time. It uses an idea similar to that for range-minimum queries; there are a constant number of recursions before stopping at a subproblem of a limited size. The bit array  $B$  is partitioned into *large blocks* of size  $l = \lg^2 n$  bits and *small blocks* of size  $s = \lg n / 2$  bits. For each large block, the rank of its first bit is stored in a separate table  $R_l[0 \dots n/l]$ ; each such entry takes  $\lg n$  bits for a total of  $(n/l) \lg n = n / \lg n$  bits of storage. Within a large block, another directory  $R_s[0 \dots l/s]$  stores the rank of each of the  $l/s = 2 \lg n$  small blocks it contains. The difference here is that it only needs  $\lg l = \lg \lg^2 n = 2 \lg \lg n$  bits for each entry, since only the differences from the rank of the first bit in the containing large block need to be stored. Thus, this table takes a total of  $(n/s) \lg l = 4n \lg \lg n / \lg n$  bits. A lookup table  $R_p$  can then be used that stores the answer to every possible rank query on a bit string of length  $s$  for  $i \in [0, s]$ ; this requires  $2^s s \lg s = O(\sqrt{n} \lg n \lg \lg n)$  bits of storage space. Thus, since each of these auxiliary tables take  $o(n)$  space, this data structure supports rank queries in  $O(1)$  time and  $n + o(n)$  bits of space. To answer a query for  $\mathbf{rank}_1(x)$  in constant time, a constant time algorithm computes

$$\mathbf{rank}_1(x) = R_l[\lfloor x/l \rfloor] + R_s[\lfloor x/s \rfloor] + R_p[x \lfloor x/s \rfloor, x \bmod s]$$

In practice, the lookup table  $R_p$  can be replaced by bitwise operations and smaller tables to perform find the number of bits set in the small blocks. This is often beneficial, since succinct data structures find their uses in large data sets, in which case cache misses become much more frequent and the chances of the lookup table being evicted from closer CPU caches becomes higher.<sup>[5]</sup> Select queries can be easily supported by doing a binary search on the same auxiliary structure used for **rank**; however, this takes  $O(\lg n)$  time in the worst case. A more complicated structure using  $3n/\lg \lg n + O(\sqrt{n} \lg n \lg \lg n) = o(n)$  bits of additional storage can be used to support **select** in constant time.<sup>[6]</sup> In practice, many of these solutions have hidden constants in the  $O(\cdot)$  notation which dominate before any asymptotic advantage becomes apparent; implementations using broadword operations and word-aligned blocks often perform better in practice.<sup>[7]</sup>

## Entropy-compressed dictionaries

The  $n + o(n)$  space approach can be improved by noting that there are  $\binom{n}{m}$  distinct  $m$ -subsets of  $[n]$  (or binary strings of length  $n$  with exactly  $m$  1's), and thus  $\mathcal{B}(m, n) = \lceil \lg \binom{n}{m} \rceil$  is an information theoretic lower bound on the number of bits needed to store  $\mathcal{B}$ . There is a succinct (static) dictionary which attains this bound, namely using  $\mathcal{B}(m, n) + o(\mathcal{B}(m, n))$  space.<sup>[8]</sup> This structure can be extended to support **rank** and **select** queries and takes  $\mathcal{B}(m, n) + O(m + n \lg \lg n / \lg n)$  space.<sup>[2]</sup> This bound can be reduced to a space/time tradeoff by reducing the storage space of the dictionary to  $\mathcal{B}(m, n) + O(nt^t / \lg^t n + n^{3/4})$  with queries taking  $O(t)$  time.<sup>[9]</sup>

## Examples

When a sequence of variable-length items needs to be encoded, the items can simply be placed one after another, with no delimiters. A separate binary string consisting of 1s in the positions where an item begins, and 0s everywhere else is encoded along with it. Given this string, the *select* function can quickly determine where each item begins, given its index.<sup>[10]</sup>

Another example is the representation of a binary tree: an arbitrary binary tree on  $n$  nodes can be represented in  $2n + o(n)$  bits while supporting a variety of operations on any node, which includes finding its parent, its left and right child, and returning the size of its subtree, each in constant time. The number of different binary trees on  $n$  nodes is  $\binom{2n}{n} / (n + 1)$ . For large  $n$ , this is about  $4^n$ ; thus we need at least about  $\log_2(4^n) = 2n$  bits to encode it. A succinct binary tree therefore would occupy only 2 bits per node.

## References

- [1] Jacobson, G. J (1988). *Succinct static data structures*.
- [2] Raman, R.; V. Raman, S. S Rao (2002). "Succinct indexable dictionaries with applications to encoding k-ary trees and multisets" (<http://www.cs.cmu.edu/afs/cs.cmu.edu/project/aladdin/wwwlocal/hash/RaRaRa02.pdf>). *Proceedings of the thirteenth annual ACM-SIAM symposium on Discrete algorithms*. pp. 233–242. ISBN 0-89871-513-X. .
- [3] Sadakane, K.; R. Grossi (2006). "Squeezing succinct data structures into entropy bounds" ([http://www.dmi.unisa.it/people/cerulli/www/WSPages/WSFiles/Abs/S3/S33\\_abs\\_Grossi.pdf](http://www.dmi.unisa.it/people/cerulli/www/WSPages/WSFiles/Abs/S3/S33_abs_Grossi.pdf)). *Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithm*. pp. 1230–1239. ISBN 0-89871-605-5. .
- [4] Jacobson, G. (1989). *Space-efficient static trees and graphs* (<http://www.cs.cmu.edu/afs/cs/project/aladdin/wwwlocal/compression/00063533.pdf>). .
- [5] González, R.; S. Grabowski, V. Mäkinen, G. Navarro (2005). "Practical implementation of rank and select queries" (<http://www.dcc.uchile.cl/~gnavarro/algoritmos/ps/wea05.pdf>). *Poster Proceedings Volume of 4th Workshop on Efficient and Experimental Algorithms (WEA)*. pp. 27–38. .
- [6] Clark, D. (1998). *Compact pat trees* (<https://uwspace.uwaterloo.ca/bitstream/10012/64/1/nq21335.pdf>). .
- [7] Vigna, S. (2008). "Broadword implementation of rank/select queries" (<http://sux.dsi.unimi.it/paper.pdf>). *Experimental Algorithms*: 154–168. .
- [8] Brodnik, A.; J. I Munro (1999). "Membership in constant time and almost-minimum space" (<http://www.cs.cmu.edu/afs/cs.cmu.edu/project/aladdin/wwwlocal/compression/BM99.pdf>). *SIAM J. Comput.* **28** (5): 1627–1640. doi:10.1137/S0097539795294165. .

- [9] Pătraşcu, M. (2008). "Succincter" (<http://people.csail.mit.edu/mip/papers/succinct/succinct.pdf>). *Foundations of Computer Science, 2008. FOCS'08. IEEE 49th Annual IEEE Symposium on.* pp. 305–313. .
- [10] Belazzougui, Djamel. "Hash, displace, and compress" (<http://cmph.sourceforge.net/papers/esa09.pdf>). .

## Implicit data structure

---

In computer science, an **implicit data structure** is a data structure that uses very little memory besides the actual data elements i.e. very little information other than main data is stored in these structures. These are storage schemes which retain no pointers and represent the file of  $n$   $k$ -key records as a simple  $n$  by  $k$  array  $n$  thus retrieve faster. In implicit data structures the only structural information to be given is to allow the array to grow and shrink as  $n$ . No extra information is required. It is called "implicit" because most of the structure of the elements is expressed implicitly by their order. Another term used interchangeably is **space efficient**. Definitions of "very little" are vague and can mean from  $O(1)$  to  $O(\log n)$  extra space. Everything is accessed in-place, by reading bits at various positions in the data. To achieve optimal coding, we use bits instead of bytes. Implicit data structures are frequently also succinct data structures.

Although one may argue that disk space is no longer a problem and we should not concern ourselves with improving space utilization, the issue that implicit data structures are designed to improve is main memory utilization. Hard disks, or any other means of large data capacity, I/O devices, are orders of magnitudes slower than main memory. Hence, the higher percentage of a task can fit in buffers in main memory the less dependence is on slow I/O devices. Hence, if a larger chunk of an implicit data structure fits in main memory the operations performed on it can be faster even if the asymptotic running time is not as good as its space-oblivious counterpart. Furthermore, since the CPU-cache is usually much smaller than main-memory, implicit data structures can improve cache-efficiency and thus running speed, especially if the method used improves locality. Keys are scanned very efficiently. Downloading indexed data in mobiles becomes easier.

### Implicit data structure for weighted element

For presentation of elements with different weight several data structures are required. The structure uses one more location besides required for values of elements. The first structure supports worst case search time in terms of rank of weight of elements w.r.t multi-set of weights. If the elements are drawn from uniform distribution, then variation of this structure takes average time. The same result obtain for the data structures in which the intervals between consecutive values have access probabilities.

### Examples

Examples of implicit data structures include

- Binary heap
- Beap

### Further reading

- See publications of Hervé Brönnimann <sup>[1]</sup>, J. Ian Munro <sup>[2]</sup>, Greg Frederickson <sup>[3]</sup>

### References

- [1] <http://photon.poly.edu/~hbr/>
- [2] <http://www.cs.uwaterloo.ca/~imunro/>
- [3] <http://www.cs.purdue.edu/people/gnf>
-

# Compressed data structure

---

The term **compressed data structure** arises in the computer science subfields of algorithms, data structures, and theoretical computer science. It refers to a data structure whose operations are roughly as fast as those of a conventional data structure for the problem, but whose size can be substantially smaller. The size of the compressed data structure is typically highly dependent upon the entropy of the data being represented.

Important examples of compressed data structures include the compressed suffix array<sup>[1][2]</sup> and the FM-index,<sup>[3]</sup> both of which can represent an arbitrary text of characters  $T$  for pattern matching. Given any input pattern  $P$ , they support the operation of finding if and where  $P$  appears in  $T$ . The search time is proportional to the sum of the length of pattern  $P$ , a very slow-growing function of the length of the text  $T$ , and the number of reported matches. The space they occupy is roughly equal to the size of the text  $T$  in entropy-compressed form, such as that obtained by Prediction by Partial Matching or gzip. Moreover, both data structures are self-indexing, in that they can reconstruct the text  $T$  in a random access manner, and thus the underlying text  $T$  can be discarded. In other words, they simultaneously provide a compressed and quickly searchable representation of the text  $T$ . They represent a substantial space improvement over the conventional suffix tree and suffix array, which occupy many times more space than the size of  $T$ . They also support searching for arbitrary patterns, as opposed to the inverted index, which can support only word-based searches. In addition, inverted indexes do not have the self-indexing feature.

An important related notion is that of a succinct data structure, which uses space roughly equal to the information-theoretic minimum, which is a worst-case notion of the space needed to represent the data. In contrast, the size of a compressed data structure depends upon the particular data being represented. When the data are compressible, as is often the case in practice for natural language text, the compressed data structure can occupy substantially less space than the information-theoretic minimum.

## References

- [1] R. Grossi and J. S. Vitter, Compressed Suffix Arrays and Suffix Trees with Applications to Text Indexing and String Matching], *Proceedings of the 32nd ACM Symposium on Theory of Computing*, May 2000, 397-406. Journal version in *SIAM Journal on Computing*, 35(2), 2005, 378-407.
- [2] R. Grossi, A. Gupta, and J. S. Vitter, High-Order Entropy-Compressed Text Indexes, *Proceedings of the 14th Annual SIAM/ACM Symposium on Discrete Algorithms*, January 2003, 841-850.
- [3] P. Ferragina and G. Manzini, Opportunistic Data Structures with Applications, *Proceedings of the 41st IEEE Symposium on Foundations of Computer Science*, November 2000, 390-398. Journal version in Indexing Compressed Text, *Journal of the ACM*, 52(4), 2005, 552-581.

# Search data structure

---

In computer science, a **search data structure** is any data structure that allows the efficient retrieval of specific items from a set of items, such as a specific record from a database.

The simplest, most general, and least efficient search structure is merely an unordered sequential list of all the items. Locating the desired item in such a list, by the linear search method, inevitably requires a number of operations proportional to the number  $n$  of items, in the worst case as well as in the average case. Useful search data structures allow faster retrieval; however, they are limited to queries of some specific kind. Moreover, since the cost of building such structures is at least proportional to  $n$ , they only pay off if several queries are to be performed on the same database (or on a database that changes little between queries).

**Static** search structures are designed for answering many queries on a fixed database; **dynamic** structures also allow insertion, deletion, or modification of items between successive queries. In the dynamic case, one must also consider the cost of fixing the search structure to account for the changes in the database.

## Classification

The simplest kind of query is to locate a record that has a specific field (the *key*) equal to a specified value  $v$ . Other common kinds of query are "find the item with smallest (or largest) key value", "find the item with largest key value not exceeding  $v$ ", "find all items with key values between specified bounds  $v_{\min}$  and  $v_{\max}$ ".

In certain databases the key values may be points in some multi-dimensional space. For example, the key may be a geographic position (latitude and longitude) on the Earth. In that case, common kinds of queries are *find the record with a key closest to a given point  $v$* , or *find all items whose key lies at a given distance from  $v$* , or *find all items within a specified region  $R$  of the space*.

A common special case of the latter are simultaneous range queries on two or more simple keys, such as "find all employee records with salary between 50,000 and 100,000 and hired between 1995 and 2007".

## Single ordered keys

- Array if the key values span a moderately compact interval.
- Priority-sorted list; see linear search
- Key-sorted array; see binary search
- Self-balancing binary search tree
- Hash table

## Finding the smallest element

- Heap

## Asymptotic amortized worst-case analysis

In this table, the asymptotic notation  $O(f(n))$  means "not exceeding some fixed multiple of  $f(n)$  in the worst case."

---

	Insert	Delete	Search	Find maximum	Space usage
Unsorted array	$O(1)$	$O(1)$	$O(n)$	$O(n)$	$O(n)$
Value-indexed array	$O(1)$	$O(1)$	$O(1)$	$O(n)$	$O(n)$
Sorted array	$O(n)$	$O(n)$	$O(\log n)$	$O(1)$	$O(n)$
Unsorted linked list	$O(1)^*$	$O(1)^*$	$O(n)$	$O(n)$	$O(n)$
Sorted linked list	$O(n)^*$	$O(1)^*$	$O(n)$	$O(1)$	$O(n)$
Self-balancing binary tree	$O(\log n)$	$O(\log n)$	$O(\log n)$	$O(\log n)$	$O(n)$
Heap	$O(\log n)$	$O(\log n)^{**}$	$O(n)$	$O(1)$	$O(n)$
Hash table	$O(1)$	$O(1)$	$O(1)$	$O(n)$	$O(n)$

\* The cost to add or delete an element into a known location in the list (i.e. if you have an iterator to the location) is  $O(1)$ . If you don't know the location, then you need to traverse the list to the location of deletion/insertion, which takes  $O(n)$  time. \*\* The deletion cost is  $O(\log n)$  for the minimum or maximum,  $O(n)$  for an arbitrary element.

This table is only an approximate summary; for each data structure there are special situations and variants that may lead to different costs. Also two or more data structures can be combined to obtain lower costs.

## Footnotes

# Persistent data structure

In computing, a **persistent data structure** is a data structure that always preserves the previous version of itself when it is modified. Such data structures are effectively immutable, as their operations do not (visibly) update the structure in-place, but instead always yield a new updated structure. (A persistent data structure is *not* a data structure committed to persistent storage, such as a disk; this is a different and unrelated sense of the word "persistent.")

A data structure is partially persistent if all versions can be accessed but only the newest version can be modified. The data structure is fully persistent if every version can be both accessed and modified. If there is also a meld or merge operation that can create a new version from two previous versions, the data structure is called confluent persistent. Structures that are not persistent are called ephemeral.<sup>[1]</sup>

These types of data structures are particularly common in logical and functional programming, and in a purely functional program all data is immutable, so all data structures are automatically fully persistent.<sup>[1]</sup> Persistent data structures can also be created using in-place updating of data and these may, in general, use less time or storage space than their purely functional counterparts.

While persistence can be achieved by simple copying, this is inefficient in time and space, because most operations make only small changes to a data structure. A better method is to exploit the similarity between the new and old versions to share structure between them, such as using the same subtree in a number of tree structures. However, because it rapidly becomes infeasible to determine how many previous versions share which parts of the structure, and because it is often desirable to discard old versions, this necessitates an environment with garbage collection.

## Partially persistent

In partially persistence model, we may query any previous version of the data structure, but we may only update the latest version. This implies a linear ordering among the versions.

*Three methods on balanced binary search tree:*

### Fat Node

Fat node method is to record all changes made to node fields in the nodes themselves, without erasing old values of the fields. This requires that we allow nodes to become arbitrarily “fat”. In other words, each fat node contains the same information and pointer fields as an ephemeral node, along with space for an arbitrary number of extra field values. Each extra field value has an associated field name and a version stamp which indicates the version in which the named field was changed to have the specified value. Besides, each fat node has its own version stamp, indicating the version in which the node was created. The only purpose of nodes having version stamps is to make sure that each node only contains one value per field name per version. In order to navigate through the structure, each original field value in a node has a version stamp of zero.

### Complexity of Fat Node

With using fat node method, it requires  $O(1)$  space for every modification: just store the new data. Each modification takes  $O(1)$  additional time to store the modification at the end of the modification history. This is an amortized time bound, assuming we store the modification history in a growable array. For access time, we must find the right version at each node as we traverse the structure. If we made  $m$  modifications, then each access operation has  $O(\log m)$  slowdown, and a query in our region computation will take  $O(\log^2 n)$  time. Since we have arranged the modification history, using time stamp, so it takes  $O(\log m)$  time to find the last modification before an arbitrary time stamp.

### Path Copying

Path copy is to make a copy of all nodes on the path which contains the node we about to insert or delete. Then you have to cascade the change back through the data structure: all nodes that pointed to the old node must be modified to point to the new node instead. These modifications cause more cascading changes, and so on, until we reach to the root. Maintain an array of roots indexed by timestamp. The data structure pointed to by time  $t$ 's root is exactly time  $t$ 's data structure.

### Complexity of Path Copying

With  $m$  modifications, this costs  $O(\log m)$  additive lookup time. Modification time and space are bounded by the size of the structure, since a single modification may cause the entire structure to be copied. That's  $O(m)$  for one update, and thus  $O(n^2)$  preprocessing time.

### A combination

Sleator, Tarjan et al. came up with a way to combine the advantages of fat nodes and path copying, getting  $O(1)$  access slowdown and  $O(1)$  modification space and time.

In each node, we store one modification box. This box can hold one modification to the node—either a modification to one of the pointers, or to the node's key, or to some other piece of node-specific data—and a timestamp for when that modification was applied. Initially, every node's modification box is empty.

Whenever we access a node, we check the modification box, and compare its timestamp against the access time. (The access time specifies the version of the data structure that we care about.) If the modification box is empty, or the access time is before the modification time, then we ignore the modification box and just deal with the normal

part of the node. On the other hand, if the access time is after the modification time, then we use the value in the modification box, overriding that value in the node. (Say the modification box has a new left pointer. Then we'll use it instead of the normal left pointer, but we'll still use the normal right pointer.)

Modifying a node works like this. (We assume that each modification touches one pointer or similar field.) If the node's modification box is empty, then we fill it with the modification. Otherwise, the modification box is full. We make a copy of the node, but using only the latest values. (That is, we overwrite one of the node's fields with the value that was stored in the modification box.) Then we perform the modification directly on the new node, without using the modification box. (We overwrite one of the new node's fields, and its modification box stays empty.) Finally, we cascade this change to the node's parent, just like path copying. (This may involve filling the parent's modification box, or making a copy of the parent recursively. If the node has no parent—it's the root—we add the new root to a sorted array of roots.)

With this algorithm, given any time  $t$ , at most one modification box exists in the data structure with time  $t$ . Thus, a modification at time  $t$  splits the tree into three parts: one part contains the data from before time  $t$ , one part contains the data from after time  $t$ , and one part was unaffected by the modification.

### Complexity of the combination

Time and space for modifications require amortized analysis. A modification takes  $O(1)$  amortized space, and  $O(1)$  amortized time. To see why, use a potential function  $\phi$ , where  $\phi(T)$  is the number of full live nodes in  $T$ . The live nodes of  $T$  are just the nodes that are reachable from the current root at the current time (that is, after the last modification). The full live nodes are the live nodes whose modification boxes are full.

Each modification involves some number of copies, say  $k$ , followed by 1 change to a modification box. (Well, not quite—you could add a new root—but that doesn't change the argument.) Consider each of the  $k$  copies. Each costs  $O(1)$  space and time, but decreases the potential function by one. (First, the node we copy must be full and live, so it contributes to the potential function. The potential function will only drop, however, if the old node isn't reachable in the new tree. But we know it isn't reachable in the new tree—the next step in the algorithm will be to modify the node's parent to point at the copy. Finally, we know the copy's modification box is empty. Thus, we've replaced a full live node with an empty live node, and  $\phi$  goes down by one.) The final step fills a modification box, which costs  $O(1)$  time and increases  $\phi$  by one.

Putting it all together, the change in  $\phi$  is  $\Delta\phi = 1 - k$ . Thus, we've paid  $O(k + \Delta\phi) = O(1)$  space and  $O(k + \Delta\phi + 1) = O(1)$  time.

## Fully persistent

In fully persistent model, both updates and queries are allowed on any version of the data structure.

## Confluently persistent

In confluently persistent model, we use combinations to combine input of more than one previous versions to output a new single version. Rather than a branching tree, combinations of versions induce a DAG (direct acyclic graph) structure on the version graph.

## Examples of persistent data structures

Perhaps the simplest persistent data structure is the singly linked list or *cons*-based list, a simple list of objects formed by each carrying a reference to the next in the list. This is persistent because we can take a *tail* of the list, meaning the last  $k$  items for some  $k$ , and add new nodes on to the front of it. The tail will not be duplicated, instead becoming shared between both the old list and the new list. So long as the contents of the tail are immutable, this sharing will be invisible to the program.



Many common reference-based data structures, such as red-black trees,<sup>[2]</sup> and stacks,<sup>[3]</sup> can easily be adapted to create a persistent version. Some others need slightly more effort, for example: Queue, Double-ended queues (dequeue), Min-Dequeue (which have additional operation min returning minimal element in constant time without incurring additional complexity on standard operations of queuing and dequeuing on both ends), Random access list (with constant cons/head as single linked list, but with additional operation of random access with sub-linear, most often logarithmic, complexity), Random access queue, Random access double-ended queue and Random access stack (as well Random access Min-List, Min-Queue, Min-Dequeue, Min-Stack).

There also exist persistent data structures which use destructible operations, making them impossible to implement efficiently in purely functional languages (like Haskell), but possible in languages like C or Java. These types of data structures can be avoided with proper design. One primary advantage to using purely persistent data structures is that they often behave better in multi-threaded environments.

## Linked lists

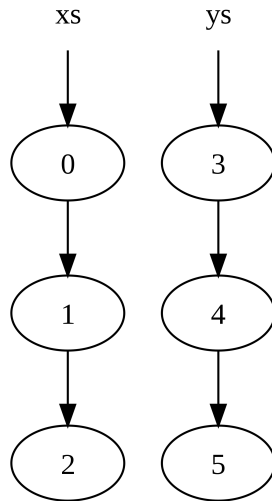
*This example is taken from Okasaki. See the bibliography.*

Singly linked lists are the bread-and-butter data structure in functional languages. In ML-derived languages and Haskell, they are purely functional because once a node in the list has been allocated, it cannot be modified, only copied or destroyed. Note that ML itself is **not** purely functional.

Consider the two lists:

```
xs = [0, 1, 2]
ys = [3, 4, 5]
```

These would be represented in memory by:

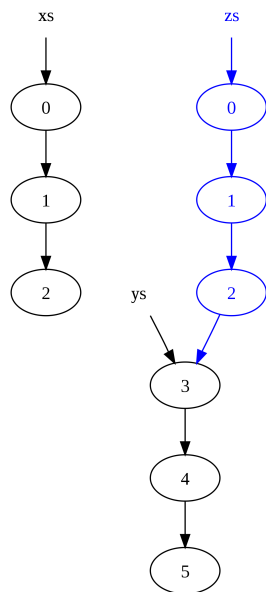


where a circle indicates a node in the list (the arrow out representing the second element of the node which is a pointer to another node).

Now concatenating the two lists:

```
zs = xs ++ ys
```

results in the following memory structure:



Notice that the nodes in list `xs` have been copied, but the nodes in `ys` are shared. As a result, the original lists (`xs` and `ys`) persist and have not been modified.

The reason for the copy is that the last node in `xs` (the node containing the original value 2) cannot be modified to point to the start of `ys`, because that would change the value of `xs`.

## Trees

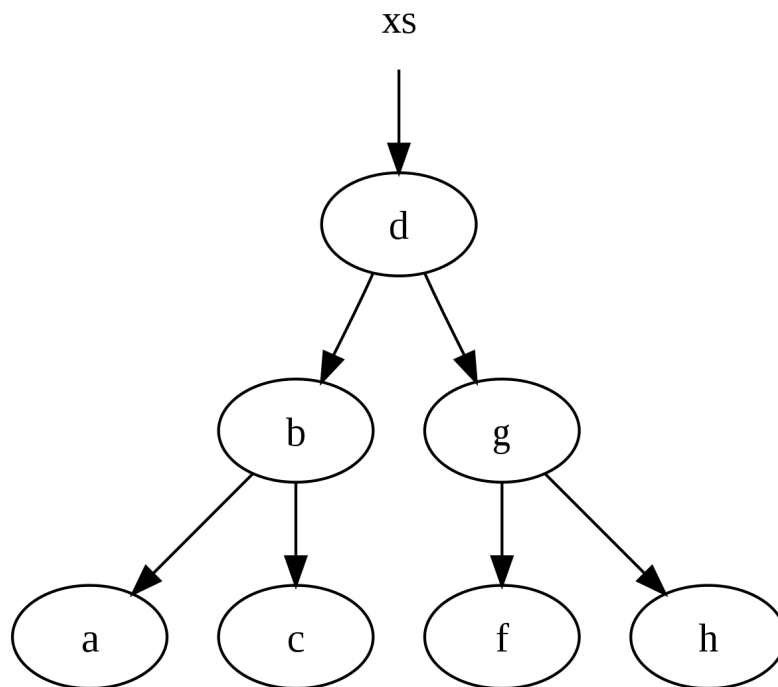
*This example is taken from Okasaki. See the bibliography.*

Consider a binary tree used for fast searching, where every node has the recursive invariant that subnodes on the left are less than the node, and subnodes on the right are greater than the node.

For instance, the set of data

```
xs = [a, b, c, d, f, g, h]
```

might be represented by the following binary search tree:



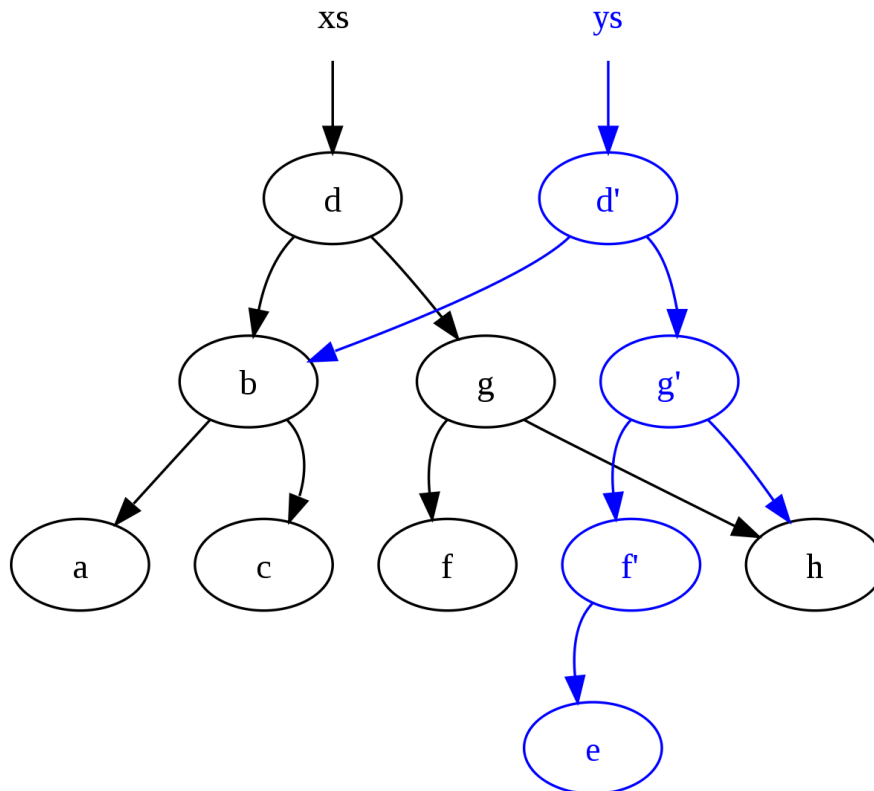
A function which inserts data into the binary tree and maintains the invariant is:

```
fun insert (x, E) = T (E, x, E)
  | insert (x, s as T (a, y, b)) =
    if x < y then T (insert (x, a), y, b)
    else if x > y then T (a, y, insert (x, b))
    else s
```

After executing

```
ys = insert ("e", xs)
```

we end up with the following:



Notice two points: Firstly the original tree ( $xs$ ) persists. Secondly many common nodes are shared between the old tree and the new tree. Such persistence and sharing is difficult to manage without some form of garbage collection (GC) to automatically free up nodes which have no live references, and this is why GC is a feature commonly found in functional programming languages.

## Reference cycles

Since every value in a purely functional computation is built up out of existing values, it would seem that it is impossible to create a cycle of references. In that case, the reference graph (the graph of the references from object to object) could only be a directed acyclic graph. However, in most functional languages, functions can be defined recursively; this capability allows recursive structures using functional suspensions. In lazy languages, such as Haskell, all data structures are represented as implicitly suspended thunks; in these languages any data structure can be recursive because a value can be defined in terms of itself. Some other languages, such as OCaml, allow the explicit definition of recursive values.

## References

- [1] Kaplan, Haim (2001). "Persistent data structures" (<http://www.math.tau.ac.il/~haimk/papers/persistent-survey.ps>). *Handbook on Data Structures and Applications* (CRC Press). .
- [2] Neil Sarnak, Robert E. Tarjan (1986). "Planar Point Location Using Persistent Search Trees" (<http://www.link.cs.cmu.edu/15859-f07/papers/point-location.pdf>). *Communications of the ACM* **29** (7): 669–679. doi:10.1145/6138.6151. .
- [3] Chris Okasaki. *Purely Functional Data Structures (thesis)* (<http://www.cs.cmu.edu/~rwh/theses/okasaki.pdf>). .

## Further reading

- Persistent Data Structures and Managed References (<http://www.infoq.com/presentations/Value-Identity-State-Rich-Hickey>) - video presentation by Rich Hickey on Clojure's use of persistent data structures and how they support concurrency

- Making Data Structures Persistent (<http://www.cs.cmu.edu/~sleator/papers/Persistence.htm>) by James R. Driscoll, Neil Sarnak, Daniel D. Sleator, Robert E. Tarjan
- Fully persistent arrays for efficient incremental updates and voluminous reads (<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.34.1317>)
- Real-Time Deques, Multihead Turing Machines, and Purely Functional Programming (<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.51.2895>)
- *Purely functional data structures* by Chris Okasaki, Cambridge University Press, 1998, ISBN 0-521-66350-4.
- Purely Functional Data Structures (<http://www.cs.cmu.edu/~rwh/theses/okasaki.pdf>) thesis by Chris Okasaki (PDF format)
- Fully Persistent Lists with Catenation (<http://www.cs.cmu.edu/~sleator/papers/fully-persistent-lists.pdf>) by James R. Driscoll, Daniel D. Sleator, Robert E. Tarjan (PDF)
- Persistent Data Structures (<http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-854j-advanced-algorithms-fall-2005/lecture-notes/persistent.pdf>) from MIT open course Advanced Algorithms (<http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-854j-advanced-algorithms-fall-2005/>)

## External links

- Lightweight Java implementation of Persistent Red-Black Trees (<http://wiki.edinburghhacklab.com/PersistentRedBlackTreeSet>)

# Concurrent data structure

---

In computer science, a **concurrent data structure** is a particular way of storing and organizing data for access by multiple computing threads (or processes) on a computer.

Historically, such data structures were used on uniprocessor machines with operating systems that supported multiple computing threads (or processes). The term concurrency captured the multiplexing/interleaving of the threads' operations on the data by the operating system, even though the processors never issued two operations that accessed the data simultaneously.

Today, as multiprocessor computer architectures that provide parallelism become the dominant computing platform (through the proliferation of multi-core processors), the term has come to stand mainly for data structures that can be accessed by multiple threads which may actually access the data simultaneously because they run on different processors that communicate with one another. The concurrent data structure (sometimes also called a *shared data structure*) is usually considered to reside in an abstract storage environment called shared memory, though this memory may be physically implemented as either a "tightly coupled" or a distributed collection of storage modules.

## Basic principles

Concurrent data structures, intended for use in parallel or distributed computing environments, differ from "sequential" data structures, intended for use on a processor machine, in several ways.<sup>[1]</sup> Most notably, in a sequential environment one specifies the data structure's properties and checks that they are implemented correctly, by providing **safety properties**. In a concurrent environment, the specification must also describe **liveness properties** which an implementation must provide. Safety properties usually state that something bad never happens, while liveness properties state that something good keeps happening. These properties can be expressed, for example, using Linear Temporal Logic.

The type of liveness requirements tend to define the data structure. The method calls can be blocking or non-blocking. Data structures are not restricted to one type or the other, and can allow combinations where some

---

method calls are blocking and others are non-blocking (examples can be found in the Java concurrency software library).

The safety properties of concurrent data structures must capture their behavior given the many possible interleavings of methods called by different threads. It is quite intuitive to specify how abstract data structures behave in a sequential setting in which there are no interleavings. Therefore, many mainstream approaches for arguing the safety properties of a concurrent data structure (such as serializability, linearizability, sequential consistency, and quiescent consistency<sup>[1]</sup>) specify the structures properties sequentially, and map its concurrent executions to a collection of sequential ones.

In order to guarantee the safety and liveness properties, concurrent data structures must typically (though not always) allow threads to reach consensus as to the results of their simultaneous data access and modification requests. To support such agreement, concurrent data structures are implemented using special primitive synchronization operations (see synchronization primitives) available on modern multiprocessor machines that allow multiple threads to reach consensus. This consensus can be reached achieved in a blocking manner by using locks, or without locks, in which case it is non-blocking. There is a wide body of theory on the design of concurrent data structures (see bibliographical references).

## Design and Implementation

Concurrent data structures are significantly more difficult to design and to verify as being correct than their sequential counterparts.

The primary source of this additional difficulty is concurrency, exacerbated by the fact that threads must be thought of as being completely asynchronous: they are subject to operating system preemption, page faults, interrupts, and so on.

On today's machines, the layout of processors and memory, the layout of data in memory, the communication load on the various elements of the multiprocessor architecture all influence performance. Furthermore, there is a tension between correctness and performance: algorithmic enhancements that seek to improve performance often make it more difficult to design and verify a correct data structure implementation.

A key measure for performance is scalability, captured by the speedup of the implementation. Speedup is a measure of how effectively the application is utilizing the machine it is running on. On a machine with  $P$  processors, the speedup is the ratio of the structures execution time on a single processor to its execution time on  $T$  processors. Ideally, we want linear speedup: we would like to achieve a speedup of  $P$  when using  $P$  processors. Data structures whose speedup grows with  $P$  are called **scalable**. The extent to which one can scale the performance of a concurrent data structure is captured by a formula known as Amdahl's law and more refined versions of it such as Gustafson's law.

A key issue with the performance of concurrent data structures is the level of memory contention: the overhead in traffic to and from memory as a result of multiple threads concurrently attempting to access the same locations in memory. This issue is most acute with blocking implementations in which locks control access to memory. In order to acquire a lock, a thread must repeatedly attempt to modify that location. On a cache-coherent multiprocessor (one in which processors have local caches that are updated by hardware in order to keep them consistent with the latest values stored) this results in long waiting times for each attempt to modify the location, and is exacerbated by the additional memory traffic associated with unsuccessful attempts to acquire the lock.

## References

- [1] Mark Moir and Nir Shavit (2007). " *Concurrent Data Structures* (<http://www.cs.tau.ac.il/~shanir/concurrent-data-structures.pdf>)". In Dinesh Metha and Sartaj Sahni. *'Handbook of Data Structures and Applications'* (1st ed.). Chapman and Hall/CRC Press. pp. 47-14 — 47-30.

## Further reading

- Nancy Lynch "Distributed Computing"
- Hagit Attiya and Jennifer Welch "Distributed Computing: Fundamentals, Simulations And Advanced Topics, 2nd Ed"
- Doug Lea, "Concurrent Programming in Java: Design Principles and Patterns"
- Maurice Herlihy and Nir Shavit, "The Art of Multiprocessor Programming"
- Mattson, Sanders, and Massingil "Patterns for Parallel Programming"

## External links

- Multithreaded data structures for parallel computing, Part 1 ([http://www.ibm.com/developerworks/aix/library/au-multithreaded\\_structures1/index.html](http://www.ibm.com/developerworks/aix/library/au-multithreaded_structures1/index.html)) (Designing concurrent data structures) by Arpan Sen
- Multithreaded data structures for parallel computing: Part 2 ([http://www.ibm.com/developerworks/aix/library/au-multithreaded\\_structures2/index.html](http://www.ibm.com/developerworks/aix/library/au-multithreaded_structures2/index.html)) (Designing concurrent data structures without mutexes) by Arpan Sen
- libcds (<http://libcds.sourceforge.net/>) - C++ library of lock-free containers and safe memory reclamation schema

---

# Abstract data types

---

## Abstract data type

---

In computer science, an **abstract data type** (ADT) is a mathematical model for a certain class of data structures that have similar behavior; or for certain data types of one or more programming languages that have similar semantics. An abstract data type is defined indirectly, only by the operations that may be performed on it and by mathematical constraints on the effects (and possibly cost) of those operations.<sup>[1]</sup>

For example, an abstract stack data structure could be defined by three operations: `push`, that inserts some data item onto the structure, `pop`, that extracts an item from it (with the constraint that each `pop` always returns the most recently pushed item that has not been popped yet), and `peek`, that allows data on top of the structure to be examined without removal. When analyzing the efficiency of algorithms that use stacks, one may also specify that all operations take the same time no matter how many items have been pushed into the stack, and that the stack uses a constant amount of storage for each element.

Abstract data types are purely theoretical entities, used (among other things) to simplify the description of abstract algorithms, to classify and evaluate data structures, and to formally describe the type systems of programming languages. However, an ADT may be implemented by specific data types or data structures, in many ways and in many programming languages; or described in a formal specification language. ADTs are often implemented as modules: the module's interface declares procedures that correspond to the ADT operations, sometimes with comments that describe the constraints. This information hiding strategy allows the implementation of the module to be changed without disturbing the client programs.

The term **abstract data type** can also be regarded as a generalised approach of a number of algebraic structures, such as lattices, groups, and rings.<sup>[2]</sup> This can be treated as part of subject area of artificial intelligence. The notion of abstract data types is related to the concept of data abstraction, important in object-oriented programming and design by contract methodologies for software development .

### Defining an abstract data type (ADT)

An abstract data type is defined as a mathematical model of the data objects that make up a data type as well as the functions that operate on these objects. There are no standard conventions for defining them. A broad division may be drawn between "imperative" and "functional" definition styles.

### Imperative abstract data type definitions

In the "imperative" view, which is closer to the philosophy of imperative programming languages, an abstract data structure is conceived as an entity that is *mutable* — meaning that it may be in different *states* at different times. Some operations may change the state of the ADT; therefore, the order in which operations are evaluated is important, and the same operation on the same entities may have different effects if executed at different times — just like the instructions of a computer, or the commands and procedures of an imperative language. To underscore this view, it is customary to say that the operations are *executed* or *applied*, rather than *evaluated*. The imperative style is often used when describing abstract algorithms. This is described by Donald E. Knuth and can be referenced from here [The Art of Computer Programming](#).

---



### Abstract variable

Imperative ADT definitions often depend on the concept of an *abstract variable*, which may be regarded as the simplest non-trivial ADT. An abstract variable  $V$  is a mutable entity that admits two operations:

- $\text{store}(V, x)$  where  $x$  is a *value* of unspecified nature; and
- $\text{fetch}(V)$ , that yields a value;

with the constraint that

- $\text{fetch}(V)$  always returns the value  $x$  used in the most recent  $\text{store}(V, x)$  operation on the same variable  $V$ .

As in so many programming languages, the operation  $\text{store}(V, x)$  is often written  $V \leftarrow x$  (or some similar notation), and  $\text{fetch}(V)$  is implied whenever a variable  $V$  is used in a context where a value is required. Thus, for example,  $V \leftarrow V + 1$  is commonly understood to be a shorthand for  $\text{store}(V, \text{fetch}(V) + 1)$ .

In this definition, it is implicitly assumed that storing a value into a variable  $U$  has no effect on the state of a distinct variable  $V$ . To make this assumption explicit, one could add the constraint that

- if  $U$  and  $V$  are distinct variables, the sequence  $\{ \text{store}(U, x); \text{store}(V, y) \}$  is equivalent to  $\{ \text{store}(V, y); \text{store}(U, x) \}$ .

More generally, ADT definitions often assume that any operation that changes the state of one ADT instance has no effect on the state of any other instance (including other instances of the same ADT) — unless the ADT axioms imply that the two instances are connected (aliased) in that sense. For example, when extending the definition of abstract variable to include abstract records, the operation that selects a field from a record variable  $R$  must yield a variable  $V$  that is aliased to that part of  $R$ .

The definition of an abstract variable  $V$  may also restrict the stored values  $x$  to members of a specific set  $X$ , called the *range* or *type* of  $V$ . As in programming languages, such restrictions may simplify the description and analysis of algorithms, and improve their readability.

Note that this definition does not imply anything about the result of evaluating  $\text{fetch}(V)$  when  $V$  is *un-initialized*, that is, before performing any  $\text{store}$  operation on  $V$ . An algorithm that does so is usually considered invalid, because its effect is not defined. (However, there are some important algorithms whose efficiency strongly depends on the assumption that such a  $\text{fetch}$  is legal, and returns some arbitrary value in the variable's range.)

### Instance creation

Some algorithms need to create new instances of some ADT (such as new variables, or new stacks). To describe such algorithms, one usually includes in the ADT definition a  $\text{create}()$  operation that yields an instance of the ADT, usually with axioms equivalent to

- the result of  $\text{create}()$  is distinct from any instance  $S$  in use by the algorithm.

This axiom may be strengthened to exclude also partial aliasing with other instances. On the other hand, this axiom still allows implementations of  $\text{create}()$  to yield a previously created instance that has become inaccessible to the program.

### Preconditions, postconditions, and invariants

In imperative-style definitions, the axioms are often expressed by *preconditions*, that specify when an operation may be executed; *postconditions*, that relate the states of the ADT before and after the execution of each operation; and *invariants*, that specify properties of the ADT that are *not* changed by the operations.

### Example: abstract stack (imperative)

As another example, an imperative definition of an abstract stack could specify that the state of a stack  $S$  can be modified only by the operations

- $\text{push}(S, x)$ , where  $x$  is some *value* of unspecified nature; and
- $\text{pop}(S)$ , that yields a value as a result;

with the constraint that

- For any value  $x$  and any abstract variable  $V$ , the sequence of operations  $\{ \text{push}(S, x); V \leftarrow \text{pop}(S) \}$  is equivalent to  $\{ V \leftarrow x \}$ ;

Since the assignment  $\{ V \leftarrow x \}$ , by definition, cannot change the state of  $S$ , this condition implies that  $\{ V \leftarrow \text{pop}(S) \}$  restores  $S$  to the state it had before the  $\{ \text{push}(S, x) \}$ . From this condition and from the properties of abstract variables, it follows, for example, that the sequence

$$\{ \text{push}(S, x); \text{push}(S, y); U \leftarrow \text{pop}(S); \text{push}(S, z); V \leftarrow \text{pop}(S); W \leftarrow \text{pop}(S); \}$$

where  $x, y$ , and  $z$  are any values, and  $U, V, W$  are pairwise distinct variables, is equivalent to

$$\{ U \leftarrow y; V \leftarrow z; W \leftarrow x \}$$

Here it is implicitly assumed that operations on a stack instance do not modify the state of any other ADT instance, including other stacks; that is,

- For any values  $x, y$ , and any distinct stacks  $S$  and  $T$ , the sequence  $\{ \text{push}(S, x); \text{push}(T, y) \}$  is equivalent to  $\{ \text{push}(T, y); \text{push}(S, x) \}$ .

A stack ADT definition usually includes also a Boolean-valued function  $\text{empty}(S)$  and a  $\text{create}()$  operation that returns a stack instance, with axioms equivalent to

- $\text{create}() \neq S$  for any stack  $S$  (a newly created stack is distinct from all previous stacks)
- $\text{empty}(\text{create}())$  (a newly created stack is empty)
- $\text{not empty}(\text{push}(S, x))$  (pushing something into a stack makes it non-empty)

### Single-instance style

Sometimes an ADT is defined as if only one instance of it existed during the execution of the algorithm, and all operations were applied to that instance, which is not explicitly notated. For example, the abstract stack above could have been defined with operations  $\text{push}(x)$  and  $\text{pop}()$ , that operate on "the" only existing stack. ADT definitions in this style can be easily rewritten to admit multiple coexisting instances of the ADT, by adding an explicit instance parameter (like  $S$  in the previous example) to every operation that uses or modifies the implicit instance.

On the other hand, some ADTs cannot be meaningfully defined without assuming multiple instances. This is the case when a single operation takes two distinct instances of the ADT as parameters. For an example, consider augmenting the definition of the stack ADT with an operation  $\text{compare}(S, T)$  that checks whether the stacks  $S$  and  $T$  contain the same items in the same order.

## Functional ADT definitions

Another way to define an ADT, closer to the spirit of functional programming, is to consider each state of the structure as a separate entity. In this view, any operation that modifies the ADT is modeled as a mathematical function that takes the old state as an argument, and returns the new state as part of the result. Unlike the "imperative" operations, these functions have no side effects. Therefore, the order in which they are evaluated is immaterial, and the same operation applied to the same arguments (including the same input states) will always return the same results (and output states).

In the functional view, in particular, there is no way (or need) to define an "abstract variable" with the semantics of imperative variables (namely, with `fetch` and `store` operations). Instead of storing values into variables, one passes them as arguments to functions.

### Example: abstract stack (functional)

For example, a complete functional-style definition of a stack ADT could use the three operations:

- `push`: takes a stack state and an arbitrary value, returns a stack state;
- `top`: takes a stack state, returns a value;
- `pop`: takes a stack state, returns a stack state;

with the following axioms:

- $\text{top}(\text{push}(s, x)) = x$  (pushing an item onto a stack leaves it at the top)
- $\text{pop}(\text{push}(s, x)) = s$  (`pop` undoes the effect of `push`)

In a functional-style definition there is no need for a `create` operation. Indeed, there is no notion of "stack instance". The stack states can be thought of as being potential states of a single stack structure, and two stack states that contain the same values in the same order are considered to be identical states. This view actually mirrors the behavior of some concrete implementations, such as linked lists with hash cons.

Instead of `create()`, a functional definition of a stack ADT may assume the existence of a special stack state, the *empty stack*, designated by a special symbol like  $\Lambda$  or  $()$ ; or define a `bottom()` operation that takes no arguments and returns this special stack state. Note that the axioms imply that

- $\text{push}(\Lambda, x) \neq \Lambda$

In a functional definition of a stack one does not need an `empty` predicate: instead, one can test whether a stack is empty by testing whether it is equal to  $\Lambda$ .

Note that these axioms do not define the effect of `top`( $s$ ) or `pop`( $s$ ), unless  $s$  is a stack state returned by a `push`. Since `push` leaves the stack non-empty, those two operations are undefined (hence invalid) when  $s = \Lambda$ . On the other hand, the axioms (and the lack of side effects) imply that  $\text{push}(s, x) = \text{push}(t, y)$  if and only if  $x = y$  and  $s = t$ .

As in some other branches of mathematics, it is customary to assume also that the stack states are only those whose existence can be proved from the axioms in a finite number of steps. In the stack ADT example above, this rule means that every stack is a *finite* sequence of values, that becomes the empty stack ( $\Lambda$ ) after a finite number of `pops`. By themselves, the axioms above do not exclude the existence of infinite stacks (that can be `popped` forever, each time yielding a different state) or circular stacks (that return to the same state after a finite number of `pops`). In particular, they do not exclude states  $s$  such that  $\text{pop}(s) = s$  or  $\text{push}(s, x) = s$  for some  $x$ . However, since one cannot obtain such stack states with the given operations, they are assumed "not to exist".

## Advantages of abstract data typing

- Encapsulation

Abstraction provides a promise that any implementation of the ADT has certain properties and abilities; knowing these is all that is required to make use of an ADT object. The user does not need any technical knowledge of how the implementation works to use the ADT. In this way, the implementation may be complex but will be encapsulated in a simple interface when it is actually used.

- Localization of change

Code that uses an ADT object will not need to be edited if the implementation of the ADT is changed. Since any changes to the implementation must still comply with the interface, and since code using an ADT may only refer to properties and abilities specified in the interface, changes may be made to the implementation without requiring any changes in code where the ADT is used.

- Flexibility

Different implementations of an ADT, having all the same properties and abilities, are equivalent and may be used somewhat interchangeably in code that uses the ADT. This gives a great deal of flexibility when using ADT objects in different situations. For example, different implementations of an ADT may be more efficient in different situations; it is possible to use each in the situation where they are preferable, thus increasing overall efficiency.

## Typical operations

Some operations that are often specified for ADTs (possibly under other names) are

- `compare(s,t)`, that tests whether two structures are equivalent in some sense;
- `hash(s)`, that computes some standard hash function from the instance's state;
- `print(s)` or `show(s)`, that produces a human-readable representation of the structure's state.

In imperative-style ADT definitions, one often finds also

- `create()`, that yields a new instance of the ADT;
- `initialize(s)`, that prepares a newly created instance  $s$  for further operations, or resets it to some "initial state";
- `copy(s,t)`, that puts instance  $s$  in a state equivalent to that of  $t$ ;
- `clone(t)`, that performs  $s \leftarrow \text{new}()$ , `copy(s,t)`, and returns  $s$ ;
- `free(s)` or `destroy(s)`, that reclaims the memory and other resources used by  $s$ ;

The `free` operation is not normally relevant or meaningful, since ADTs are theoretical entities that do not "use memory". However, it may be necessary when one needs to analyze the storage used by an algorithm that uses the ADT. In that case one needs additional axioms that specify how much memory each ADT instance uses, as a function of its state, and how much of it is returned to the pool by `free`.

## Examples

Some common ADTs, which have proved useful in a great variety of applications, are

- Container
  - Deque
  - List
  - Map
  - Multimap
  - Multiset
  - Priority queue
  - Queue
-

- Set
- Stack
- String
- Tree

Each of these ADTs may be defined in many ways and variants, not necessarily equivalent. For example, a stack ADT may or may not have a `count` operation that tells how many items have been pushed and not yet popped. This choice makes a difference not only for its clients but also for the implementation.

## Implementation

Implementing an ADT means providing one procedure or function for each abstract operation. The ADT instances are represented by some concrete data structure that is manipulated by those procedures, according to the ADT's specifications.

Usually there are many ways to implement the same ADT, using several different concrete data structures. Thus, for example, an abstract stack can be implemented by a linked list or by an array.

An ADT implementation is often packaged as one or more modules, whose interface contains only the signature (number and types of the parameters and results) of the operations. The implementation of the module — namely, the bodies of the procedures and the concrete data structure used — can then be hidden from most clients of the module. This makes it possible to change the implementation without affecting the clients.

When implementing an ADT, each instance (in imperative-style definitions) or each state (in functional-style definitions) is usually represented by a handle of some sort.<sup>[3]</sup>

Modern object-oriented languages, such as C++ and Java, support a form of abstract data types. When a class is used as a type, it is an abstract type that refers to a hidden representation. In this model an ADT is typically implemented as a class, and each instance of the ADT is an object of that class. The module's interface typically declares the constructors as ordinary procedures, and most of the other ADT operations as methods of that class. However, such an approach does not easily encapsulate multiple representational variants found in an ADT. It also can undermine the extensibility of object-oriented programs. In a pure object-oriented program that uses interfaces as types, types refer to behaviors not representations.

## Example: implementation of the stack ADT

As an example, here is an implementation of the stack ADT above in the C programming language.

### Imperative-style interface

An imperative-style interface might be:

```
typedef struct stack_Rep stack_Rep;      /* Type: instance
representation (an opaque record). */
typedef stack_Rep *stack_T;              /* Type: handle to a stack
instance (an opaque pointer). */
typedef void *stack_Item;                /* Type: value that can be
stored in stack (arbitrary address). */

stack_T stack_create(void);              /* Create new stack
instance, initially empty. */
void stack_push(stack_T s, stack_Item e); /* Add an item at the top of
the stack. */
stack_Item stack_pop(stack_T s);         /* Remove the top item from
the stack and return it . */
```

```
int stack_empty(stack_T ts);           /* Check whether stack is
empty. */
```

This implementation could be used in the following manner:

```
#include <stack.h>                     /* Include the stack interface. */
stack_T t = stack_create();           /* Create a stack instance. */
int foo = 17;                         /* An arbitrary datum. */
t = stack_push(t, &foo);              /* Push the address of 'foo' onto the
stack. */
...
void *e = stack_pop(t);               /* Get the top item and delete it from
the stack. */
if (stack_empty(t)) { ... }          /* Do something if stack is empty. */
...
```

This interface can be implemented in many ways. The implementation may be arbitrarily inefficient, since the formal definition of the ADT, above, does not specify how much space the stack may use, nor how long each operation should take. It also does not specify whether the stack state  $t$  continues to exist after a call  $s \leftarrow \text{pop}(t)$ .

In practice the formal definition should specify that the space is proportional to the number of items pushed and not yet popped; and that every one of the operations above must finish in a constant amount of time, independently of that number. To comply with these additional specifications, the implementation could use a linked list, or an array (with dynamic resizing) together with two integers (an item count and the array size)

### Functional-style interface

Functional-style ADT definitions are more appropriate for functional programming languages, and vice-versa. However, one can provide a functional style interface even in an imperative language like C. For example:

```
typedef struct stack_Rep stack_Rep;    /* Type: stack state
representation (an opaque record). */
typedef stack_Rep *stack_T;           /* Type: handle to a stack
state (an opaque pointer). */
typedef void *stack_Item;             /* Type: item (arbitrary
address). */

stack_T stack_empty(void);            /* Returns the empty stack
state. */
stack_T stack_push(stack_T s, stack_Item x); /* Adds x at the top of s,
returns the resulting state. */
stack_Item stack_top(stack_T s);      /* Returns the item
currently at the top of s. */
stack_T stack_pop(stack_T s);         /* Remove the top item
from s, returns the resulting state. */
```

The main problem is that C lacks garbage collection, and this makes this style of programming impractical; moreover, memory allocation routines in C are slower than allocation in a typical garbage collector, thus the performance impact of so many allocations is even greater.

## ADT libraries

Many modern programming languages, such as C++ and Java, come with standard libraries that implement several common ADTs, such as those listed above.

## Built-in abstract data types

The specification of some programming languages is intentionally vague about the representation of certain built-in data types, defining only the operations that can be done on them. Therefore, those types can be viewed as "built-in ADTs". Examples are the arrays in many scripting languages, such as Awk, Lua, and Perl, which can be regarded as an implementation of the Map or Table ADT.

## References

- [1] Barbara Liskov, Programming with Abstract Data Types, in Proceedings of the ACM SIGPLAN Symposium on Very High Level Languages, pp. 50--59, 1974, Santa Monica, California
- [2] Rudolf Lidl (2004). *Abstract Algebra*. Springer. ISBN 81-8128-149-7., Chapter 7, section 40.
- [3] Robert Sedgewick (1998). *Algorithms in C*. Addison/Wesley. ISBN 0-201-31452-5., definition 4.4.

## Further

- Mitchell, John C.; Plotkin, Gordon (July 1988). "Abstract Types Have Existential Type" (<http://theory.stanford.edu/~jcm/papers/mitch-plotkin-88.pdf>). *ACM Transactions on Programming Languages and Systems* **10** (3).

## External links

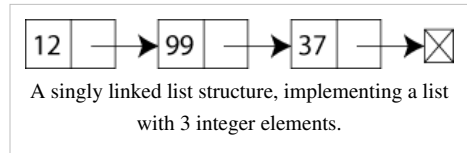
- Abstract data type (<http://www.nist.gov/dads/HTML/abstractDataType.html>) in NIST Dictionary of Algorithms and Data Structures

# List

In computer science, a **list** or **sequence** is an abstract data type that implements an ordered collection of values, where the same value may occur more than once. An instance of a list is a computer representation of the mathematical concept of a finite sequence. Each instance of a value in the list is usually called an **item**, **entry**, or **element** of the list; if the same value occurs multiple times, each occurrence is considered a distinct item.

The name **list** is also used for several concrete data structures that can be used to implement abstract lists, especially linked lists.

The so-called **static** list structures allow only inspection and enumeration of the values. A **mutable** or **dynamic** list may allow items to be inserted, replaced, or deleted during the list's existence.



Many programming languages provide support for **list data types**, and have special syntax and semantics for lists and list operations. A list can often be constructed by writing the items in sequence, separated by commas, semicolons, or spaces, within a pair of delimiters such as parentheses '()', brackets '[]', braces '{}', or angle brackets '<>'. Some languages may allow list types to be indexed or sliced like array types. In object-oriented programming languages, lists are usually provided as instances of subclasses of a generic "list" class. List data types are often implemented using arrays or linked lists of some sort, but other data structures may be more appropriate for some applications. In some contexts, such as in Lisp programming, the term list may refer specifically to a linked list rather than an array.

In type theory and functional programming, abstract lists are usually defined inductively by four operations: *nil* that yields the empty list, *cons*, which adds an item at the beginning of a list, *head*, that returns the first element of a list, and *tail* that returns a list minus its first element. Formally, Peano's natural numbers can be defined as abstract lists with elements of unit type.

## Operations

Implementation of the list data structure may provide some of the following operations:

- a constructor for creating an empty list;
- an operation for testing whether or not a list is empty;
- an operation for prepending an entity to a list
- an operation for appending an entity to a list
- an operation for determining the first component (or the "head") of a list
- an operation for referring to the list consisting of all the components of a list except for its first (this is called the "tail" of the list.)

## Characteristics

Lists have the following properties:

- The **size** of **lists**. It indicates how many elements there are in the list.
- **Equality** of lists:
  - In mathematics, sometimes equality of lists is defined simply in terms of object identity: two lists are equal if and only if they are the same object.
  - In modern programming languages, equality of lists is normally defined in terms of structural equality of the corresponding entries, except that if the lists are typed, then the list types may also be relevant.
- Lists may be **typed**. This implies that the entries in a list must have types that are compatible with the list's type. It is common that lists are typed when they are implemented using arrays.



- Each element in the list has an **index**. The first element commonly has index 0 or 1 (or some other predefined integer). Subsequent elements have indices that are 1 higher than the previous element. The last element has index  $\text{<initial index>} + \text{<size>} - 1$ .
  - It is possible to retrieve the element at a particular index.
  - It is possible to traverse the list in the order of increasing index.
  - It is possible to change the element at a particular index to a different value, without affecting any other elements.
  - It is possible to insert an element at a particular index. The indices of higher elements at that are increased by 1.
  - It is possible to remove an element at a particular index. The indices of higher elements at that are decreased by 1.

## Implementations

Lists are typically implemented either as linked lists (either singly or doubly linked) or as arrays, usually variable length or dynamic arrays.

The standard way of implementing lists, originating with the programming language Lisp, is to have each element of the list contain both its value and a pointer indicating the location of the next element in the list. This results in either a linked list or a tree, depending on whether the list has nested sublists. Some older Lisp implementations (such as the Lisp implementation of the Symbolics 3600) also supported "compressed lists" (using CDR coding) which had a special internal representation (invisible to the user). Lists can be manipulated using iteration or recursion. The former is often preferred in imperative programming languages, while the latter is the norm in functional languages.

Lists can be implemented as self-balancing binary search trees holding index-value pairs, providing equal-time access to any element (e.g. all residing in the fringe, and internal nodes storing the right-most child's index, used to guide the search), taking the time logarithmic in the list's size, but as long as it doesn't change much will provide the illusion of random access and enable swap, prefix and append operations in logarithmic time as well.

## Programming language support

Some languages do not offer a list data structure, but offer the use of associative arrays or some kind of table to emulate lists. For example, Lua provides tables. Although Lua stores lists that have numerical indices as arrays internally, they still appear as hash tables.

In Lisp, lists are the fundamental data type and can represent both program code and data. In most dialects, the list of the first three prime numbers could be written as `(list 2 3 5)`. In several dialects of Lisp, including Scheme, a list is a collection of pairs, consisting of a value and a pointer to the next pair (or null value), making a singly linked list.

## Applications

As the name implies, lists can be used to store a list of records. The items in a list can be sorted for the purpose of fast search (binary search).

Because in computing, lists are easier to realize than sets, a finite set in mathematical sense can be realized as a list with additional restrictions, that is, duplicate elements are disallowed and such that order is irrelevant. If the list is sorted, it speeds up determining if a given item is already in the set but in order to ensure the order, it requires more time to add new entry to the list. In efficient implementations, however, sets are implemented using self-balancing binary search trees or hash tables, rather than a list.

---

## Abstract definition

The abstract list type  $L$  with elements of some type  $E$  (a monomorphic list) is defined by the following functions:

$\text{nil}: () \rightarrow L$   
 $\text{cons}: E \times L \rightarrow L$   
 $\text{first}: L \rightarrow E$   
 $\text{rest}: L \rightarrow L$

with the axioms

$\text{first} (\text{cons} (e, l)) = e$   
 $\text{rest} (\text{cons} (e, l)) = l$

for any element  $e$  and any list  $l$ . It is implicit that

$\text{cons} (e, l) \neq l$   
 $\text{cons} (e, l) \neq e$   
 $\text{cons} (e_1, l_1) = \text{cons} (e_2, l_2)$  if  $e_1 = e_2$  and  $l_1 = l_2$

Note that  $\text{first} (\text{nil} ())$  and  $\text{rest} (\text{nil} ())$  are not defined.

These axioms are equivalent to those of the abstract stack data type.

In type theory, the above definition is more simply regarded as an inductive type defined in terms of constructors: *nil* and *cons*. In algebraic terms, this can be represented as the transformation  $1 + E \times L \rightarrow L$ . *first* and *rest* are then obtained by pattern matching on the *cons* constructor and separately handling the *nil* case.

## The list monad

The list type forms a monad with the following functions (using  $E^*$  rather than  $L$  to represent monomorphic lists with elements of type  $E$ ):

$\text{return}: A \rightarrow A^* = a \mapsto \text{cons } a \text{ nil}$   
 $\text{bind}: A^* \rightarrow (A \rightarrow B^*) \rightarrow B^* = l \mapsto f \mapsto \begin{cases} \text{nil} & \text{if } l = \text{nil} \\ \text{append } (f a) (\text{bind } l' f) & \text{if } l = \text{cons } a l' \end{cases}$

where *append* is defined as:

$\text{append}: A^* \rightarrow A^* \rightarrow A^* = l_1 \mapsto l_2 \mapsto \begin{cases} l_2 & \text{if } l_1 = \text{nil} \\ \text{cons } a (\text{append } l'_1 l_2) & \text{if } l_1 = \text{cons } a l'_1 \end{cases}$

Alternatively, the monad may be defined in terms of operations *return*, *fmap* and *join*, with:

$\text{fmap}: (A \rightarrow B) \rightarrow (A^* \rightarrow B^*) = f \mapsto l \mapsto \begin{cases} \text{nil} & \text{if } l = \text{nil} \\ \text{cons } (f a) (\text{fmap } f l') & \text{if } l = \text{cons } a l' \end{cases}$   
 $\text{join}: A^{**} \rightarrow A^* = l \mapsto \begin{cases} \text{nil} & \text{if } l = \text{nil} \\ \text{append } a (\text{join } l') & \text{if } l = \text{cons } a l' \end{cases}$

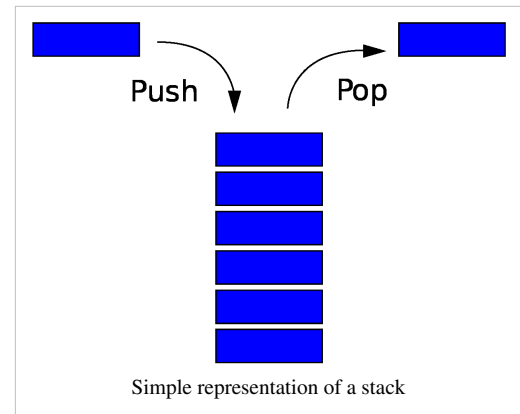
Note that *fmap*, *join*, *append* and *bind* are well-defined, since they're applied to progressively deeper arguments at each recursive call.

The list type is an additive monad, with *nil* as the monadic zero and *append* as monadic sum.

Lists form a monoid under the *append* operation. The identity element of the monoid is the empty list, *nil*. In fact, this is the free monoid over the set of list elements.

# Stack

In computer science, a **stack** is a last in, first out (LIFO) abstract data type and linear data structure. A stack can have any abstract data type as an element, but is characterized by two fundamental operations, called *push* and *pop* (or *pull*). The push operation adds a new item to the top of the stack, or initializes the stack if it is empty. If the stack is full and does not contain enough space to accept the given item, the stack is then considered to be in an overflow state. The pop operation removes an item from the top of the stack. A pop either reveals previously concealed items, or results in an empty stack, but if the stack is empty then it goes into underflow state (It means no items are present in stack to be removed). A stack pointer is the register which holds the value of the stack. The stack pointer always points to the top value of the stack.



A stack is a *restricted data structure*, because only a small number of operations are performed on it. The nature of the pop and push operations also means that stack elements have a natural order. Elements are removed from the stack in the reverse order to the order of their addition: therefore, the lower elements are those that have been on the stack the longest.<sup>[1]</sup>

## History

The stack was first proposed in 1946, in the computer design of Alan M. Turing (who used the terms "bury" and "unbury") as a means of calling and returning from subroutines. In 1957, the Germans Klaus Samelson and Friedrich L. Bauer patented the idea.<sup>[2]</sup> The same concept was developed, independently, by the Australian Charles Leonard Hamblin in the first half of 1957.<sup>[3]</sup>

## Abstract definition

A stack is a basic computer science data structure and can be defined in an abstract, implementation-free manner, or it can be generally defined as a linear list of items in which all additions and deletion are restricted to one end that is Top.

This is a VDM (*Vienna Development Method*) description of a stack:<sup>[4]</sup>

Function signatures:

```
init: -> Stack
push: N x Stack -> Stack
top: Stack -> (N U ERROR)
remove: Stack -> Stack
isempty: Stack -> Boolean
```

(where N indicates an element (natural numbers in this case), and U indicates set union)

Semantics:

```
top(init()) = ERROR
top(push(i, s)) = i
remove(init()) = init()
remove(push(i, s)) = s
```

```
isempty(init()) = true
isempty(push(i, s)) = false
```

## Inessential operations

In many implementations, a stack has more operations than "push" and "pop". An example is "top of stack", or "peek", which observes the top-most element *without* removing it from the stack.<sup>[5]</sup> Since this can be done with a "pop" and a "push" with the same data, it is not essential. An underflow condition can occur in the "stack top" operation if the stack is empty, the same as "pop". Often implementations have a function which just returns whether the stack is empty.

## Software stacks

### Implementation

In most high level languages, a stack can be easily implemented either through an array or a linked list. What identifies the data structure as a stack in either case is not the implementation but the interface: the user is only allowed to pop or push items onto the array or linked list, with few other helper operations. The following will demonstrate both implementations, using C.

#### Array

The **array implementation** aims to create an array where the first element (usually at the zero-offset) is the bottom. That is, `array[0]` is the first element pushed onto the stack and the last element popped off. The program must keep track of the size, or the length of the stack. The stack itself can therefore be effectively implemented as a two-element structure in C:

```
typedef struct {
    size_t size;
    int items[STACKSIZE];
} STACK;
```

The `push()` operation is used both to initialize the stack, and to store values to it. It is responsible for inserting (copying) the value into the `ps->items[]` array and for incrementing the element counter (`ps->size`). In a responsible C implementation, it is also necessary to check whether the array is already full to prevent an overrun.

```
void push(STACK *ps, int x)
{
    if (ps->size == STACKSIZE) {
        fputs("Error: stack overflow\n", stderr);
        abort();
    } else
        ps->items[ps->size++] = x;
}
```

The `pop()` operation is responsible for removing a value from the stack, and decrementing the value of `ps->size`. A responsible C implementation will also need to check that the array is not already empty.

```
int pop(STACK *ps)
{
    if (ps->size == 0) {
        fputs("Error: stack underflow\n", stderr);
    }
}
```

```

        abort();
    } else
        return ps->items[--ps->size];
}

```

If we use a dynamic array, then we can implement a stack that can grow or shrink as much as needed. The size of the stack is simply the size of the dynamic array. A dynamic array is a very efficient implementation of a stack, since adding items to or removing items from the end of a dynamic array is amortized  $O(1)$  time.

### Linked list

The **linked-list** implementation is equally simple and straightforward. In fact, a simple singly linked list is sufficient to implement a stack—it only requires that the head node or element can be removed, or popped, and a node can only be inserted by becoming the new head node.

Unlike the array implementation, our structure typedef corresponds not to the entire stack structure, but to a single node:

```

typedef struct stack {
    int data;
    struct stack *next;
} STACK;

```

Such a node is identical to a typical singly linked list node, at least to those that are implemented in C.

The `push()` operation both initializes an empty stack, and adds a new node to a non-empty one. It works by receiving a data value to push onto the stack, along with a target stack, creating a new node by allocating memory for it, and then inserting it into a linked list as the new head:

```

void push(STACK **head, int value)
{
    STACK *node = malloc(sizeof(STACK)); /* create a new node */

    if (node == NULL) {
        fputs("Error: no space available for node\n", stderr);
        abort();
    } else { /* initialize node */
        node->data = value;
        node->next = empty(*head) ? NULL : *head; /* insert new head if
any */
        *head = node;
    }
}

```

A `pop()` operation removes the head from the linked list, and assigns the pointer to the head to the previous second node. It checks whether the list is empty before popping from it:

```

int pop(STACK **head)
{
    if (empty(*head)) { /* stack is empty */
        fputs("Error: stack underflow\n", stderr);
        abort();
    } else { //pop a node

```

```
    STACK *top = *head;
    int value = top->data;
    *head = top->next;
    free(top);
    return value;
}
```

## Stacks and programming languages

Some languages, like LISP and Python, do not call for stack implementations, since **push** and **pop** functions are available for any list. All Forth-like languages (such as Adobe PostScript) are also designed around language-defined stacks that are directly visible to and manipulated by the programmer. Examples from Common Lisp:

```
(setf list (list 'a 'b 'c))
;; [] (A B C)
(pop list)
;; [] A
list
;; [] (B C)
(push 'new list)
;; [] (NEW B C)
```

C++'s Standard Template Library provides a "stack" templated class which is restricted to only push/pop operations. Java's library contains a `Stack` class that is a specialization of `Vector`---this could be considered a design flaw, since the inherited `get()` method from `Vector` ignores the LIFO constraint of the `Stack`. PHP has an `SplStack`<sup>[6]</sup> class.

## Hardware stacks

A common use of stacks at the architecture level is as a means of allocating and accessing memory.

### Basic architecture of a stack

A typical stack is an area of computer memory with a fixed origin and a variable size. Initially the size of the stack is zero. A *stack pointer*, usually in the form of a hardware register, points to the most recently referenced location on the stack; when the stack has a size of zero, the stack pointer points to the origin of the stack.

The two operations applicable to all stacks are:

- a *push* operation, in which a data item is placed at the location pointed to by the stack pointer, and the address in the stack pointer is adjusted by the size of the data item;
- a *pop* or *pull* operation: a data item at the current location pointed to by the stack pointer is removed, and the stack pointer is adjusted by the size of the data item.

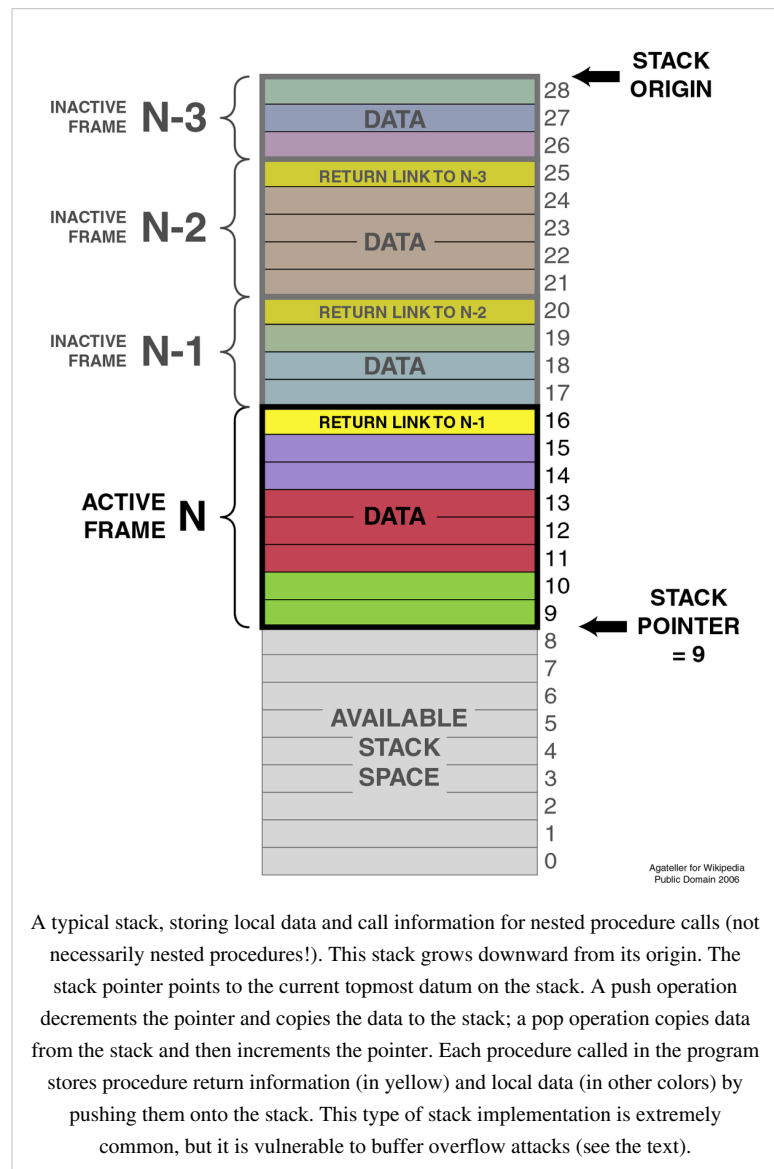
There are many variations on the basic principle of stack operations. Every stack has a fixed location in memory at which it begins. As data items are added to the stack, the stack pointer is displaced to indicate the current extent of the stack, which expands away from the origin.

Stack pointers may point to the origin of a stack or to a limited range of addresses

either above or below the origin (depending on the direction in which the stack grows); however, the stack pointer cannot cross the origin of the stack. In other words, if the origin of the stack is at address 1000 and the stack grows downwards (towards addresses 999, 998, and so on), the stack pointer must never be incremented beyond 1000 (to 1001, 1002, etc.). If a pop operation on the stack causes the stack pointer to move past the origin of the stack, a *stack underflow* occurs. If a push operation causes the stack pointer to increment or decrement beyond the maximum extent of the stack, a *stack overflow* occurs.

Some environments that rely heavily on stacks may provide additional operations, for example:

- *Duplicate*: the top item is popped, and then pushed again (twice), so that an additional copy of the former top item is now on top, with the original below it.
- *Peek*: the topmost item is inspected (or returned), but the stack pointer is not changed, and the stack size does not change (meaning that the item remains on the stack). This is also called **top** operation in many articles.



A typical stack, storing local data and call information for nested procedure calls (not necessarily nested procedures!). This stack grows downward from its origin. The stack pointer points to the current topmost datum on the stack. A push operation decrements the pointer and copies the data to the stack; a pop operation copies data from the stack and then increments the pointer. Each procedure called in the program stores procedure return information (in yellow) and local data (in other colors) by pushing them onto the stack. This type of stack implementation is extremely common, but it is vulnerable to buffer overflow attacks (see the text).

- *Swap or exchange*: the two topmost items on the stack exchange places.
- *Rotate (or Roll)*: the  $n$  topmost items are moved on the stack in a rotating fashion. For example, if  $n=3$ , items 1, 2, and 3 on the stack are moved to positions 2, 3, and 1 on the stack, respectively. Many variants of this operation are possible, with the most common being called *left rotate* and *right rotate*.

Stacks are either visualized growing from the bottom up (like real-world stacks), or, with the top of the stack in a fixed position (see image [note in the image, the top (28) is the stack 'bottom', since the stack 'top' is where items are pushed or popped from]), a coin holder, a Pez dispenser, or growing from left to right, so that "topmost" becomes "rightmost". This visualization may be independent of the actual structure of the stack in memory. This means that a *right rotate* will move the first element to the third position, the second to the first and the third to the second. Here are two equivalent visualizations of this process:

apple		banana
banana	===right rotate==>	cucumber
cucumber		apple
cucumber		apple
banana	===left rotate==>	cucumber
apple		banana

A stack is usually represented in computers by a block of memory cells, with the "bottom" at a fixed location, and the stack pointer holding the address of the current "top" cell in the stack. The top and bottom terminology are used irrespective of whether the stack actually grows towards lower memory addresses or towards higher memory addresses.

Pushing an item on to the stack adjusts the stack pointer by the size of the item (either decrementing or incrementing, depending on the direction in which the stack grows in memory), pointing it to the next cell, and copies the new top item to the stack area. Depending again on the exact implementation, at the end of a push operation, the stack pointer may point to the next unused location in the stack, or it may point to the topmost item in the stack. If the stack points to the current topmost item, the stack pointer will be updated before a new item is pushed onto the stack; if it points to the next available location in the stack, it will be updated *after* the new item is pushed onto the stack.

Popping the stack is simply the inverse of pushing. The topmost item in the stack is removed and the stack pointer is updated, in the opposite order of that used in the push operation.

## Hardware support

### Stack in main memory

Most CPUs have registers that can be used as stack pointers. Processor families like the x86, Z80, 6502, and many others have special instructions that implicitly use a dedicated (hardware) stack pointer to conserve opcode space. Some processors, like the PDP-11 and the 68000, also have special addressing modes for implementation of stacks, typically with a semi-dedicated stack pointer as well (such as A7 in the 68000). However, in most processors, several different registers may be used as additional stack pointers as needed (whether updated via addressing modes or via add/sub instructions).

### Stack in registers or dedicated memory

The x87 floating point architecture is an example of a set of registers organised as a stack where direct access to individual registers (relative the current top) is also possible. As with stack-based machines in general, having the top-of-stack as an implicit argument allows for a small machine code footprint with a good usage of bus bandwidth and code caches, but it also prevents some types of optimizations possible on processors permitting random access to the register file for all (two or three) operands. A stack structure also makes superscalar implementations with



register renaming (for speculative execution) somewhat more complex to implement, although it is still feasible, as exemplified by modern x87 implementations.

Sun SPARC, AMD Am29000, and Intel i960 are all examples of architectures using register windows within a register-stack as another strategy to avoid the use of slow main memory for function arguments and return values.

There are also a number of small microprocessors that implements a stack directly in hardware and some microcontrollers have a fixed-depth stack that is not directly accessible. Examples are the PIC microcontrollers, the Computer Cowboys MuP21, the Harris RTX line, and the Novix NC4016. Many stack-based microprocessors were used to implement the programming language Forth at the microcode level. Stacks were also used as a basis of a number of mainframes and mini computers. Such machines were called stack machines, the most famous being the Burroughs B5000.

## Applications

Stacks have numerous applications. We see stacks in everyday life, from the books in our library, to the sheaf of papers that we keep in our printer tray. All of them follow the *Last In First Out* (LIFO) logic, that is when we add a book to a pile of books, we add it to the top of the pile, whereas when we remove a book from the pile, we generally remove it from the top of the pile.

Given below are a few applications of stacks in the world of computers:

### Converting a decimal number into a binary number

The logic for transforming a decimal number into a binary number is as follows:

1. Read a number
2. Iteration (while number is greater than zero)
  1. Find out the remainder after dividing the number by 2
  2. Print the remainder
  3. Divide the number by 2
3. End the iteration

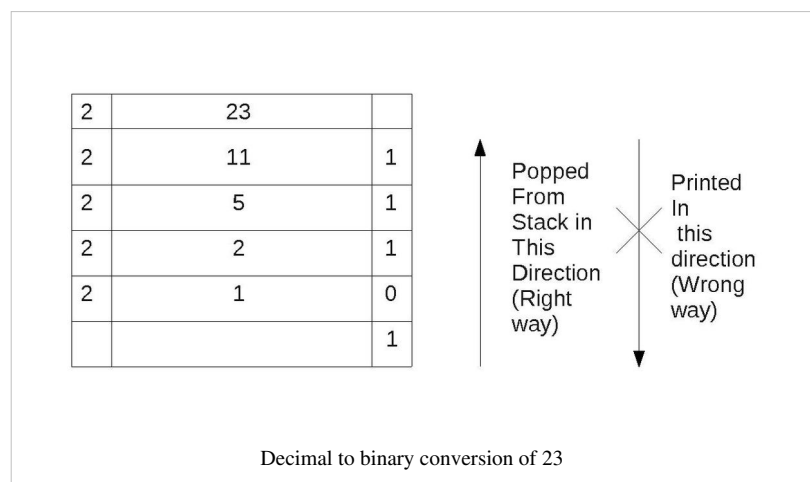
However, there is a problem with this

logic. Suppose the number, whose binary form we want to find is 23. Using this logic, we get the result as 11101, instead of getting 10111.

To solve this problem, we use a stack.<sup>[7]</sup> We make use of the *LIFO* property of the stack. Initially we *push* the binary digit formed into the stack, instead of printing it directly. After the entire number has been converted into the binary form, we *pop* one digit at a time from the stack and print it. Therefore we get the decimal number converted into its proper binary form.

#### Algorithm:

```
function outputInBinary(Integer n)
    Stack s = new Stack
    while n > 0 do
        Integer bit = n modulo 2
        s.push(bit)
    if s is full then
```



```
        return error
    end if
    n = floor(n / 2)
end while
while s is not empty do
    output(s.pop())
end while
end function
```

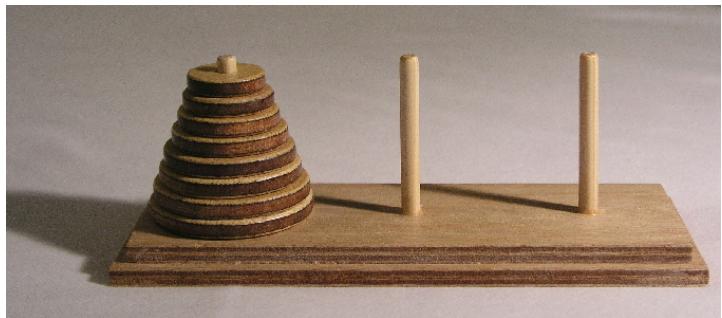
## Towers of Hanoi

One of the most interesting applications of stacks can be found in solving a puzzle called Tower of Hanoi. According to an old Brahmin story, the existence of the universe is calculated in terms of the time taken by a number of monks, who are working all the time, to move 64 disks from one pole to another. But there are some rules about how this should be done, which are:

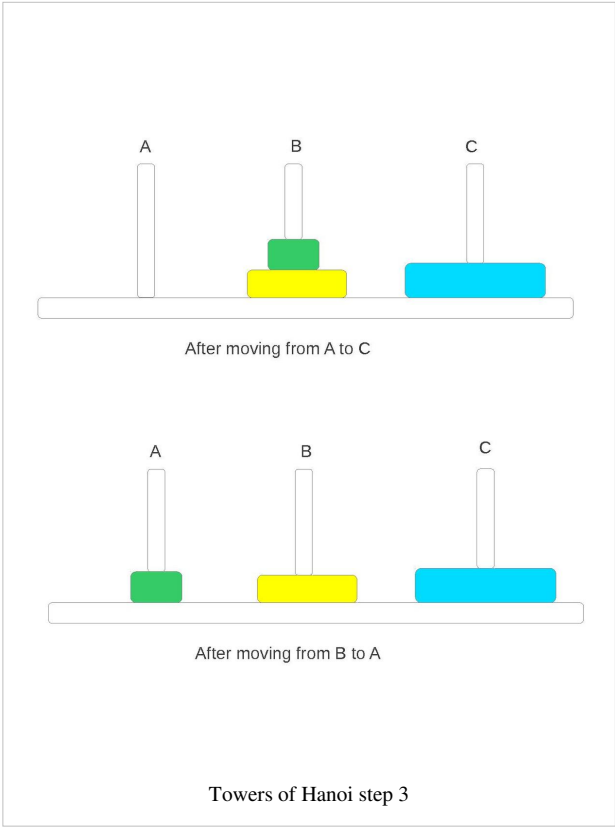
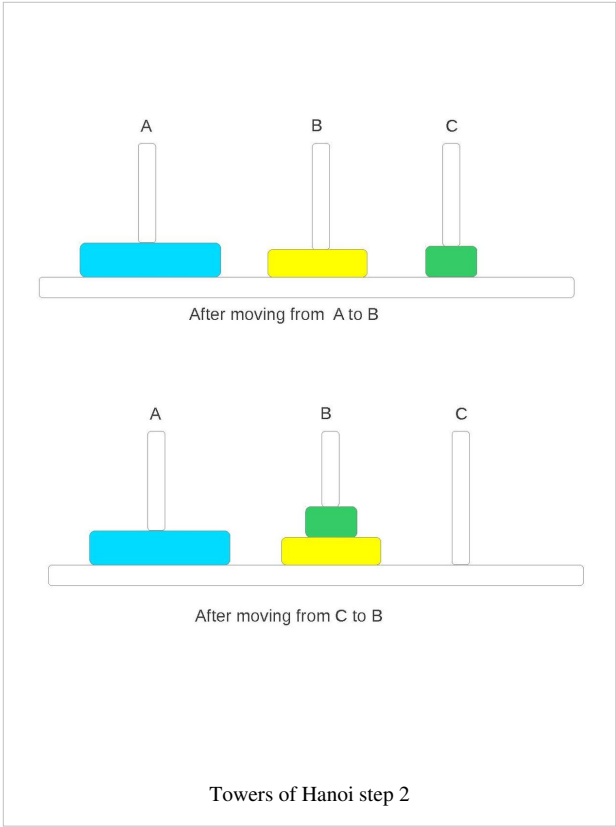
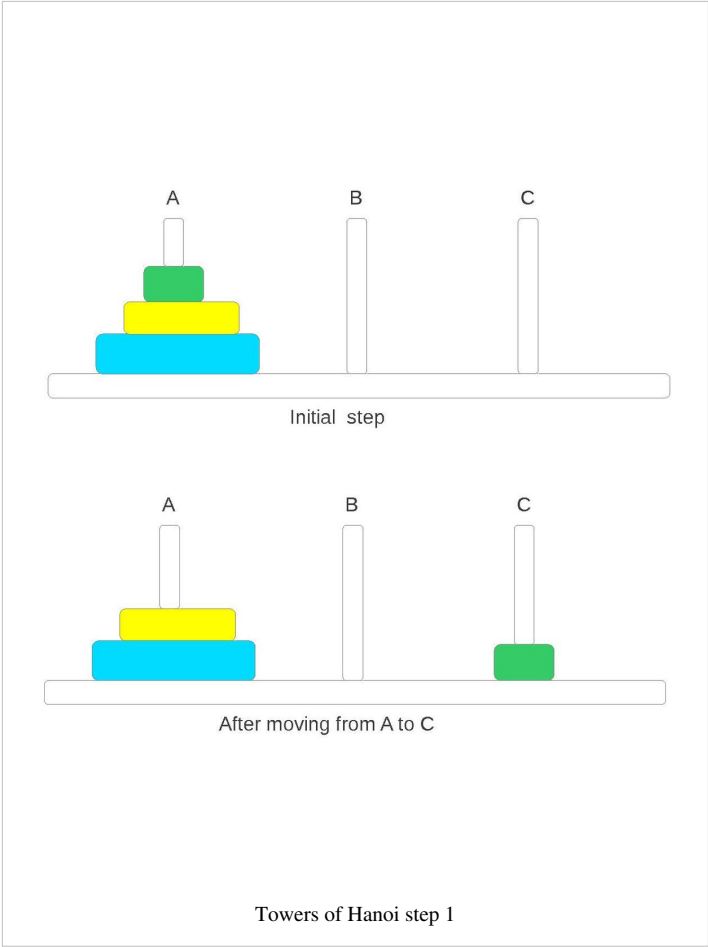
1. move only one disk at a time.
2. for temporary storage, a third pole may be used.
3. a disk of larger diameter may not be placed on a disk of smaller diameter.<sup>[8]</sup>

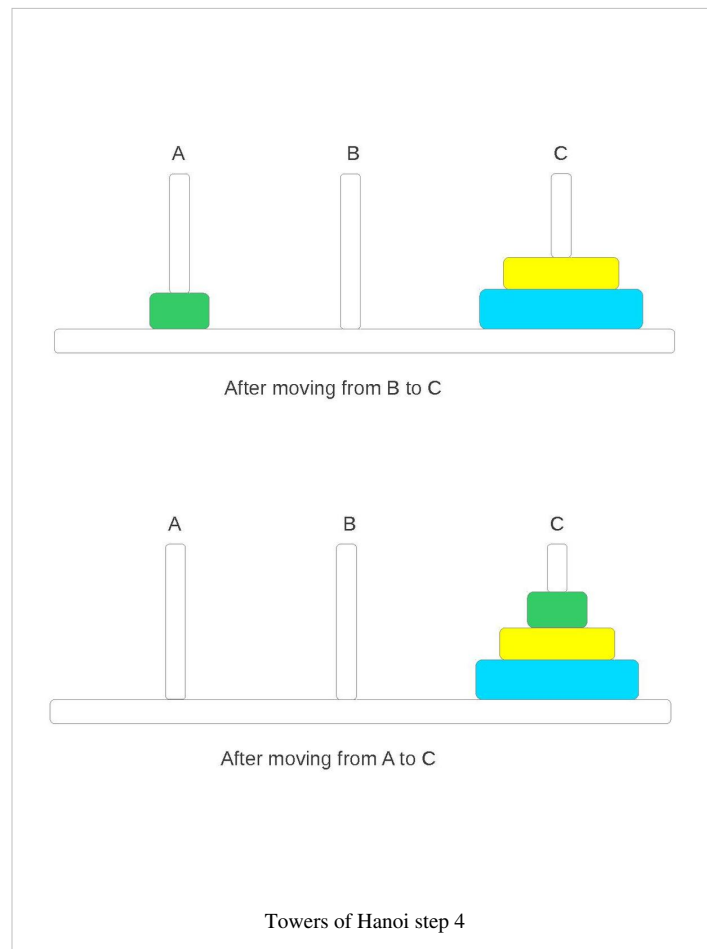
For algorithm of this puzzle see Tower of Hanoi.

Assume that A is first tower, B is second tower & C is third tower.



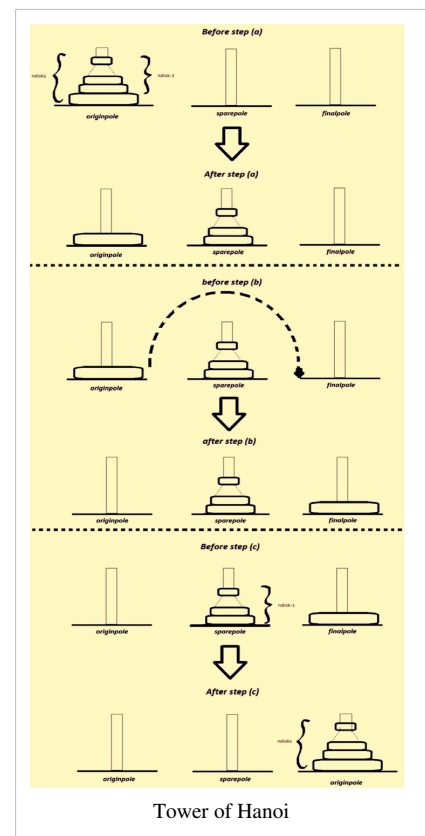
Towers of Hanoi





### Output: (when there are 3 disks)

Let 1 be the smallest disk, 2 be the disk of medium size and 3 be the largest disk.



Move disk	From peg	To peg
1	A	C
2	A	B
1	C	B
3	A	C
1	B	A
2	B	C
1	A	C

The C++ code for this solution can be implemented in two ways:

#### First implementation (using stacks implicitly by recursion)

```
void TowersofHanoi(int n, int a, int b, int c)
{
    if(n > 0)
    {
        TowersofHanoi(n-1, a, c, b);    //recursion
        cout << " Move top disk from tower " <<
            a << " to tower " << b << endl ;
        TowersofHanoi(n-1, c, b, a);    //recursion
    }
}
```

[9]

#### Second implementation (using stacks explicitly)

```
// Global variable , tower [1:3] are three towers
arrayStack<int> tower[4];

void TowerofHanoi(int n)
{
    // Preprocessor for moveAndShow.
    for (int d = n; d > 0; d--)        //initialize
        tower[1].push(d);              //add disk d to tower 1
    moveAndShow(n, 1, 2, 3);           /*move n disks from tower 1 to
tower 3 using
                                     tower 2 as intermediate tower*/
}

void moveAndShow(int n, int a, int b, int c)
{
    // Move the top n disks from tower a to tower b showing states.
    // Use tower c for intermediate storage.
    if(n > 0)
    {
```

```

        moveAndShow(n-1, a, c, b);           //recursion
        int d = tower[x].top();             //move a disc from top of tower
x to top of
        tower[x].pop();                     //tower y
        tower[y].push(d);
        showState();                       //show state of 3 towers
        moveAndShow(n-1, c, b, a);         //recursion
    }
}

```

However complexity for above written implementations is  $O(2^n)$ . So it's obvious that problem can only be solved for small values of  $n$  (generally  $n \leq 30$ ).

In case of the monks, the number of turns taken to transfer 64 disks, by following the above rules, will be 18,446,744,073,709,551,615; which will surely take a lot of time!<sup>[8][9]</sup>

## Expression evaluation and syntax parsing

Calculators employing reverse Polish notation use a stack structure to hold values. Expressions can be represented in prefix, postfix or infix notations and conversion from one form to another may be accomplished using a stack. Many compilers use a stack for parsing the syntax of expressions, program blocks etc. before translating into low level code. Most programming languages are context-free languages, allowing them to be parsed with stack based machines.

### Evaluation of an infix expression that is fully parenthesized

Input:  $((2 * 5) - (1 * 2)) / (11 - 9)$

Output: 4

Analysis: Five types of input characters

1. Opening bracket
2. Numbers
3. Operators
4. Closing bracket
5. New line character

Data structure requirement: A character stack

Algorithm

```

1. Read one input character

2. Actions at end of each input

Opening brackets      (2.1) Push into stack and then Go to step (1)
Number               (2.2) Push into stack and then Go to step (1)
Operator             (2.3) Push into stack and then Go to step (1)
Closing brackets     (2.4) Pop from character stack

                     (2.4.1) if it is opening bracket, then discard it, Go to step (1)
                     (2.4.2) Pop is used four times

                             The first popped element is assigned to op2
                             The second popped element is assigned to op
                             The third popped element is assigned to op1
                             The fourth popped element is the remaining opening bracket, which can be discarded
                             Evaluate op1 op op2

```

```

Convert the result into character and
push into the stack
Go to step (2.4)
New line character      (2.5) Pop from stack and print the answer
STOP

```

Result: The evaluation of the fully parenthesized infix expression is printed as follows:

Input String:  $((2 * 5) - (1 * 2)) / (11 - 9)$

Input Symbol	Stack (from bottom to top)	Operation
(	(	
(	((	
(	(( (	
2	(( ( 2	
*	(( ( 2 *	
5		(( ( 2 * 5
)	(( 10	2 * 5 = 10 and <i>push</i>
-	(( 10 -	
(	(( 10 - (	
1	(( 10 - ( 1	
*	(( 10 - ( 1 *	
2	(( 10 - ( 1 * 2	
)	(( 10 - 2	1 * 2 = 2 & <i>Push</i>
)	( 8	10 - 2 = 8 & <i>Push</i>
/	( 8 /	
(	( 8 / (	
11	( 8 / ( 11	
-	( 8 / ( 11 -	
9	( 8 / ( 11 - 9	
)	( 8 / 2	11 - 9 = 2 & <i>Push</i>
)	4	8 / 2 = 4 & <i>Push</i>
New line	Empty	<i>Pop &amp; Print</i>

### Evaluation of infix expression which is not fully parenthesized

Input:  $(2 * 5 - 1 * 2) / (11 - 9)$

Output: 4

Analysis: There are five types of input characters which are:

1. Opening brackets
2. Numbers
3. Operators
4. Closing brackets
5. New line character (\n)

We do not know what to do if an operator is read as an input character. By implementing the priority rule for operators, we have a solution to this problem.

The *Priority rule*: we should perform a comparative priority check if an operator is read, and then push it. If the stack *top* contains an operator of priority higher than or equal to the priority of the input operator, then we *pop* it and print it. We keep on performing the priority check until the *top* of stack either contains an operator of lower priority or if it does not contain an operator.

Data Structure Requirement for this problem: a character stack and an integer stack

Algorithm:

```

1. Read an input character
2. Actions that will be performed at the end of each input
   Opening brackets      (2.1)  Push it into character stack and then Go to step (1)
   Number                (2.2)  Push into integer stack, Go to step (1)
   Operator              (2.3)  Do the comparative priority check
                           (2.3.1) if the character stack's top contains an operator with equal
                                or higher priority, then pop it into op
                                Pop a number from integer stack into op2
                                Pop another number from integer stack into op1
                                Calculate op1 op op2 and push the result into the integer
                                stack
   Closing brackets      (2.4)  Pop from the character stack
                           (2.4.1) if it is an opening bracket, then discard it and Go to
                                step (1)
                           (2.4.2) To op, assign the popped element
                                Pop a number from integer stack and assign it op2
                                Pop another number from integer stack and assign it
                                to op1
                                Calculate op1 op op2 and push the result into the integer
                                stack
                                Convert into character and push into stack
                                Go to the step (2.4)
   New line character    (2.5)  Print the result after popping from the stack
                           STOP

```

Result: The evaluation of an infix expression that is not fully parenthesized is printed as follows:

Input String:  $(2 * 5 - 1 * 2) / (11 - 9)$



Input Symbol	Character Stack (from bottom to top)	Integer Stack (from bottom to top)	Operation performed
(	(		
2	(	2	
*	( *		Push as * has higher priority
5	( *	2 5	
-	( *		Since '-' has less priority, we do $2 * 5 = 10$
	( -	10	We push 10 and then push '-'
1	( -	10 1	
*	( - *	10 1	Push * as it has higher priority
2	( - *	10 1 2	
)	( -	10 2	Perform $1 * 2 = 2$ and push it
	(	8	Pop - and $10 - 2 = 8$ and push, Pop (
/	/	8	
(	/(	8	
11	/(	8 11	
-	/( -	8 11	
9	/( -	8 11 9	
)	/	8 2	Perform $11 - 9 = 2$ and push it
New line		4	Perform $8 / 2 = 4$ and push it
		4	Print the output, which is 4

[10]

### Evaluation of prefix expression

Input: / - \* 2 5 \* 1 2 - 11 9

Output: 4

Analysis: there are three types of input characters

1. Numbers
2. Operators
3. New line character (\n)

Data structure requirement: a character stack and an integer stack

Algorithm:

1. Read one character input at a time and keep pushing it into the character stack until the new line character is reached
2. Perform pop from the character stack. If the stack is empty, go to step (3)
  - Number (2.1) Push in to the integer stack and then go to step (1)
  - Operator (2.2) Assign the operator to op
    - Pop a number from integer stack and assign it to op1
    - Pop another number from integer stack and assign it to op2
    - Calculate  $op1 \text{ op } op2$  and push the output into the integer stack. Go to step (2)

3. Pop the result from the integer stack and display the result

Result: the evaluation of prefix expression is printed as follows:

Input String: / - \* 2 5 \* 1 2 - 11 9

Input Symbol	Character Stack (from bottom to top)	Integer Stack (from bottom to top)	Operation performed
/	/		
-	/		
*	/ - *		
2	/ - * 2		
5	/ - * 2 5		
*	/ - * 2 5 *		
1	/ - * 2 5 * 1		
2	/ - * 2 5 * 1 2		
-	/ - * 2 5 * 1 2 -		
11	/ - * 2 5 * 1 2 - 11		
9	/ - * 2 5 * 1 2 - 11 9		
\n	/ - * 2 5 * 1 2 - 11	9	
	/ - * 2 5 * 1 2 -	9 11	
	/ - * 2 5 * 1 2	2	11 - 9 = 2
	/ - * 2 5 * 1	2 2	
	/ - * 2 5 *	2 2 1	
	/ - * 2 5	2 2	1 * 2 = 2
	/ - * 2	2 2 5	
	/ - *	2 2 5 2	
	/ -	2 2 10	5 * 2 = 10
	/	2 8	10 - 2 = 8
	Stack is empty	4	8 / 2 = 4
		Stack is empty	Print 4

[10]

### Evaluation of postfix expression

The calculation:  $1 + 2 * 4 + 3$  can be written down like this in postfix notation with the advantage of no precedence rules and parentheses needed:

1 2 4 \* + 3 +

The expression is evaluated from the left to right using a stack:

1. when encountering an operand: push it
2. when encountering an operator: pop two operands, evaluate the result and push it.

Like the following way (the *Stack* is displayed after *Operation* has taken place):

Input	Operation	Stack (after op)
1	Push operand	1
2	Push operand	2, 1
4	Push operand	4, 2, 1
*	Multiply	8, 1
+	Add	9
3	Push operand	3, 9
+	Add	12

The final result, 12, lies on the top of the stack at the end of the calculation.

Example in C

```
#include<stdio.h>

int main()
{
    int a[100], i;
    printf("To pop enter -1\n");
    for(i = 0;;)
    {
        printf("Push ");
        scanf("%d", &a[i]);
        if(a[i] == -1)
        {
            if(i == 0)
            {
                printf("Underflow\n");
            }
            else
            {
                printf("pop = %d\n", a[--i]);
            }
        }
        else
        {
            i++;
        }
    }
}
```

**Evaluation of postfix expression (Pascal)**

This is an implementation in Pascal, using marked sequential file as data archives.

```
{
programmer : clx321
file      : stack.pas
unit      : Pstack.tpu
}
program TestStack;
{this program uses ADT of Stack, I will assume that the unit of ADT of
Stack has already existed}

uses
    PStack;    {ADT of STACK}

{dictionary}
const
    mark = '.';

var
    data : stack;
    f : text;
    cc : char;
    ccInt, cc1, cc2 : integer;

{functions}
IsOperand (cc : char) : boolean;    {JUST Prototype}
    {return TRUE if cc is operand}
ChrToInt (cc : char) : integer;    {JUST Prototype}
    {change char to integer}
Operator (cc1, cc2 : integer) : integer;    {JUST Prototype}
    {operate two operands}

{algorithms}
begin
    assign (f, cc);
    reset (f);
    read (f, cc);    {first elmt}
    if (cc = mark) then
        begin
            writeln ('empty archives !');
        end
    else
        begin
            repeat
                if (IsOperand (cc)) then
                    begin
                        ccInt := ChrToInt (cc);
```

```

        push (ccInt, data);
    end
else
    begin
        pop (cc1, data);
        pop (cc2, data);
        push (data, Operator (cc2, cc1));
    end;
    read (f, cc);    {next elmt}
until (cc = mark);
end;
close (f);
end
}

```

### Conversion of an Infix expression that is fully parenthesized into a Postfix expression

**Input:** (((8 + 1) - (7 - 4)) / (11 - 9))

**Output:** 8 1 + 7 4 - - 11 9 - /

**Analysis:** There are five types of input characters which are:

- \* Opening brackets
- \* Numbers
- \* Operators
- \* Closing brackets
- \* New line character (\n)

**Requirement:** A character stack

**Algorithm:**

1. Read an character input
2. Actions to be performed at end of each input
 

Opening brackets	(2.1) Push into stack and then Go to step (1)
Number	(2.2) Print and then Go to step (1)
Operator	(2.3) Push into stack and then Go to step (1)
Closing brackets	(2.4) Pop it from the stack
	(2.4.1) If it is an operator, print it, Go to step (2.4)
	(2.4.2) If the popped element is an opening bracket,
	discard it and go to step (1)
New line character	(2.5) STOP

Therefore, the final output after conversion of an infix expression to a postfix expression is as follows:

Input	Operation	Stack (after op)	Output on monitor
(	(2.1) Push operand into stack	(	
(	(2.1) Push operand into stack	((	
(	(2.1) Push operand into stack	(( (	
8	(2.2) Print it		8
+	(2.3) Push operator into stack	(( (+	8
1	(2.2) Print it		8 1
)	(2.4) Pop from the stack: Since popped element is '+' print it	(( (	8 1 +
	(2.4) Pop from the stack: Since popped element is '(' we ignore it and read next character	((	8 1 +
-	(2.3) Push operator into stack	(( (-	
(	(2.1) Push operand into stack	(( (- (	
7	(2.2) Print it		8 1 + 7
-	(2.3) Push the operator in the stack	(( (- (-	
4	(2.2) Print it		8 1 + 7 4
)	(2.4) Pop from the stack: Since popped element is '-' print it	(( (- (	8 1 + 7 4 -
	(2.4) Pop from the stack: Since popped element is '(' we ignore it and read next character	(( (-	
)	(2.4) Pop from the stack: Since popped element is '-' print it	(( (	8 1 + 7 4 - -
	(2.4) Pop from the stack: Since popped element is '(' we ignore it and read next character	(	
/	(2.3) Push the operand into the stack	( (/	
(	(2.1) Push into the stack	(( (/ (	
11	(2.2) Print it		8 1 + 7 4 - - 11
-	(2.3) Push the operand into the stack	(( (/ (-	
9	(2.2) Print it		8 1 + 7 4 - - 11 9
)	(2.4) Pop from the stack: Since popped element is '-' print it	(( (/ (	8 1 + 7 4 - - 11 9 -
	(2.4) Pop from the stack: Since popped element is '(' we ignore it and read next character	(( (/	
)	(2.4) Pop from the stack: Since popped element is '/' print it	(	8 1 + 7 4 - - 11 9 - /
	(2.4) Pop from the stack: Since popped element is '(' we ignore it and read next character	Stack is empty	
New line character	(2.5) STOP		

## Rearranging railroad cars

### Problem Description

This is one useful application of stacks. Consider that a freight train has  $n$  railroad cars, each to be left at different station. They're numbered 1 through  $n$  and freight train visits these stations in order  $n$  through 1. Obviously, the railroad cars are labeled by their destination. To facilitate removal of the cars from the train, we must rearrange them in ascending order of their number (i.e. 1 through  $n$ ). When cars are in this order, they can be detached at each station. We rearrange cars at a shunting yard that has *input track*, *output track* and  $k$  holding tracks between input & output tracks (i.e. *holding track*).

### Solution Strategy

To rearrange cars, we examine the cars on the input from front to back. If the car being examined is next one in the output arrangement, we move it directly to *output track*. If not, we move it to the *holding track* & leave it there until it's time to place it to the *output track*. The holding tracks operate in a LIFO manner as the cars enter & leave these tracks from top. When rearranging cars only following moves are permitted:

- A car may be moved from front (i.e. right end) of the input track to the top of one of the *holding tracks* or to the left end of the output track.
- A car may be moved from the top of *holding track* to left end of the *output track*.

The figure shows a shunting yard with  $k = 3$ , holding tracks *H1*, *H2* & *H3*, also  $n = 9$ . The  $n$  cars of freight train begin in the input track & are to end up in the output track in order 1 through  $n$  from right to left. The cars initially are in the order 5,8,1,7,4,2,9,6,3 from back to front. Later cars are rearranged in desired order.

### A Three Track Example

- Consider the input arrangement from figure, here we note that the car 3 is at the front, so it can't be output yet, as it to be preceded by cars 1 & 2. So car 3 is detached & moved to holding track **H1**.
- The next car 6 can't be output & it is moved to holding track **H2**. Because we have to output car 3 before car 6 & this will not possible if we move car 6 to holding track **H1**.
- Now it's obvious that we move car 9 to **H3**.

The requirement of rearrangement of cars on any holding track is that the cars should be preferred to arrange in ascending order from top to bottom.

- So car 2 is now moved to holding track H1 so that it satisfies the previous statement. If we move car 2 to H2 or H3, then we've no place to move cars 4,5,7,8. *The least restrictions on future car placement arise when the new car  $\lambda$  is moved to the holding track that has a car at its top with smallest label  $\Psi$  such that  $\lambda < \Psi$ . We may call it an assignment rule to decide whether a particular car belongs to a specific holding track.*
- When car 4 is considered, there are three places to move the car H1,H2,H3. The top of these tracks are 2,6,9. So using above mentioned Assignment rule, we move car 4 to H2.
- The car 7 is moved to H3.
- The next car 1 has the least label, so it's moved to output track.
- Now it's time for car 2 & 3 to output which are from H1 (in short all the cars from H1 are appended to car 1 on output track).

The car 4 is moved to output track. No other cars can be moved to output track at this time.

- The next car 8 is moved to holding track H1.
- Car 5 is output from input track. Car 6 is moved to output track from H2, so is the 7 from H3, 8 from H1 & 9 from H3.

[9]

### Backtracking

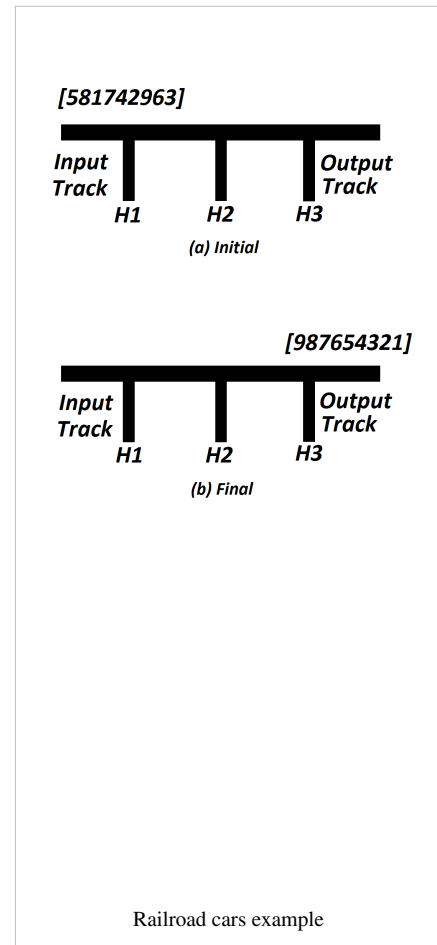
Another important application of stacks is Backtracking.

Consider a simple example of finding the correct path in a maze.

There are a series of points, from the starting point to the destination. We start from one point. To reach the final destination, there are several paths. Suppose we choose a random path. After following a certain path, we realise that the path we have chosen is wrong. So we need to find a way by which we can return back to the beginning of that path. This can be done with the use of stacks.

With the help of stacks, we remember the point where we have reached.

This is done by pushing that point into the stack. In case we end up on the wrong path, we can pop the last point from the stack and thus return back to the last point and continue our quest to find the right path. This is called backtracking.





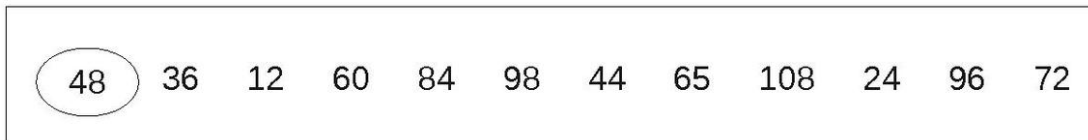
## Quicksort

Sorting means arranging the list of elements in a particular order. In case of numbers, it could be in ascending order, or in the case of letters, alphabetic order.

Quicksort is an algorithm of the *divide and conquer* type. In this method, to sort a set of numbers, we reduce it to two smaller sets, and then sort these smaller sets.

This can be explained with the help of the following example:

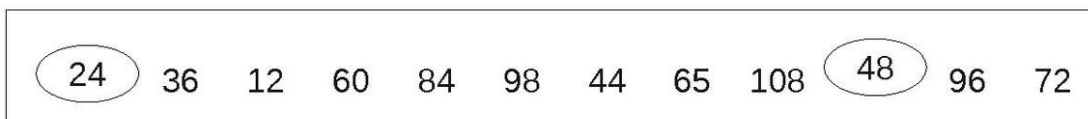
Suppose **A** is a list of the following numbers:



In the reduction step, we find the final position of one of the numbers. In this case, let us assume that we have to find the final position of 48, which is the first number in the list.

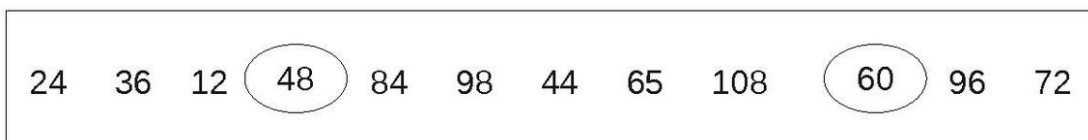
To accomplish this, we adopt the following method. Begin with the last number, and move from right to left. Compare each number with 48. If the number is smaller than 48, we stop at that number and swap it with 48.

In our case, the number is 24. Hence, we swap 24 and 48.



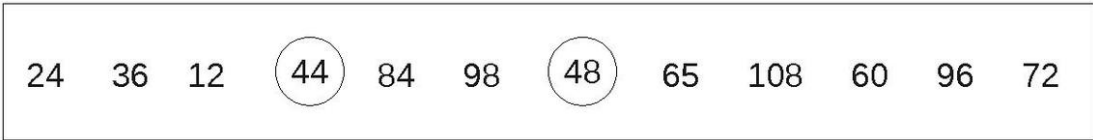
The numbers 96 and 72 to the right of 48, are greater than 48. Now beginning with 24, scan the numbers in the opposite direction, that is from left to right. Compare every number with 48 until you find a number that is greater than 48.

In this case, it is 60. Therefore we swap 48 and 60.

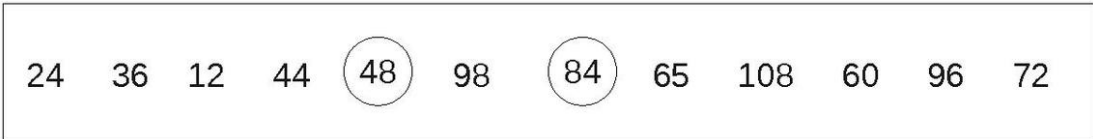


Note that the numbers 12, 24 and 36 to the left of 48 are all smaller than 48. Now, start scanning numbers from 60, in the right to left direction. As soon as you find lesser number, swap it with 48.

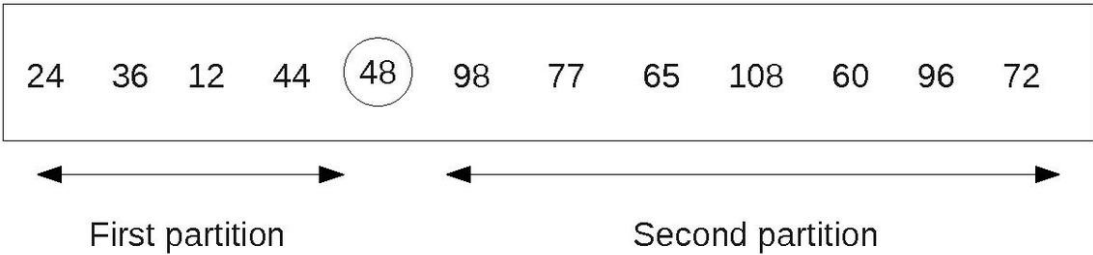
In this case, it is 44. Swap it with 48. The final result is:



Now, beginning with 44, scan the list from left to right, until you find a number greater than 48.  
Such a number is 84. Swap it with 48. The final result is:



Now, beginning with 84, traverse the list from right to left, until you reach a number lesser than 48. We do not find such a number before reaching 48. This means that all the numbers in the list have been scanned and compared with 48. Also, we notice that all numbers less than 48 are to the left of it, and all numbers greater than 48, are to its right.  
The final partitions look as follows:



Therefore, 48 has been placed in its proper position and now our task is reduced to sorting the two partitions. This above step of creating partitions can be repeated with every partition containing 2 or more elements. As we can process only a single partition at a time, we should be able to keep track of the other partitions, for future processing. This is done by using two **stacks** called LOWERBOUND and UPPERBOUND, to temporarily store these partitions. The addresses of the first and last elements of the partitions are pushed into the LOWERBOUND and UPPERBOUND stacks respectively. Now, the above reduction step is applied to the partitions only after its boundary values are *popped* from the stack.

We can understand this from the following example:

Take the above list A with 12 elements. The algorithm starts by pushing the boundary values of A, that is 1 and 12 into the LOWERBOUND and UPPERBOUND stacks respectively. Therefore the stacks look as follows:

LOWERBOUND :	1	UPPERBOUND :	12
--------------	---	--------------	----

To perform the reduction step, the values of the stack top are popped from the stack. Therefore, both the stacks become empty.

LOWERBOUND :	{empty}	UPPERBOUND :	{empty}
--------------	---------	--------------	---------

Now, the reduction step causes 48 to be fixed to the 5th position and creates two partitions, one from position 1 to 4 and the other from position 6 to 12. Hence, the values 1 and 6 are pushed into the LOWERBOUND stack and 4 and

12 are pushed into the UPPERBOUND stack.

LOWERBOUND: 1, 6

UPPERBOUND: 4, 12

For applying the reduction step again, the values at the stack top are popped. Therefore, the values 6 and 12 are popped. Therefore the stacks look like:

LOWERBOUND: 1

UPPERBOUND: 4

The reduction step is now applied to the second partition, that is from the 6th to 12th element.

A[6]	A[7]	A[8]	A[9]	A[10]	A[11]	A[12]
98	84	65	108	60	96	72
72	84	65	108	60	96	98
72	84	65	98	60	96	108
72	84	65	96	60	98	108

After the reduction step, 98 is fixed in the 11th position. So, the second partition has only one element. Therefore, we push the upper and lower boundary values of the first partition onto the stack. So, the stacks are as follows:

LOWERBOUND: 1, 6

UPPERBOUND: 4, 10

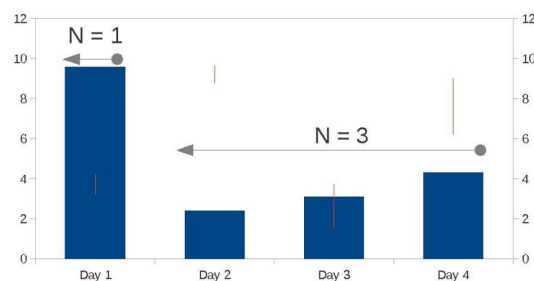
The processing proceeds in the following way and ends when the stacks do not contain any upper and lower bounds of the partition to be processed, and the list gets sorted.

[11]

## The Stock Span Problem

In the stock span problem, we will solve a financial problem with the help of stacks.

Suppose, for a stock, we have a series of  $n$  daily price quotes, the *span* of the stock's price on a particular day is defined as the maximum number of consecutive days for which the price of the stock on the current day is less than or equal to its price on that day.



The Stockspan Problem

### An algorithm which has Quadratic Time Complexity

**Input:** An array  $P$  with  $n$  elements

**Output:** An array  $S$  of  $n$  elements such that  $S[i]$  is the largest integer  $k$  such that  $k \leq i + 1$  and  $P[j] \leq P[i]$  for  $j = i - k + 1, \dots, i$

**Algorithm:**

```

1. Initialize an array P which contains the daily prices of the stocks
2. Initialize an array S which will store the span of the stock
3. for  $i = 0$  to  $i = n - 1$ 
    3.1 Initialize  $k$  to zero
    3.2 Done with a false condition
    3.3 repeat
        3.3.1 if (  $P[i - k] \leq P[i]$  ) then
            Increment  $k$  by 1
        3.3.2 else
            Done with true condition
    3.4 Till (  $k > i$  ) or done with processing
        Assign value of  $k$  to  $S[i]$  to get the span of the stock
4. Return array S

```

Now, analyzing this algorithm for running time, we observe:

- We have initialized the array  $S$  at the beginning and returned it at the end. This is a constant time operation, hence takes  $O(n)$  time
- The *repeat* loop is nested within the *for* loop. The *for* loop, whose counter is  $i$  is executed  $n$  times. The statements which are not in the repeat loop, but in the for loop are executed  $n$  times. Therefore these statements and the incrementing and condition testing of  $i$  take  $O(n)$  time.
- In repetition of  $i$  for the outer for loop, the body of the inner *repeat* loop is executed maximum  $i + 1$  times. In the worst case, element  $S[i]$  is greater than all the previous elements. So, testing for the if condition, the statement after that, as well as testing the until condition, will be performed  $i + 1$  times during iteration  $i$  for the outer for loop. Hence, the total time taken by the inner loop is  $O(n(n + 1)/2)$ , which is  $O(n^2)$

The running time of all these steps is calculated by adding the time taken by all these three steps. The first two terms are  $O(n)$  while the last term is  $O(n^2)$ . Therefore the total running time of the algorithm is  $O(n^2)$ .

### An algorithm which has Linear Time Complexity

In order to calculate the span more efficiently, we see that the span on a particular day can be easily calculated if we know the closest day before  $i$ , such that the price of the stocks on that day was higher than the price of the stocks on the present day. If there exists such a day, we can represent it by  $h(i)$  and initialize  $h(i)$  to be  $-1$ .

Therefore the span of a particular day is given by the formula,  $s = i - h(i)$ .

To implement this logic, we use a stack as an abstract data type to store the days  $i$ ,  $h(i)$ ,  $h(h(i))$  and so on. When we go from day  $i-1$  to  $i$ , we pop the days when the price of the stock was less than or equal to  $p(i)$  and then push the value of day  $i$  back into the stack.

Here, we assume that the stack is implemented by operations that take  $O(1)$  that is constant time. The algorithm is as follows:

**Input:** An array  $P$  with  $n$  elements and an empty stack  $N$

**Output:** An array  $S$  of  $n$  elements such that  $P[i]$  is the largest integer  $k$  such that  $k \leq i + 1$  and  $P[j] \leq P[i]$  for  $j = i - k + 1, \dots, i$

**Algorithm:**

```

1. Initialize an array P which contains the daily prices of the stocks
2. Initialize an array S which will store the span of the stock
3. for i = 0 to i = n - 1
    3.1 Initialize k to zero
    3.2 Done with a false condition
    3.3 while not (Stack N is empty or done with processing)
        3.3.1 if ( P[i] >= P[N.top()]) then
            Pop a value from stack N
        3.3.2 else
            Done with true condition
    3.4 if Stack N is empty
        3.4.1 Initialize h to -1
    3.5 else
        3.5.1 Initialize stack top to h
    3.6 Put the value of h - i in S[i]
    3.7 Push the value of i in N
4. Return array S

```

Now, analyzing this algorithm for running time, we observe:

- We have initialized the array S at the beginning and returned it at the end. This is a constant time operation, hence takes  $O(1)$  time
- The *while* loop is nested within the *for* loop. The *for* loop, whose counter is  $i$  is executed  $n$  times. The statements which are not in the repeat loop, but in the *for* loop are executed  $n$  times. Therefore these statements and the incrementing and condition testing of  $i$  take  $O(n)$  time.
- Now, observe the inner while loop during  $i$  repetitions of the *for* loop. The statement *done with a true condition* is done at most once, since it causes an exit from the loop. Let us say that  $t(i)$  is the number of times statement *Pop a value from stack N* is executed. So it becomes clear that *while not (Stack N is empty or done with processing)* is tested maximum  $t(i) + 1$  times.
- Adding the running time of all the operations in the while loop, we get:

$$\sum_{i=0}^{n-1} t(i) + 1$$

- An element once popped from the stack N is never pushed back into it. Therefore,

$$\sum_{i=1}^{n-1} t(i)$$

So, the running time of all the statements in the while loop is  $O(n)$

The running time of all the steps in the algorithm is calculated by adding the time taken by all these steps. The running time of each step is  $O(n)$ . Hence the running time complexity of this algorithm is  $O(n)$ .

[12]

## Runtime memory management

A number of programming languages are stack-oriented, meaning they define most basic operations (adding two numbers, printing a character) as taking their arguments from the stack, and placing any return values back on the stack. For example, PostScript has a return stack and an operand stack, and also has a graphics state stack and a dictionary stack.

Forth uses two stacks, one for argument passing and one for subroutine return addresses. The use of a return stack is extremely commonplace, but the somewhat unusual use of an argument stack for a human-readable programming language is the reason Forth is referred to as a *stack-based* language.

Many virtual machines are also stack-oriented, including the p-code machine and the Java Virtual Machine.

Almost all calling conventions – computer runtime memory environments – use a special stack (the "call stack") to hold information about procedure/function calling and nesting in order to switch to the context of the called function and restore to the caller function when the calling finishes. The functions follow a runtime protocol between caller and callee to save arguments and return value on the stack. Stacks are an important way of supporting nested or recursive function calls. This type of stack is used implicitly by the compiler to support CALL and RETURN statements (or their equivalents) and is not manipulated directly by the programmer.

Some programming languages use the stack to store data that is local to a procedure. Space for local data items is allocated from the stack when the procedure is entered, and is deallocated when the procedure exits. The C programming language is typically implemented in this way. Using the same stack for both data and procedure calls has important security implications (see below) of which a programmer must be aware in order to avoid introducing serious security bugs into a program.

## Security

Some computing environments use stacks in ways that may make them vulnerable to security breaches and attacks. Programmers working in such environments must take special care to avoid the pitfalls of these implementations.

For example, some programming languages use a common stack to store both data local to a called procedure and the linking information that allows the procedure to return to its caller. This means that the program moves data into and out of the same stack that contains critical return addresses for the procedure calls. If data is moved to the wrong location on the stack, or an oversized data item is moved to a stack location that is not large enough to contain it, return information for procedure calls may be corrupted, causing the program to fail.

Malicious parties may attempt a stack smashing attack that takes advantage of this type of implementation by providing oversized data input to a program that does not check the length of input. Such a program may copy the data in its entirety to a location on the stack, and in so doing it may change the return addresses for procedures that have called it. An attacker can experiment to find a specific type of data that can be provided to such a program such that the return address of the current procedure is reset to point to an area within the stack itself (and within the data provided by the attacker), which in turn contains instructions that carry out unauthorized operations.

This type of attack is a variation on the buffer overflow attack and is an extremely frequent source of security breaches in software, mainly because some of the most popular programming languages (such as C) use a shared stack for both data and procedure calls, and do not verify the length of data items. Frequently programmers do not write code to verify the size of data items, either, and when an oversized or undersized data item is copied to the stack, a security breach may occur.

---

## References

- [1] <http://www.cprogramming.com/tutorial/computersciencetheory/stack.html> cprogramming.com
  - [2] Dr. Friedrich Ludwig Bauer and Dr. Klaus Samelson (30. März 1957) (in german). *Verfahren zur automatischen Verarbeitung von kodierten Daten und Rechenmaschine zur Ausübung des Verfahrens*. (<http://v3.espacenet.com/origdoc?DB=EPODOC&IDX=DE1094019&F=0&QPN=DE1094019>). Deutsches Patentamt. . Retrieved 2010-10-01.
  - [3] C. L. Hamblin, "An Addressless Coding Scheme based on Mathematical Notation", N.S.W University of Technology, May 1957 (typescript)
  - [4] Jones: "Systematic Software Development Using VDM"
  - [5] Horowitz, Ellis: "Fundamentals of Data Structures in Pascal", page 67. Computer Science Press, 1984
  - [6] <http://www.php.net/manual/en/class.splstack.php>
  - [7] Richard F. Gilberg; Behrouz A. Forouzan. *Data Structures-A Pseudocode Approach with C++*. Thomson Brooks/Cole.
  - [8] Dromey, R.G. *How to Solve it by Computer*. Prentice Hall of India.
  - [9] Data structures, Algorithms and Applications in C++ by Sartaj Sahni
  - [10] Gopal, Arpita. *Magnifying Data Structures*. PHI.
  - [11] Lipschutz, Seymour. *Theory and Problems of Data Structures*. Tata McGraw Hill.
  - [12] Goodrich, Tamassia, Mount, Michael, Roberto, David. *Data Structures and Algorithms in C++*. Wiley-India.
- Stack implementation on goodsoft.org.ua ([http://goodsoft.org.ua/en/data\\_struct/stack.html](http://goodsoft.org.ua/en/data_struct/stack.html))


## Further reading

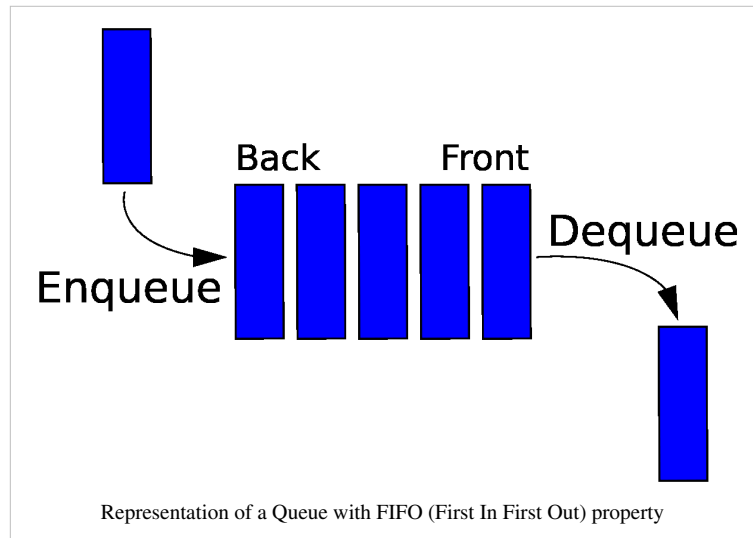
- Donald Knuth. *The Art of Computer Programming*, Volume 1: *Fundamental Algorithms*, Third Edition. Addison-Wesley, 1997. ISBN 0-201-89683-4. Section 2.2.1: Stacks, Queues, and Deques, pp. 238–243.
- Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7. Section 10.1: Stacks and queues, pp. 200–204.

## External links

- Stack Machines - the new wave ([http://www.ece.cmu.edu/~koopman/stack\\_computers/index.html](http://www.ece.cmu.edu/~koopman/stack_computers/index.html))
- Bounding stack depth (<http://www.cs.utah.edu/~regehr/stacktool>)
- Libsafe - Protecting Critical Elements of Stacks (<http://research.avayalabs.com/project/libsafe/>)
- VBScript implementation of stack, queue, deque, and Red-Black Tree (<http://www.ludvikjerabek.com/downloads.html>)
- Stack Size Analysis for Interrupt-driven Programs (<http://www.cs.ucla.edu/~palsberg/paper/sas03.pdf>) (322 KB)
- Paul E. Black, Bounded stack (<http://www.nist.gov/dads/HTML/boundedstack.html>) at the NIST Dictionary of Algorithms and Data Structures.

# Queue

In computer science, a **queue** (  /'kju:/ *KEW*) is a particular kind of abstract data type or collection in which the entities in the collection are kept in order and the principal (or only) operations on the collection are the addition of entities to the rear terminal position and removal of entities from the front terminal position. This makes the queue a First-In-First-Out (FIFO) data structure. In a FIFO data structure, the first element added to the queue will be the first one to be removed. This is equivalent to the requirement that once an element is added, all elements that were added before have to be removed before the new element can be invoked. A queue is an example of a linear data structure.



Queues provide services in computer science, transport, and operations research where various entities such as data, objects, persons, or events are stored and held to be processed later. In these contexts, the queue performs the function of a buffer.

Queues are common in computer programs, where they are implemented as data structures coupled with access routines, as an abstract data structure or in object-oriented languages as classes. Common implementations are circular buffers and linked lists.

## Representing a queue

In each of the cases, the customer or object at the front of the line was the first one to enter, while at the end of the line is the last to have entered. Every time a customer finishes paying for their items (or a person steps off the escalator, or the machine part is removed from the assembly line, etc.) that object leaves the queue from the front. This represents the queue “dequeue” function. Every time another object or customer enters the line to wait, they join the end of the line and represent the “enqueue” function. The queue “size” function would return the length of the line, and the “empty” function would return true only if there was nothing in the line.

## Queue implementation

Theoretically, one characteristic of a queue is that it does not have a specific capacity. Regardless of how many elements are already contained, a new element can always be added. It can also be empty, at which point removing an element will be impossible until a new element has been added again.

Fixed length arrays are limited in capacity, and inefficient because items need to be copied towards the head of the queue. However conceptually they are simple and work with early languages such as FORTRAN and BASIC which did not have pointers or objects. Most modern languages with objects or pointers can implement or come with libraries for dynamic lists. Such data structures may have not specified fixed capacity limit besides memory constraints. Queue **overflow** results from trying to add an element onto a full queue and queue **underflow** happens when trying to remove an element from an empty queue.

A **bounded queue** is a queue limited to a fixed number of items.



There are several efficient implementations of FIFO queues. An efficient implementation is one that can perform the operations—enqueueing and dequeueing—in  $O(1)$  time.

- Linked list
  - A doubly linked list has  $O(1)$  insertion and deletion at both ends, so is a natural choice for queues.
  - A regular singly linked list only has efficient insertion and deletion at one end. However, a small modification—keeping a pointer to the *last* node in addition to the first one—will enable it to implement an efficient queue.
- A deque implemented using a modified dynamic array

## Queues and programming languages

Some languages, like Perl and Ruby, already have operations for pushing and popping an array from both ends, so one can use **push** and **shift** functions to enqueue and dequeue a list (or, in reverse, one can use **unshift** and **pop**), although in some cases these operations are not efficient.

C++'s Standard Template Library provides a "queue" templated class which is restricted to only push/pop operations. Since J2SE5.0, Java's library contains a `Queue` interface that specifies queue operations; implementing classes include `LinkedList` and (since J2SE 1.6) `ArrayDeque`. PHP has an `SplQueue`<sup>[1]</sup> class and third party libraries like *beanstalk'd* and *Gearman*.

## References

### General

- Donald Knuth. *The Art of Computer Programming*, Volume 1: *Fundamental Algorithms*, Third Edition. Addison-Wesley, 1997. ISBN 0-201-89683-4. Section 2.2.1: Stacks, Queues, and Deques, pp. 238–243.
- Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7. Section 10.1: Stacks and queues, pp. 200–204.
- William Ford, William Topp. *Data Structures with C++ and STL*, Second Edition. Prentice Hall, 2002. ISBN 0-13-085850-1. Chapter 8: Queues and Priority Queues, pp. 386–390.
- Adam Drozdek. *Data Structures and Algorithms in C++*, Third Edition. Thomson Course Technology, 2005. ISBN 0-534-49182-0. Chapter 4: Stacks and Queues, pp. 137–169.

### Citations

[1] <http://www.php.net/manual/en/class.splqueue.php>

## External links

- STL Quick Reference (<http://www.halpernwrightsoftware.com/stdlib-scratch/quickref.html#containers14>)
- VBScript implementation of stack, queue, deque, and Red-Black Tree (<http://www.ludvikjerabek.com/downloads.html>)

Paul E. Black, Bounded queue (<http://www.nist.gov/dads/HTML/boundedqueue.html>) at the NIST Dictionary of Algorithms and Data Structures.

# Deque

In computer science, a **double-ended queue** (**dequeue**, often abbreviated to **deque**, pronounced *deck*) is an abstract data type that implements a queue for which elements can only be added to or removed from the front (head) or back (tail).<sup>[1]</sup> It is also often called a **head-tail linked list**.

## Naming conventions

*Deque* is sometimes written *dequeue*, but this use is generally deprecated in technical literature or technical writing because *dequeue* is also a verb meaning "to remove from a queue". Nevertheless, several libraries and some writers, such as Aho, Hopcroft, and Ullman in their textbook *Data Structures and Algorithms*, spell it *dequeue*. John Mitchell, author of *Concepts in Programming Languages*, also uses this terminology.

## Distinctions and sub-types

This differs from the queue abstract data type or First-In-First-Out List (FIFO), where elements can only be added to one end and removed from the other. This general data class has some possible sub-types:

- An input-restricted deque is one where deletion can be made from both ends, but insertion can only be made at one end.
- An output-restricted deque is one where insertion can be made at both ends, but deletion can be made from one end only.

Both the basic and most common list types in computing, queues and stacks can be considered specializations of deques, and can be implemented using deques.

## Operations

The following operations are possible on a deque:

operation	Ada	C++	Java	Perl	PHP	Python	Ruby	JavaScript
insert element at back	Append	push_back	offerLast	push	array_push	append	push	push
insert element at front	Prepend	push_front	offerFirst	unshift	array_unshift	appendleft	unshift	unshift
remove last element	Delete_Last	pop_back	pollLast	pop	array_pop	pop	pop	pop
remove first element	Delete_First	pop_front	pollFirst	shift	array_shift	popleft	shift	shift
examine last element	Last_Element	back	peekLast	\$array[-1]	end	<obj>[-1]	last	<obj>[<obj>.length - 1]
examine first element	First_Element	front	peekFirst	\$array[0]	reset	<obj>[0]	first	<obj>[0]

## Implementations

There are at least two common ways to efficiently implement a deque: with a modified dynamic array or with a doubly linked list.

The dynamic array approach uses a variant of a dynamic array that can grow from both ends, sometimes called **array deques**. These array deques have all the properties of a dynamic array, such as constant time random access, good locality of reference, and inefficient insertion/removal in the middle, with the addition of amortized constant time insertion/removal at both ends, instead of just one end. Three common implementations include:

- Storing deque contents in a circular buffer, and only resizing when the buffer becomes full. This decreases the frequency of resizings.
- Allocating deque contents from the center of the underlying array, and resizing the underlying array when either end is reached. This approach may require more frequent resizings and waste more space, particularly when elements are only inserted at one end.
- Storing contents in multiple smaller arrays, allocating additional arrays at the beginning or end as needed. Indexing is implemented by keeping a dynamic array containing pointers to each of the smaller arrays.

## Language support

Ada's `containers` provides the generic packages `Ada.Containers.Vectors` and `Ada.Containers.Doubly_Linked_Lists`, for the dynamic array and linked list implementations, respectively.

C++'s Standard Template Library provides the class templates `std::deque` and `std::list`, for the multiple array and linked list implementations, respectively.

As of Java 6, Java's Collections Framework provides a new `Deque` interface that provides the functionality of insertion and removal at both ends. It is implemented by classes such as `ArrayDeque` (also new in Java 6) and `LinkedList`, providing the dynamic array and linked list implementations, respectively. However, the `ArrayDeque`, contrary to its name, does not support random access.

Python 2.4 introduced the `collections` module with support for deque objects.

As of PHP 5.3, PHP's SPL extension contains the `'SplDoublyLinkedList'` class that can be used to implement Deque datastructures. Previously to make a Deque structure the array functions `array_shift/unshift/pop/push` had to be used instead.

GHC's `Data.Sequence` <sup>[2]</sup> module implements an efficient, functional deque structure in Haskell. The implementation uses 2-3 finger trees annotated with sizes. There are other (fast) possibilities to implement purely functional (thus also persistent) double queues (most using heavily lazy evaluation), see references, <sup>[3][4]</sup>, <sup>[5]</sup>.

## Complexity

- In a doubly linked list implementation and assuming no allocation/deallocation overhead, the time complexity of all deque operations is  $O(1)$ . Additionally, the time complexity of insertion or deletion in the middle, given an iterator, is  $O(1)$ ; however, the time complexity of random access by index is  $O(n)$ .
- In a growing array, the amortized time complexity of all deque operations is  $O(1)$ . Additionally, the time complexity of random access by index is  $O(1)$ ; but the time complexity of insertion or deletion in the middle is  $O(n)$ .

## Applications

One example where a deque can be used is the *A-Steal* job scheduling algorithm.<sup>[6]</sup> This algorithm implements task scheduling for several processors. A separate deque with threads to be executed is maintained for each processor. To execute the next thread, the processor gets the first element from the deque (using the "remove first element" deque operation). If the current thread forks, it is put back to the front of the deque ("insert element at front") and a new thread is executed. When one of the processors finishes execution of its own threads (i.e. its deque is empty), it can "steal" a thread from another processor: it gets the last element from the deque of another processor ("remove last element") and executes it.

## References

- [1] Donald Knuth. *The Art of Computer Programming*, Volume 1: *Fundamental Algorithms*, Third Edition. Addison-Wesley, 1997. ISBN 0-201-89683-4. Section 2.2.1: Stacks, Queues, and Deques, pp. 238–243.
- [2] <http://www.haskell.org/ghc/docs/latest/html/libraries/containers/Data-Sequence.html>
- [3] [www.cs.cmu.edu/~rwh/theses/okasaki.pdf](http://www.cs.cmu.edu/~rwh/theses/okasaki.pdf) C. Okasaki, "Purely Functional Data Structures", September 1996
- [4] Adam L. Buchsbaum and Robert E. Tarjan. Confluently persistent deques via data structural bootstrapping. *Journal of Algorithms*, 18(3):513–547, May 1995. (pp. 58, 101, 125)
- [5] Haim Kaplan and Robert E. Tarjan. Purely functional representations of catenable sorted lists. In *ACM Symposium on Theory of Computing*, pages 202–211, May 1996. (pp. 4, 82, 84, 124)
- [6] Eitan Frachtenberg, Uwe Schwiegelshohn (2007). *Job Scheduling Strategies for Parallel Processing: 12th International Workshop, JSSPP 2006*. Springer. ISBN 3-540-71034-5. See p.22.

## External links

- SGI STL Documentation: deque<T, Alloc> (<http://www.sgi.com/tech/stl/Deque.html>)
- Code Project: An In-Depth Study of the STL Deque Container ([http://www.codeproject.com/KB/stl/vector\\_vs\\_deque.aspx](http://www.codeproject.com/KB/stl/vector_vs_deque.aspx))
- Diagram of a typical STL deque implementation (<http://pages.cpsc.ucalgary.ca/~kremer/STL/1024x768/deque.html>)
- Deque implementation in C (<http://www.martinbroadhurst.com/articles/deque.html>)
- VBScript implementation of stack, queue, deque, and Red-Black Tree (<http://www.ludvikjerabek.com/downloads.html>)

# Priority queue

---

In computer science, a **priority queue** is an abstract data type which is like a regular queue or stack data structure, but where additionally each element has a "priority" associated with it. In a priority queue, an element with high priority is served before an element with low priority. If two elements have the same priority, they are served according to their order in the queue.

- *stack* — elements are pulled in last-in first-out-order (e.g. a stack of papers)
- *queue* — elements are pulled in first-in first-out-order (e.g. a line in a cafeteria)

It is a common misconception that a priority queue is a heap. A priority queue is an abstract concept like "a list" or "a map"; just as a list can be implemented with a linked list or an array, a priority queue can be implemented with a heap or a variety of other methods.

A priority queue must at least support the following operations:

- `insert_with_priority`: add an element to the queue with an associated priority
- `pull_highest_priority_element`: remove the element from the queue that has the *highest priority*, and return it (also known as "pop\_element(Off)", "get\_maximum\_element", or "get\_front(most)\_element"; some conventions consider lower priorities to be higher, so this may also be known as "get\_minimum\_element", and is often referred to as "get-min" in the literature; the literature also sometimes implement separate "peek\_at\_highest\_priority\_element" and "delete\_element" functions, which can be combined to produce "pull\_highest\_priority\_element")

More advanced implementations may support more complicated operations, such as *pull\_lowest\_priority\_element*, inspecting the first few highest- or lowest-priority elements (peeking at the highest priority element can be made  $O(1)$  time in nearly all implementations), clearing the queue, clearing subsets of the queue, performing a batch insert, merging two or more queues into one, incrementing priority of any element, etc.

## Similarity to queues

One can imagine a priority queue as a modified queue, but when one would get the next element off the queue, the highest-priority element is retrieved first.

Stacks and queues may be modeled as particular kinds of priority queues. In a stack, the priority of each inserted element is monotonically increasing; thus, the last element inserted is always the first retrieved. In a queue, the priority of each inserted element is monotonically decreasing; thus, the first element inserted is always the first retrieved.

## Implementation

### Naive implementations

There are a variety of simple, usually inefficient, ways to implement a priority queue. They provide an analogy to help one understand what a priority queue is. For instance, one can keep all the elements in an unsorted list. Whenever the highest-priority element is requested, search through all elements for the one with the highest priority. (In big O notation:  $O(1)$  insertion time,  $O(n)$  pull time due to search.)

### Usual implementation

To improve performance, priority queues typically use a heap as their backbone, giving  $O(\log n)$  performance for inserts and removals, and  $O(n)$  to build initially. Alternatively, when a self-balancing binary search tree is used, insertion and removal also take  $O(\log n)$  time, although building trees from existing sequences of elements takes  $O(n)$

---

$\log n$ ) time; this typical where one might already have access to these data structures, such as with third-party or standard libraries.

Note that from a computational-complexity standpoint, priority queues are congruent to sorting algorithms. See the next section for how efficient sorting algorithms can create efficient priority queues.

There are several specialized heap data structures that either supply additional operations or outperform these approaches. The binary heap uses  $O(\log n)$  time for both operations, but also allow queries of the element of highest priority without removing it in constant time. Binomial heaps add several more operations, but require  $O(\log n)$  time for requests. Fibonacci heaps can insert elements, query the highest priority element, and increase an element's priority in amortized constant time<sup>[1]</sup> though deletions are still  $O(\log n)$ ). Brodal queues can do this in worst-case constant time.

While relying on a heap is a common way to implement priority queues, for integer data, faster implementations exist. This can even apply to data-types that have a finite range, such as floats:

- When the set of keys is  $\{1, 2, \dots, C\}$ , a van Emde Boas tree would support the *minimum*, *maximum*, *insert*, *delete*, *search*, *extract-min*, *extract-max*, *predecessor* and *successor* operations in  $O(\log \log C)$  time, but has a space cost for small queues of about  $O(2^{m/2})$ , where  $m$  is the number of bits in the priority value.<sup>[2]</sup>
- The Fusion tree algorithm by Fredman and Willard implements the *minimum* operation in  $O(1)$  time and *insert* and *extract-min* operations in  $O(\sqrt{\log n})$  time.<sup>[3]</sup>

For applications that do many "peek" operations for every "extract-min" operation, the time complexity for peek actions can be reduced to  $O(1)$  in all tree and heap implementations by caching the highest priority element after every insertion and removal. For insertion, this adds at most a constant cost, since the newly inserted element is compared only to the previously cached minimum element. For deletion, this at most adds an additional "peek" cost, which is typically cheaper than the deletion cost, so overall time complexity is not significantly impacted.

## Equivalence of priority queues and sorting algorithms

### Using a priority queue to sort

The semantics of priority queues naturally suggest a sorting method: insert all the elements to be sorted into a priority queue, and sequentially remove them; they will come out in sorted order. This is actually the procedure used by several sorting algorithms, once the layer of abstraction provided by the priority queue is removed. This sorting method is equivalent to the following sorting algorithms:

- Heapsort if the priority queue is implemented with a heap.
- Smoothsort if the priority queue is implemented with a Leonardo heap.
- Selection sort if the priority queue is implemented with an unordered array.
- Insertion sort if the priority queue is implemented with an ordered array.
- Tree sort if the priority queue is implemented with a self-balancing binary search tree.

## Using a sorting algorithm to make a priority queue

A sorting algorithm can also be used to implement a priority queue. Specifically, Thorup says<sup>[4]</sup>:

We present a general deterministic linear space reduction from priority queues to sorting implying that if we can sort up to  $n$  keys in  $S(n)$  time per key, then there is a priority queue supporting *delete* and *insert* in  $O(S(n))$  time and find-min in constant time.

That is, if there is a sorting algorithm which can sort in  $O(S)$  time per key, where  $S$  is some function of  $n$  and word size<sup>[5]</sup>, then one can use the given procedure to create a priority queue where pulling the highest-priority element is  $O(1)$  time, and inserting new elements (and deleting elements) is  $O(S)$  time. For example if one has an  $O(n \lg(\lg(n)))$  sort algorithm, one can easily create a priority queue with  $O(1)$  pulling and  $O(\lg(\lg(n)))$  insertion.

## Libraries

A priority queue is often considered to be a "container data structure".

The Standard Template Library (STL), and the C++ 1998 standard, specifies `priority_queue` as one of the STL container adaptor class templates. It implements a max-priority-queue. Unlike actual STL containers, it does not allow iteration of its elements (it strictly adheres to its abstract data type definition). STL also has utility functions for manipulating another random-access container as a binary max-heap. The Boost\_(C++\_libraries) also have an implementation in the library heap.

Python's `heapq`<sup>[6]</sup> module implements a binary min-heap on top of a list.

Java's library contains a `PriorityQueue` class, which implements a min-priority-queue.

Go's library contains a `container/heap`<sup>[7]</sup> module, which implements a min-heap on top of any compatible data structure.

The Standard PHP Library extension contains the class `SplPriorityQueue`<sup>[8]</sup>.

Apple's Core Foundation framework contains a `CFBinaryHeap`<sup>[9]</sup> structure, which implements a min-heap.

## Applications

### Bandwidth management

Priority queuing can be used to manage limited resources such as bandwidth on a transmission line from a network router. In the event of outgoing traffic queuing due to insufficient bandwidth, all other queues can be halted to send the traffic from the highest priority queue upon arrival. This ensures that the prioritized traffic (such as real-time traffic, e.g. an RTP stream of a VoIP connection) is forwarded with the least delay and the least likelihood of being rejected due to a queue reaching its maximum capacity. All other traffic can be handled when the highest priority queue is empty. Another approach used is to send disproportionately more traffic from higher priority queues.

Many modern protocols for Local Area Networks also include the concept of Priority Queues at the Media Access Control (MAC) sub-layer to ensure that high-priority applications (such as VoIP or IPTV) experience lower latency than other applications which can be served with Best effort service. Examples include IEEE 802.11e (an amendment to IEEE 802.11 which provides Quality of Service) and ITU-T G.hn (a standard for high-speed Local area network using existing home wiring (power lines, phone lines and coaxial cables)).

Usually a limitation (policer) is set to limit the bandwidth that traffic from the highest priority queue can take, in order to prevent high priority packets from choking off all other traffic. This limit is usually never reached due to high level control instances such as the Cisco Callmanager, which can be programmed to inhibit calls which would exceed the programmed bandwidth limit.

## Discrete event simulation

Another use of a priority queue is to manage the events in a discrete event simulation. The events are added to the queue with their simulation time used as the priority. The execution of the simulation proceeds by repeatedly pulling the top of the queue and executing the event thereon.

*See also:* Scheduling (computing), queueing theory

## Dijkstra's algorithm

When the graph is stored in the form of adjacency list or matrix, priority queue can be used to extract minimum efficiently when implementing Dijkstra's algorithm, although one also needs the ability to alter the priority of a particular vertex in the priority queue efficiently.

## Huffman coding

Huffman coding requires one to repeatedly obtain the two lowest-frequency trees. A priority queue makes this efficient.

## A\* and SMA\* search algorithms

The A\* search algorithm finds the shortest path between two vertices or nodes of a weighted graph, trying out the most promising routes first. The priority queue (also known as the *fringe*) is used to keep track of unexplored routes; the one for which a lower bound on the total path length is smallest is given highest priority. If memory limitations make A\* impractical, the SMA\* algorithm can be used instead, with a double-ended priority queue to allow removal of low-priority items.

## ROAM triangulation algorithm

The Real-time Optimally Adapting Meshes (ROAM) algorithm computes a dynamically changing triangulation of a terrain. It works by splitting triangles where more detail is needed and merging them where less detail is needed. The algorithm assigns each triangle in the terrain a priority, usually related to the error decrease if that triangle would be split. The algorithm uses two priority queues, one for triangles that can be split and another for triangles that can be merged. In each step the triangle from the split queue with the highest priority is split, or the triangle from the merge queue with the lowest priority is merged with its neighbours.

## References

- [1] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7. Chapter 20: Fibonacci Heaps, pp.476–497. Third edition p518.
- [2] P. van Emde Boas. Preserving order in a forest in less than logarithmic time. In *Proceedings of the 16th Annual Symposium on Foundations of Computer Science*, pages 75–84. IEEE Computer Society, 1975.
- [3] Michael L. Fredman and Dan E. Willard. Surpassing the information theoretic bound with fusion trees. *Journal of Computer and System Sciences*, 48(3):533–551, 1994.
- [4] Mikkel Thorup. 2007. Equivalence between priority queues and sorting. J. ACM 54, 6, Article 28 (December 2007). DOI=10.1145/1314690.1314692 (<http://doi.acm.org/10.1145/1314690.1314692>)
- [5] <http://courses.csail.mit.edu/6.851/spring07/scribe/lec17.pdf>
- [6] <http://docs.python.org/library/heapq.html>
- [7] <http://golang.org/pkg/container/heap/>
- [8] <http://us2.php.net/manual/en/class.splpriorityqueue.php>
- [9] <http://developer.apple.com/library/mac/#documentation/CoreFoundation/Reference/CFBinaryHeapRef/Reference/reference.html>



## Further reading

- Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7. Section 6.5: Priority queues, pp.138–142.

## External links

- C++ reference for `std::priority_queue` ([http://en.cppreference.com/w/cpp/container/priority\\_queue](http://en.cppreference.com/w/cpp/container/priority_queue))
- Descriptions (<http://leekillough.com/heaps/>) by Lee Killough
- PQlib (<http://bitbucket.org/trijezdci/pqlib/src/>) - Open source Priority Queue library for C
- libpqueue (<http://github.com/vy/libpqueue>) is a generic priority queue (heap) implementation (in C) used by the Apache HTTP Server project.
- Survey of known priority queue structures (<http://www.theturingmachine.com/algorithms/heaps.html>) by Stefan Xenos
- UC Berkeley - Computer Science 61B - Lecture 24: Priority Queues (<http://video.google.com/videoplay?docid=3499489585174920878>) (video) - introduction to priority queues using binary heap
- Double-Ended Priority Queues (<http://www.cise.ufl.edu/~sahni/dsaa/enrich/c13/double.htm>) by Sartaj Sahni
- Priority queues and min priority queues in C++ (<http://www.technical-recipes.com/2011/priority-queues-and-min-priority-queues-in-c/>)

# Map

---

In computer science, an **associative array**, **map**, or **dictionary** is an abstract data type composed of a collection of (key,value) pairs, such that each possible key appears at most once in the collection.

Operations associated with this data type allow:<sup>[1][2]</sup>

- the addition of pairs to the collection
- the removal of pairs from the collection
- the modification of the values of existing pairs
- the lookup of the value associated with a particular key

The **dictionary problem** is the task of designing a data structure that implements an associative array. A standard solution to the dictionary problem is a hash table; in some cases it is also possible to solve the problem using directly addressed arrays, binary search trees, or other more specialized structures.<sup>[1][2][3]</sup>

Many programming languages include associative arrays as primitive data types, and they are available in software libraries for many others. Content-addressable memory is a form of direct hardware-level support for associative arrays.

Associative arrays have many applications including such fundamental programming patterns as memoization and the decorator pattern.<sup>[4]</sup>

---

## Operations

In an associative array, the association between a key and a value is often known as a "binding", and the same word "binding" may also be used to refer to the process of creating a new association.

The operations that are usually defined for an associative array are:<sup>[1][2]</sup>

- **Add** or **insert**: add a new {key, value} pair to the collection, binding the new key to its new value. The arguments to this operation are the key and the value.
- **Reassign**: replace the value in one of the (key,value) pairs that are already in the collection, binding an old key to a new value. As with an insertion, the arguments to this operation are the key and the value.
- **Remove** or **delete**: remove a (key,value) pair from the collection, unbinding a given key from its value. The argument to this operation is the key.
- **Lookup**: find the value (if any) that is bound to a given key. The argument to this operation is the key, and the value is returned from the operation. If no value is found, some associative array implementations raise an exception.

In addition, associative arrays may also include other operations such as determining the number of bindings or constructing an iterator to loop over all the bindings. Usually, for such an operation, the order in which the bindings are returned may be arbitrary.

A multimap generalizes an associative array by allowing multiple values to be associated with a single key.<sup>[5]</sup> A bidirectional map is a related abstract data type in which the bindings operate in both directions: each value must be associated with a unique key, and a second lookup operation takes a value as argument and looks up the key associated with that value.

## Example

Suppose that the set of loans made by a library is to be represented in a data structure. Each book in a library may be checked out only by a single library patron at a time. However, a single patron may be able to check out multiple books. Therefore, the information about which books are checked out to which patrons may be represented by an associative array, in which the books are the keys and the patrons are the values. For instance (using the notation from Python in which a binding is represented by placing a colon between the key and the value), the current checkouts may be represented by an associative array

```
{
    "Great Expectations": "John",
    "Pride and Prejudice": "Alice",
    "Wuthering Heights": "Alice"
}
```

A lookup operation with the key "Great Expectations" in this array would return the name of the person who checked out that book, John. If John returns his book, that would cause a deletion operation in the associative array, and if Pat checks out another book, that would cause an insertion operation, leading to a different state:

```
{
    "Pride and Prejudice": "Alice",
    "The Brothers Karamazov": "Pat",
    "Wuthering Heights": "Alice"
}
```

In this new state, the same lookup as before, with the key "Great Expectations", would raise an exception, because this key is no longer present in the array.

## Implementation

For dictionaries with very small numbers of bindings, it may make sense to implement the dictionary using an association list, a linked list of bindings. With this implementation, the time to perform the basic dictionary operations is linear in the total number of bindings; however, it is easy to implement and the constant factors in its running time are small.<sup>[1][6]</sup> Another very simple implementation technique, usable when the keys are restricted to a narrow range of integers, is direct addressing into an array: the value for a given key  $k$  is stored at the array cell  $A[k]$ , or if there is no binding for  $k$  then the cell stores a special sentinel value that indicates the absence of a binding. As well as being simple, this technique is fast: each dictionary operation takes constant time. However, the space requirement for this structure is the size of the entire keyspace, making it impractical unless the keyspace is small.<sup>[3]</sup>

The most frequently used general purpose implementation of an associative array is with a hash table: an array of bindings, together with a hash function that maps each possible key into an array index. The basic idea of a hash table is that the binding for a given key is stored at the position given by applying the hash function to that key, and that lookup operations are performed by looking at that cell of the array and using the binding found there. However, hash table based dictionaries must be prepared to handle collisions that occur when two keys are mapped by the hash function to the same index, and many different collision resolution strategies have been developed for dealing with this situation, often based either on open addressing (looking at a sequence of hash table indices instead of a single index, until finding either the given key or an empty cell) or on hash chaining (storing a small association list instead of a single binding in each hash table cell).<sup>[1][2][3]</sup>

Dictionaries may also be stored in binary search trees or in data structures specialized to a particular type of keys such as radix trees, tries, Judy arrays, or van Emde Boas trees, but these implementation methods are less efficient than hash tables as well as placing greater restrictions on the types of data that they can handle. The advantages of these alternative structures come from their ability to handle operations beyond the basic ones of an associative array, such as finding the binding whose key is the closest to a queried key, when the query is not itself present in the set of bindings.

## Language support

Associative arrays can be implemented in any programming language as a package and many language systems provide them as part of their standard library. In some languages, they are not only built into the standard system, but have special syntax, often using array-like subscripting.

Built-in syntactic support for associative arrays was introduced by SNOBOL4, under the name "table". MUMPS made multi-dimensional associative arrays, optionally persistent, its key data structure. SETL supported them as one possible implementation of sets and maps. Most modern scripting languages, starting with AWK and including Perl, Tcl, JavaScript, Python, Ruby, and Lua, support associative arrays as a primary container type. In many more languages, they are available as library functions without special syntax.

In Smalltalk, Objective-C, .NET<sup>[7]</sup>, Python, and REALbasic they are called *dictionaries*; in Perl and Ruby they are called *hashes*; in C++, Java, and Go they are called *maps* (see `map` (C++), `unordered_map` (C++), and `Map`); in Common Lisp and Windows PowerShell, they are called *hash tables* (since both typically use this implementation). In PHP, all arrays can be associative, except that the keys are limited to integers and strings. In JavaScript (see also JSON), all objects behave as associative arrays. In Lua, they are called *tables*, and are used as the primitive building block for all data structures. In Visual FoxPro, they are called *Collections*.

## References

- [1] Goodrich, Michael T.; Tamassia, Roberto (2006), "9.1 The Map Abstract Data Type", *Data Structures & Algorithms in Java* (4th ed.), Wiley, pp. 368–371.
- [2] Mehlhorn, Kurt; Sanders, Peter (2008), "4 Hash Tables and Associative Arrays", *Algorithms and Data Structures: The Basic Toolbox*, Springer, pp. 81–98.
- [3] Cormen, Thomas H.; Leiserson, Charles E.; Rivest, Ronald L.; Stein, Clifford (2001), "11 Hash Tables", *Introduction to Algorithms* (2nd ed.), MIT Press and McGraw-Hill, pp. 221–252, ISBN 0-262-03293-7.
- [4] Goodrich & Tamassia (2006), pp. 597–599.
- [5] Goodrich & Tamassia (2006), pp. 389–397.
- [6] "When should I use a hash table instead of an association list?" (<http://www.faqs.org/faqs/lisp-faq/part2/section-2.html>). lisp-faq/part2. 1996-02-20. .
- [7] "Dictionary<TKey, TValue> Class" (<http://msdn.microsoft.com/en-us/library/xfhwa508.aspx>). MSDN. .

## External links

- NIST's Dictionary of Algorithms and Data Structures: Associative Array (<http://www.nist.gov/dads/HTML/assocarray.html>)

# Bidirectional map

---

In computer science, a **bidirectional map** is an associative data structure in which both types can be used as key.

## External links

- [http://www.boost.org/doc/libs/1\\_47\\_0/libs/bimap/doc/html/index.html](http://www.boost.org/doc/libs/1_47_0/libs/bimap/doc/html/index.html)
- <http://commons.apache.org/collections/api-release/org/apache/commons/collections/BidiMap.html>
- [http://cablemodem.fibertel.com.ar/mcape/oss/projects/mc\\_projects/boost\\_projects/boost\\_bimap.html#](http://cablemodem.fibertel.com.ar/mcape/oss/projects/mc_projects/boost_projects/boost_bimap.html#)
- <http://www.codeproject.com/KB/stl/bimap.aspx>
- <http://guava-libraries.googlecode.com/svn/tags/release09/javadoc/com/google/common/collect/BiMap.html>

# Multimap

---

In computer science, a **multimap** (sometimes also **multihash**) is a generalization of a map or associative array abstract data type in which more than one value may be associated with and returned for a given key. Both map and multimap are particular cases of containers (see for example C++ Standard Template Library containers). Often the multimap is implemented as a map with lists or sets as the map values.

## Examples

- In a student enrollment system, where students may be enrolled in multiple classes simultaneously, there might be an association for each enrollment of a student in a course, where the key is the student ID and the value is the course ID. If a student is enrolled in three courses, there will be three associations containing the same key.
- The index of a book may report any number of references for a given index term, and thus may be coded as a multimap from index terms to any number of reference locations.
- Querystrings may have multiple values associated with a single field. This is commonly generated when a web form allows multiple check boxes or selections to be chosen in response to a single form element.

## Language support

C++'s Standard Template Library provides the `multimap` container for the sorted multimap using a self-balancing binary search tree,<sup>[1]</sup> and SGI's STL extension provides the `hash_multimap` container, which implements a multimap using a hash table.<sup>[2]</sup>

Apache Commons Collections provides a `MultiMap` interface for Java.<sup>[3]</sup> It also provides a `MultiValueMap` implementing class that makes a `MultiMap` out of a `Map` object and a type of `Collection`.<sup>[4]</sup>

Google Collections also provides an interface `Multimap` and implementations.<sup>[5]</sup>

The Scala programming language's API also provides `Multimap` and implementations<sup>[6]</sup>

## References

- [1] "`multimap<Key, Data, Compare, Alloc>`" (<http://www.sgi.com/tech/stl/Multimap.html>). *Standard Template Library Programmer's Guide*. Silicon Graphics International. .
  - [2] "`hash_multimap<Key, HashFcn, EqualKey, Alloc>`" ([http://www.sgi.com/tech/stl/hash\\_multimap.html](http://www.sgi.com/tech/stl/hash_multimap.html)). *Standard Template Library Programmer's Guide*. Silicon Graphics International. .
  - [3] "Interface `MultiMap`" (<http://commons.apache.org/collections/api-release/org/apache/commons/collections/MultiMap.html>). *Commons Collections 3.2.1 API, Apache Commons*. .
  - [4] "Class `MultiValueMap`" (<http://commons.apache.org/collections/api-release/org/apache/commons/collections/map/MultiValueMap.html>). *Commons Collections 3.2.1 API, Apache Commons*. .
  - [5] "Interface `Multimap<K,V>`" (<http://google-collections.googlecode.com/svn/trunk/javadoc/com/google/common/collect/Multimap.html>). *Google Collections Library 1.0*. .
  - [6] "`Scala.collection.mutable.MultiMap`" (<http://www.scala-lang.org/api/current/scala/collection/mutable/MultiMap.html>). *Scala stable API*. .
-

# Set

---

In computer science, a **set** is an abstract data structure that can store certain values, without any particular order, and no repeated values. It is a computer implementation of the mathematical concept of a finite set. Unlike most other collection types, rather than retrieving a specific element from a set, one typically tests a value for membership in a set.

Some set data structures are designed for **static** or **frozen sets** that do not change after they are constructed. Static sets allow only query operations on their elements — such as checking whether a given value is in the set, or enumerating the values in some arbitrary order. Other variants, called **dynamic** or **mutable sets**, allow also the insertion and/or deletion of elements from the set.

An abstract data structure is a collection, or aggregate, of data. The data may be booleans, numbers, characters, or other data structures. If one considers the structure yielded by packaging<sup>[1]</sup> or indexing,<sup>[2]</sup> there are four basic data structures.<sup>[3][4]</sup>

1. unpackaged, unindexed: bunch
2. packaged, unindexed: set
3. unpackaged, indexed: string (sequence)
4. packaged, indexed: list (array)

In this view, the contents of a set are a bunch, and isolated data items are elementary bunches (elements). Whereas sets *contain* elements, bunches *consist of* elements.

Further structuring may be achieved by considering the multiplicity of elements (sets become multisets, bunches become hyperbunches)<sup>[5]</sup> or their homogeneity (a record is a set of fields, not necessarily all of the same type).

## Implementations

A set can be implemented in many ways. For example, one can use a list, ignoring the order of the elements and taking care to avoid repeated values. Sets are often implemented using various flavors of trees, tries, or hash tables.

A set can be seen, and implemented, as a (partial) associative array, in which the value of each key-value pair has the unit type.

## Type theory

In type theory, sets are generally identified with their indicator function: accordingly, a set of values of type  $A$  may be denoted by  $2^A$  or  $\mathcal{P}(A)$ . (Subtypes and subsets may be modeled by refinement types, and quotient sets may be

replaced by setoids.) The characteristic function  $F$  of a set  $S$  is defined as:  $F(x) = \begin{cases} 1, & \text{if } x \in S \\ 0, & \text{if } x \notin S \end{cases}$

In theory, many other abstract data structures can be viewed as set structures with additional operations and/or additional axioms imposed on the standard operations. For example, an abstract heap can be viewed as a set structure with a  $\min(S)$  operation that returns the element of smallest value.

---

## Operations

### Core set-theoretical operations

One may define the operations of the algebra of sets:

- `union( $S, T$ )`: returns the union of sets  $S$  and  $T$ .
- `intersection( $S, T$ )`: returns the intersection of sets  $S$  and  $T$ .
- `difference( $S, T$ )`: returns the difference of sets  $S$  and  $T$ .
- `subset( $S, T$ )`: a predicate that tests whether the set  $S$  is a subset of set  $T$ .

### Static sets

Typical operations that may be provided by a static set structure  $S$  are:

- `is_element_of( $x, S$ )`: checks whether the value  $x$  is in the set  $S$ .
- `is_empty( $S$ )`: checks whether the set  $S$  is empty.
- `size( $S$ )` or `cardinality( $S$ )`: returns the number of elements in  $S$ .
- `iterate( $S$ )`: returns a function that returns one more value of  $S$  at each call, in some arbitrary order.
- `enumerate( $S$ )`: returns a list containing the elements of  $S$  in some arbitrary order.
- `build( $x_1, x_2, \dots, x_n$ )`: creates a set structure with values  $x_1, x_2, \dots, x_n$ .
- `create_from( $collection$ )`: creates a new set structure containing all the elements of the given collection or all the elements returned by the given iterator.

### Dynamic sets

Dynamic set structures typically add:

- `create()`: creates a new, initially empty set structure.
  - `create_with_capacity( $n$ )`: creates a new set structure, initially empty but capable of holding up to  $n$  elements.
- `add( $S, x$ )`: adds the element  $x$  to  $S$ , if it is not present already.
- `remove( $S, x$ )`: removes the element  $x$  from  $S$ , if it is present.
- `capacity( $S$ )`: returns the maximum number of values that  $S$  can hold.

Some set structures may allow only some of these operations. The cost of each operation will depend on the implementation, and possibly also on the particular values stored in the set, and the order in which they are inserted.

### Additional operations

There are many other operations that can (in principle) be defined in terms of the above, such as:

- `pop( $S$ )`: returns an arbitrary element of  $S$ , deleting it from  $S$ .
- `map( $F, S$ )`: returns the set of distinct values resulting from applying function  $F$  to each element of  $S$ .
- `filter( $P, S$ )`: returns the subset containing all elements of  $S$  that satisfy a given predicate  $P$ .
- `fold( $A_0, F, S$ )`: returns the value  $A_{|S|}$  after applying  $A_{i+1} := F(A_i, e)$  for each element  $e$  of  $S$ .
- `clear( $S$ )`: delete all elements of  $S$ .
- `equal( $S_1, S_2$ )`: checks whether the two given sets are equal (i.e. contain all and only the same elements).
- `hash( $S$ )`: returns a hash value for the static set  $S$  such that if `equal( $S_1, S_2$ )` then `hash( $S_1$ ) = hash( $S_2$ )`.

Other operations can be defined for sets with elements of a special type:

- `sum( $S$ )`: returns the sum of all elements of  $S$  for some definition of "sum". For example, over integers or reals, it may be defined as `fold(0, add,  $S$ )`.
- `nearest( $S, x$ )`: returns the element of  $S$  that is closest in value to  $x$  (by some metric).

## Implementations

Sets can be implemented using various data structures, which provide different time and space trade-offs for various operations. Some implementations are designed to improve the efficiency of very specialized operations, such as `nearest` or `union`. Implementations described as "general use" typically strive to optimize the `element_of`, `add`, and `delete` operations.

Sets are commonly implemented in the same way as associative arrays, namely, a self-balancing binary search tree for sorted sets (which has  $O(\log n)$  for most operations), or a hash table for unsorted sets (which has  $O(1)$  average-case, but  $O(n)$  worst-case, for most operations). A sorted linear hash table<sup>[6]</sup> may be used to provide deterministically ordered sets.

Other popular methods include arrays. In particular a subset of the integers  $1..n$  can be implemented efficiently as an  $n$ -bit bit array, which also support very efficient union and intersection operations. A Bloom map implements a set probabilistically, using a very compact representation but risking a small chance of false positives on queries.

The Boolean set operations can be implemented in terms of more elementary operations (`pop`, `clear`, and `add`), but specialized algorithms may yield lower asymptotic time bounds. If sets are implemented as sorted lists, for example, the naive algorithm for `union(S, T)` will take code proportional to the length  $m$  of  $S$  times the length  $n$  of  $T$ ; whereas a variant of the list merging algorithm will do the job in time proportional to  $m+n$ . Moreover, there are specialized set data structures (such as the union-find data structure) that are optimized for one or more of these operations, at the expense of others.

## Language support

One of the earliest languages to support sets was Pascal; many languages now include it, whether in the core language or in a standard library.

- Java offers the `Set` interface to support sets (with the `HashSet` class implementing it using a hash table), and the `SortedSet` sub-interface to support sorted sets (with the `TreeSet` class implementing it using a binary search tree).
- Apple's Foundation framework (part of Cocoa) provides the Objective-C classes `NSSet`<sup>[7]</sup>, `NSMutableSet`<sup>[8]</sup>, `NSCountedSet`<sup>[9]</sup>, `NSOrderedSet`<sup>[10]</sup>, and `NSMutableOrderedSet`<sup>[11]</sup>. The CoreFoundation APIs provide the `CFSet`<sup>[12]</sup> and `CFMutableSet`<sup>[13]</sup> types for use in C.
- Python has built-in `set` and `frozenset` types<sup>[14]</sup> since 2.4, and since Python 3.0 and 2.7, supports non-empty set literals using a curly-bracket syntax, e.g.: `{ x, y, z }`.
- The .NET Framework provides the generic `HashSet`<sup>[15]</sup> and `SortedSet`<sup>[16]</sup> classes that implement the generic `ISet`<sup>[17]</sup> interface.
- Smalltalk's class library includes `Set` and `IdentitySet`, using equality and identity for inclusion test respectively. Many dialects provide variations for compressed storage (`NumberSet`, `CharacterSet`), for ordering (`OrderedSet`, `SortedSet`, etc.) or for weak references (`WeakIdentitySet`).
- Ruby's standard library includes a `set`<sup>[18]</sup> module which contains `Set` and `SortedSet` classes that implement sets using hash tables, the latter allowing iteration in sorted order.
- OCaml's standard library contains a `Set` module, which implements a functional set data structure using binary search trees.
- The GHC implementation of Haskell provides a `Data.Set`<sup>[19]</sup> module, which implements a functional set data structure using binary search trees.
- The Tcl Tcllib package provides a `set` module which implements a set data structure based upon TCL lists.

As noted in the previous section, in languages which do not directly support sets but do support associative arrays, sets can be emulated using associative arrays, by using the elements as keys, and using a dummy value as the values, which are ignored.



## In C++

In C++, the Standard Template Library (STL) provides the `set` template class, which implements a sorted set using a binary search tree; SGI's STL also provides the `hash_set` template class, which implements a set using a hash table.

In sets, the elements themselves are the keys, in contrast to sequenced containers, where elements are accessed using their (relative or absolute) position. Set elements must have a strict weak ordering.

Some of the member functions in C++ and their description is given in the table below:

### **set** member functions

Signature(s)	Description
<code>iterator begin();</code>	Returns an iterator to the first element of the set.
<code>iterator end();</code>	Returns an iterator just before the end of the set.
<code>bool empty() const;</code>	Checks if the set container is empty (i.e. has a size of 0)
<code>iterator find(const key_type &amp;x) const;</code>	Searches the container for an element <code>x</code> and if found, returns an iterator to it, or else returns an iterator to <code>set::end</code>
<code>void insert ( Input Iterator first, Input Iterator last );</code> <code>pair&lt;iterator, bool&gt; insert ( const value_type&amp; a );</code> <code>iterator insert ( position(iterator), const value_type&amp; a );</code>	Inserts an element into the set. The first version returns pair, with its member <code>pair::first</code> set to an iterator pointing to either the newly inserted element or to the element that already had its same value in the set. The <code>pair::second</code> element in the pair is set to true if a new element was inserted or false if an element with the same value existed. And the second version returns an iterator either pointing to newly inserted element or to element having same value in set.
<code>void clear();</code>	Removes all elements in the set, making its size 0.

## Multiset

A variation of the set is the **multiset** or **bag**, which is the same as a set data structure, but allows repeated ("equal") values (duplicates). The set of all bags over type `T` is given by the expression `bag T`. It is possible for objects in computer science to be considered "equal" under some equivalence relation but still distinct under another relation. Some types of multiset implementations will store distinct equal objects as separate items in the data structure; while others will collapse it down to one version (the first one encountered) and keep a positive integer count of the multiplicity of the element.

- C++'s Standard Template Library provides the `multiset` class for the sorted multiset, and SGI's STL provides the `hash_multiset` class, which implements a multiset using a hash table.
- For Java, third-party libraries provide multiset functionality:
  - Apache Commons Collections provides the `Bag` <sup>[20]</sup> and `SortedBag` interfaces, with implementing classes like `HashBag` and `TreeBag`.
  - Google Collections provides the `Multiset` <sup>[21]</sup> interface, with implementing classes like `HashMultiset` and `TreeMultiset`.
- Apple provides the `NSCountedSet` <sup>[9]</sup> class as part of Cocoa, and the `CFBag` <sup>[22]</sup> and `CFMutableBag` <sup>[23]</sup> types as part of CoreFoundation.
- Python's standard library includes `collections.Counter` <sup>[24]</sup>, which is similar to a multiset.
- Smalltalk includes the `Bag` class, which can be instantiated to use either identity or equality as predicate for inclusion test.

Where a multiset data structure is not available, a workaround is to use a regular set, but override the equality predicate of its items to always return "not equal" on distinct objects (however, such will still not be able to store multiple occurrences of the same object) or use an associative array mapping the values to their integer multiplicities (this will not be able to distinguish between equal elements at all).

Typical operations on bags:

- Bag Membership - If  $B : \text{bag } T$  and  $x : T$  then the predicate  $x \text{ in } B$  is true if, and only if,  $x$  appears in  $B$  at least once.
- Sub-bags - If  $B_1, B_2 : \text{bag } T$  then the predicate  $B_1 \sqsubseteq B_2$  is true if each element that occurs in  $B_1$  occurs in  $B_2$  no more often than it occurs in  $B_2$ .
- Counting Bags - If  $B : \text{bag } T$  and  $x : T$  then the number of times  $x$  occurs in  $B$  (a natural number) is given by the expression  $B \# x$ .
- Scaling Bags - If  $B : \text{bag } T$  and  $n : \mathbf{N}$  then  $n \otimes B$  is a bag which contains the same elements as  $B$ , except that every element that occurs  $m$  times in  $B$  occurs  $n * m$  times in  $n \otimes B$ .
- Bag Union - If  $B_1, B_2 : \text{bag } T$  then  $B_1 \sqcup B_2$  is a bag that contains just those values that occur in either  $B_1$  or  $B_2$ , except that the number of times a value  $x$  occurs in  $B_1 \sqcup B_2$  is equal to  $(B_1 \# x) + (B_2 \# x)$ .

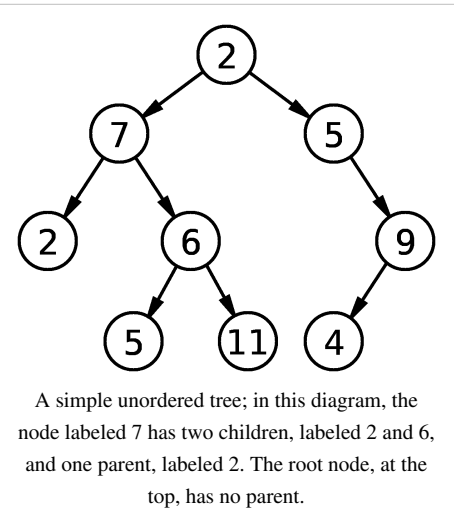
## References

- [1] "Packaging" consists in supplying a container for an aggregation of objects in order to turn them into a single object. Consider a function call: without packaging, a function can be called to act upon a bunch only by passing each bunch element as a separate argument, which complicates the function's signature considerably (and is just not possible in some programming languages). By packaging the bunch's elements into a set, the function may now be called upon a single, elementary argument: the set object (the bunch's package).
- [2] Indexing is possible when the elements being considered are totally ordered. Being without order, the elements of a multiset (for example) do not have lesser/greater or preceding/succeeding relationships: they can only be compared in absolute terms (same/different).
- [3] Hehner, Eric C. R. (1981), "Bunch Theory: A Simple Set Theory for Computer Science", *Information Processing Letters* **12** (1)
- [4] Hehner, Eric C. R. (2004), *A Practical Theory of Programming, second edition* (<http://www.cs.utoronto.ca/~hehner/aPToP/>),
- [5] Hehner, Eric C. R. (2012), *A Practical Theory of Programming, 2012-3-30 edition* (<http://www.cs.toronto.edu/~hehner/aPToP/>),
- [6] Wang, Thomas (1997), *Sorted Linear Hash Table* (<http://www.concentric.net/~Ttwang/tech/sorthash.htm>),
- [7] [http://developer.apple.com/documentation/Cocoa/Reference/Foundation/Classes/NSSet\\_Class/](http://developer.apple.com/documentation/Cocoa/Reference/Foundation/Classes/NSSet_Class/)
- [8] [http://developer.apple.com/documentation/Cocoa/Reference/Foundation/Classes/NSMutableSet\\_Class/](http://developer.apple.com/documentation/Cocoa/Reference/Foundation/Classes/NSMutableSet_Class/)
- [9] [http://developer.apple.com/documentation/Cocoa/Reference/Foundation/Classes/NSCountedSet\\_Class/](http://developer.apple.com/documentation/Cocoa/Reference/Foundation/Classes/NSCountedSet_Class/)
- [10] [http://developer.apple.com/library/mac/#documentation/Foundation/Reference/NSOrderedSet\\_Class/Reference/Reference.html](http://developer.apple.com/library/mac/#documentation/Foundation/Reference/NSOrderedSet_Class/Reference/Reference.html)
- [11] [https://developer.apple.com/library/mac/#documentation/Foundation/Reference/NSMutableOrderedSet\\_Class/Reference/Reference.html](https://developer.apple.com/library/mac/#documentation/Foundation/Reference/NSMutableOrderedSet_Class/Reference/Reference.html)
- [12] <http://developer.apple.com/documentation/CoreFoundation/Reference/CFSetRef/>
- [13] <http://developer.apple.com/documentation/CoreFoundation/Reference/CFMutableSetRef/>
- [14] <http://docs.python.org/library/stdtypes.html#set-types-set-frozenset>
- [15] <http://msdn.microsoft.com/en-us/library/bb359438.aspx>
- [16] <http://msdn.microsoft.com/en-us/library/dd412070.aspx>
- [17] <http://msdn.microsoft.com/en-us/library/dd412081.aspx>
- [18] <http://ruby-doc.org/stdlib/libdoc/set/rdoc/index.html>
- [19] <http://hackage.haskell.org/packages/archive/containers/0.2.0.1/doc/html/Data-Set.html>
- [20] <http://commons.apache.org/collections/api-release/org/apache/commons/collections/Bag.html>
- [21] <http://google-collections.googlecode.com/svn/trunk/javadoc/com/google/common/collect/Multiset.html>
- [22] <http://developer.apple.com/documentation/CoreFoundation/Reference/CFBagRef/>
- [23] <http://developer.apple.com/documentation/CoreFoundation/Reference/CFMutableBagRef/>
- [24] <http://docs.python.org/library/collections.html#collections.Counter>

# Tree

In computer science, a **tree** is a widely used data structure that simulates a hierarchical tree structure with a set of linked nodes.

A tree can be defined recursively (locally) as a collection of nodes (starting at a root node), where each node is a data structure consisting of a value, together with a list of nodes (the "children"), with the constraints that no node is duplicated. A tree can be defined abstractly as a whole (globally) as an ordered tree, with a value assigned to each node. Both these perspectives are useful: while a tree can be analyzed mathematically as a whole, when actually represented as a data structure it is usually represented and worked with separately by node (rather than as a list of nodes and an adjacency list of edges between nodes, as one may represent a digraph, for instance). For example, looking at a tree as a whole, one can talk about "the parent node" of a given node, but in general as a data structure a given node only contains the list of its children, but does not contain a reference to its parent (if any).



## Definition

### Recursive

Recursively, a tree is defined as a node (the root), which itself consists of a value (of some data type, possibly empty), together with a list of nodes (possibly empty); symbolically:

```
n: v [n[1], ..., n[k]]
```

(A node  $n$  consists of a value  $v$  and a list of other nodes.)

Note that this definition is in terms of values, and is appropriate in functional languages (it assumes referential transparency); different nodes have no connections, as they are simply lists of values. If the nodes are instead *references* to nodes (so one of the "nodes" could in fact point to an earlier node), one must specify that all the nodes are different, otherwise the above defines a directed graph,<sup>[1]</sup> possibly with loops.

Indeed, given a list of nodes, and for each node a list of its children, one cannot tell if this structure is a tree or not without analyzing its global structure and checking that it is in fact topologically a tree, as defined below. In terms of references, a tree is a special kind of directed graph, with a global constraint on its topology (namely no loops as an undirected graph).

## Mathematical

Viewed as a whole, a tree data structure is an ordered tree, generally with values attached to each node. Concretely, it is:

- A rooted tree with the "away from root" direction (a more narrow term is an "arborescence"), meaning:
  - A directed graph,
  - whose underlying undirected graph is a tree (any two vertices are connected by exactly one simple path),
  - with a distinguished root (one vertex is designated as the root),
  - which determines the direction on the edges (arrows point away from the root; given an edge, the node that the edge points from is called the *parent* and the node that the edge points to is called the *child*),

together with:

- an ordering on the child nodes of a given node, and
- a value (of some data type) at each node.

Often trees have a fixed (more properly, bounded) branching factor (outdegree), particularly always having two child nodes (possibly empty, hence *at most* two *non-empty* child nodes), hence a "binary tree".

## Terminology

A **node** is a structure which may contain a value or condition, or represent a separate data structure (which could be a tree of its own). Each node in a tree has zero or more **child nodes**, which are below it in the tree (by convention, trees are drawn growing downwards). A node that has a child is called the child's **parent node** (or *ancestor node*, or superior). A node has at most one parent.

An **internal node** (also known as an **inner node** or **branch node**) is any node of a tree that has child nodes. Similarly, an **external node** (also known as an **outer node**, **leaf node**, or **terminal node**) is any node that does not have child nodes.

The topmost node in a tree is called the **root node**. Being the topmost node, the root node will not have a parent. It is the node at which algorithms on the tree begin, since as a data structure, one can only pass from parents to children. Note that some algorithms (such as post-order depth-first search) begin at the root, but first visit leaf nodes (access the value of leaf nodes), only visit the root last (i.e., they first access the children of the root, but only access the *value* of the root last). All other nodes can be reached from it by following **edges** or **links**. (In the formal definition, each such path is also unique.) In diagrams, the root node is conventionally drawn at the top. In some trees, such as heaps, the root node has special properties. Every node in a tree can be seen as the root node of the subtree rooted at that node.

The **height** of a node is the length of the longest downward path to a leaf from that node. The height of the root is the height of the tree. The **depth** of a node is the length of the path to its root (i.e., its *root path*). This is commonly needed in the manipulation of the various self balancing trees, AVL Trees in particular. The root node has depth zero, leaf nodes have height zero, and a tree with only a single node (hence both a root and leaf) has depth and height zero. Conventionally, an empty tree (tree with no nodes) has depth and height  $-1$ .

A **subtree** of a tree  $T$  is a tree consisting of a node in  $T$  and all of its descendants in  $T$ .<sup>[2][3]</sup> Nodes thus correspond to subtrees (each node corresponds to the subtree of itself and all its descendants) – the subtree corresponding to the root node is the entire tree, and each node is the root node of the subtree it determines; the subtree corresponding to any other node is called a **proper subtree** (in analogy to the term proper subset).

## Drawing graphs

Trees are often drawn in the plane. Ordered trees can be represented essentially uniquely in the plane, and are hence called *plane trees*, as follows: if one fixes a conventional order (say, counterclockwise), and arranges the child nodes in that order (first incoming parent edge, then first child edge, etc.), this yields an embedding of the tree in the plane, unique up to ambient isotopy. Conversely, such an embedding determines an ordering of the child nodes.

If one places the root at the top (parents above children, as in a family tree) and places all nodes that are a given distance from the root (in terms of number of edges: the "level" of a tree) on a given horizontal line, one obtains a standard drawing of the tree. Given a binary tree, the first child is on the left (the "left node"), and the second child is on the right (the "right node").

## Representations

There are many different ways to represent trees; common representations represent the nodes as dynamically allocated records with pointers to their children, their parents, or both, or as items in an array, with relationships between them determined by their positions in the array (e.g., binary heap).

In general a node in a tree will not have pointers to its parents, but this information can be included (expanding the data structure to also include a pointer to the parent) or stored separately. Alternatively, upward links can be included in the child node data, as in a threaded binary tree.

## Generalizations

### Digraphs

If edges (to child nodes) are thought of as references, then a tree is a special case of a digraph, and the tree data structure can be generalized to represent directed graphs by removing the constraints that a node may have at most one parent, and that no cycles are allowed. Edges are still abstractly considered as pairs of nodes, however, the terms *parent* and *child* are usually replaced by different terminology (for example, *source* and *target*). Different implementation strategies exist: a digraph can be represented by the same local data structure as a tree (node with value and list of children), assuming that "list of children" is a list of references, or globally by such structures as adjacency lists.

In graph theory, a tree is a connected acyclic graph; unless stated otherwise, in graph theory trees and graphs are assumed undirected. There is no one-to-one correspondence between such trees and trees as data structure. We can take an arbitrary undirected tree, arbitrarily pick one of its vertices as the *root*, make all its edges directed by making them point away from the root node – producing an arborescence – and assign an order to all the nodes. The result corresponds to a tree data structure. Picking a different root or different ordering produces a different one.

### Forest

While a subtree of a tree is again a tree, a disjoint union of two trees is no longer a tree, as it has no common root. If one generalizes trees to an ordered list of trees, one obtains an (ordered) **forest**. Given a node in a tree, its children define an ordered forest (the union of subtrees given by all the children, or equivalently taking the subtree given by the node itself and erasing the root). Just as subtrees are natural for recursion (as in a depth-first search), forests are natural for corecursion (as in a breadth-first search).

Via mutual recursion, a forest can be defined as a list of trees (represented by root nodes), where a node (of a tree) consists of a value and a forest (its children):

```
f: [n[1], ..., n[k]]
n: v f
```

## Traversal methods

Stepping through the items of a tree, by means of the connections between parents and children, is called **walking the tree**, and the action is a **walk** of the tree. Often, an operation might be performed when a pointer arrives at a particular node. A walk in which each parent node is traversed before its children is called a **pre-order** walk; a walk in which the children are traversed before their respective parents are traversed is called a **post-order** walk; a walk in which a node's left subtree, then the node itself, and then finally its right sub tree are traversed is called an **in-order** traversal. (This last scenario, referring to exactly two subtrees, a left subtree and a right subtree, assumes specifically a binary tree.) A **level-order** walk effectively performs a breadth-first search search over the entirety of a tree; nodes are traversed level by level, where the root node is visited first, followed by its direct child nodes and their siblings, followed by its grandchild nodes and their siblings, etc., until all nodes in the tree have been traversed.

## Common operations

- Enumerating all the items
- Enumerating a section of a tree
- Searching for an item
- Adding a new item at a certain position on the tree
- Deleting an item
- **Pruning**: Removing a whole section of a tree
- **Grafting**: Adding a whole section to a tree
- Finding the root for any node

## Common uses

- Representing hierarchical data
- Storing data in a way that makes it easily searchable (see binary search tree and tree traversal)
- Representing sorted lists of data
- As a workflow for compositing digital images for visual effects
- Routing algorithms

## Notes

[1] Properly, a rooted, ordered directed graph.

[2] This is different from the formal definition of subtree used in graph theory, which is a subgraph that forms a tree – it need not include all descendants. For example, the root node by itself is a subtree in the graph theory sense, but not in the data structure sense (unless there are no descendants).

[3] Weisstein, Eric W., " Subtree (<http://mathworld.wolfram.com/Subtree.html>)" from MathWorld.

## References

- Donald Knuth. *The Art of Computer Programming: Fundamental Algorithms*, Third Edition. Addison-Wesley, 1997. ISBN 0-201-89683-4 . Section 2.3: Trees, pp. 308–423.
- Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7 . Section 10.4: Representing rooted trees, pp. 214–217. Chapters 12–14 (Binary Search Trees, Red-Black Trees, Augmenting Data Structures), pp. 253–320.

## External links

- Description (<http://www.nist.gov/dads/HTML/tree.html>) from the Dictionary of Algorithms and Data Structures
  - STL-like C++ tree class (<http://tree.phi-sci.com>)
  - Description of tree data structures from ideainfo.8m.com (<http://ideainfo.8m.com>)
  - flash actionscript 3 opensource implementation of tree and binary tree (<http://www.dpdk.nl/opensource>) — opensource library
  - WormWeb.org: Interactive Visualization of the *C. elegans* Cell Tree (<http://wormweb.org/celllineage>) - Visualize the entire cell lineage tree of the nematode *C. elegans* (javascript)
-

---

# Arrays

---

## Array data structure

---

In computer science, an **array data structure** or simply an **array** is a data structure consisting of a collection of *elements* (values or variables), each identified by at least one *array index* or *key*. An array is stored so that the position of each element can be computed from its index tuple by a mathematical formula.<sup>[1][2][3]</sup>

For example, an array of 10 integer variables, with indices 0 through 9, may be stored as 10 words at memory addresses 2000, 2004, 2008, ... 2036, so that the element with index  $i$  has the address  $2000 + 4 \times i$ .<sup>[4]</sup>

Arrays are analogous to the mathematical concepts of vectors, matrices, and tensors. Indeed, arrays with one or two indices are often called vectors or matrices, respectively. Arrays are often used to implement tables, especially lookup tables; the word *table* is sometimes used as a synonym of *array*.

Arrays are among the oldest and most important data structures, and are used by almost every program. They are also used to implement many other data structures, such as lists and strings. They effectively exploit the addressing logic of computers. In most modern computers and many external storage devices, the memory is a one-dimensional array of words, whose indices are their addresses. Processors, especially vector processors, are often optimized for array operations.

Arrays are useful mostly because the element indices can be computed at run time. Among other things, this feature allows a single iterative statement to process arbitrarily many elements of an array. For that reason, the elements of an array data structure are required to have the same size and should use the same data representation. The set of valid index tuples and the addresses of the elements (and hence the element addressing formula) are usually,<sup>[3][5]</sup> but not always,<sup>[2]</sup> fixed while the array is in use.

The term *array* is often used to mean array data type, a kind of data type provided by most high-level programming languages that consists of a collection of values or variables that can be selected by one or more indices computed at run-time. Array types are often implemented by array structures; however, in some languages they may be implemented by hash tables, linked lists, search trees, or other data structures.

The term is also used, especially in the description of algorithms, to mean associative array or "abstract array", a theoretical computer science model (an abstract data type or ADT) intended to capture the essential properties of arrays.

## History

The first digital computers used machine-language programming to set up and access array structures for data tables, vector and matrix computations, and for many other purposes. Von Neumann wrote the first array-sorting program (merge sort) in 1945, during the building of the first stored-program computer.<sup>[6]p. 159</sup> Array indexing was originally done by self-modifying code, and later using index registers and indirect addressing. Some mainframes designed in the 1960s, such as the Burroughs B5000 and its successors, had special instructions for array indexing that included index-bounds checking..

Assembly languages generally have no special support for arrays, other than what the machine itself provides. The earliest high-level programming languages, including FORTRAN (1957), COBOL (1960), and ALGOL 60 (1960), had support for multi-dimensional arrays, and so has C (1972). In C++ (1983), class templates exist for multi-dimensional arrays whose dimension is fixed at runtime<sup>[3][5]</sup> as well as for runtime-flexible arrays.<sup>[2]</sup>

---



## Applications

Arrays are used to implement mathematical vectors and matrices, as well as other kinds of rectangular tables. Many databases, small and large, consist of (or include) one-dimensional arrays whose elements are records.

Arrays are used to implement other data structures, such as heaps, hash tables, deques, queues, stacks, strings, and VLists.

One or more large arrays are sometimes used to emulate in-program dynamic memory allocation, particularly memory pool allocation. Historically, this has sometimes been the only way to allocate "dynamic memory" portably.

Arrays can be used to determine partial or complete control flow in programs, as a compact alternative to (otherwise repetitive), multiple `IF` statements. They are known in this context as control tables and are used in conjunction with a purpose built interpreter whose control flow is altered according to values contained in the array. The array may contain subroutine pointers (or relative subroutine numbers that can be acted upon by `SWITCH` statements) - that direct the path of the execution.

## Array element identifier and addressing formulas

When data objects are stored in an array, individual objects are selected by an index that is usually a non-negative scalar integer. Indices are also called subscripts. An index *maps* the array value to a stored object.

There are three ways in which the elements of an array can be indexed:

- **0** (*zero-based indexing*): The first element of the array is indexed by subscript of 0.<sup>[7]</sup>
- **1** (*one-based indexing*): The first element of the array is indexed by subscript of 1.<sup>[8]</sup>
- **n** (*n-based indexing*): The base index of an array can be freely chosen. Usually programming languages allowing *n-based indexing* also allow negative index values and other scalar data types like enumerations, or characters may be used as an array index.

Arrays can have multiple dimensions, thus it is not uncommon to access an array using multiple indices. For example a two dimensional array `A` with three rows and four columns might provide access to the element at the 2nd row and 4th column by the expression: `A[1, 3]` (in a row major language) and `A[3, 1]` (in a column major language) in the case of a zero-based indexing system. Thus two indices are used for a two dimensional array, three for a three dimensional array, and `n` for an `n` dimensional array.

The number of indices needed to specify an element is called the dimension, dimensionality, or rank of the array.

In standard arrays, each index is restricted to a certain range of consecutive integers (or consecutive values of some enumerated type), and the address of an element is computed by a "linear" formula on the indices.

## One-dimensional arrays

A one-dimensional array (or single dimension array) is a type of linear array. Accessing its elements involves a single subscript which can either represent a row or column index.

As an example consider the C declaration `int anArrayName[10];`

Syntax : `datatype anArrayname[sizeofArray];`

In the given example the array can contain 10 elements of any value available to the `int` type. In C, the array element indices are 0-9 inclusive in this case. For example, the expressions `anArrayName[0]`, and `anArrayName[9]` are the first and last elements respectively.

For a vector with linear addressing, the element with index  $i$  is located at the address  $B + c \cdot i$ , where  $B$  is a fixed *base address* and  $c$  a fixed constant, sometimes called the *address increment* or *stride*.

If the valid element indices begin at 0, the constant  $B$  is simply the address of the first element of the array. For this reason, the C programming language specifies that array indices always begin at 0; and many programmers will call

that element "zeroth" rather than "first".

However, one can choose the index of the first element by an appropriate choice of the base address  $B$ . For example, if the array has five elements, indexed 1 through 5, and the base address  $B$  is replaced by  $B - 30c$ , then the indices of those same elements will be 31 to 35. If the numbering does not start at 0, the constant  $B$  may not be the address of any element.

## Multidimensional arrays

For a two-dimensional array, the element with indices  $i, j$  would have address  $B + c \cdot i + d \cdot j$ , where the coefficients  $c$  and  $d$  are the *row* and *column address increments*, respectively.

More generally, in a  $k$ -dimensional array, the address of an element with indices  $i_1, i_2, \dots, i_k$  is

$$B + c_1 \cdot i_1 + c_2 \cdot i_2 + \dots + c_k \cdot i_k.$$

For example: `int a[3][2];`

This means that array `a` has 3 rows and 2 columns, and the array is of integer type. Here we can store 6 elements they are stored linearly but starting from first row linear then continuing with second row. The above array will be stored as  $a_{11}, a_{12}, a_{21}, a_{22}, a_{31}, a_{32}$ .

This formula requires only  $k$  multiplications and  $k-1$  additions, for any array that can fit in memory. Moreover, if any coefficient is a fixed power of 2, the multiplication can be replaced by bit shifting.

The coefficients  $c_k$  must be chosen so that every valid index tuple maps to the address of a distinct element.

If the minimum legal value for every index is 0, then  $B$  is the address of the element whose indices are all zero. As in the one-dimensional case, the element indices may be changed by changing the base address  $B$ . Thus, if a two-dimensional array has rows and columns indexed from 1 to 10 and 1 to 20, respectively, then replacing  $B$  by  $B + c_1 - 3 c_1$  will cause them to be renumbered from 0 through 9 and 4 through 23, respectively. Taking advantage of this feature, some languages (like FORTRAN 77) specify that array indices begin at 1, as in mathematical tradition; while other languages (like Fortran 90, Pascal and Algol) let the user choose the minimum value for each index.

## Dope vectors

The addressing formula is completely defined by the dimension  $d$ , the base address  $B$ , and the increments  $c_1, c_2, \dots, c_k$ . It is often useful to pack these parameters into a record called the array's *descriptor* or *stride vector* or *dope vector*.<sup>[2][3]</sup> The size of each element, and the minimum and maximum values allowed for each index may also be included in the dope vector. The dope vector is a complete handle for the array, and is a convenient way to pass arrays as arguments to procedures. Many useful array slicing operations (such as selecting a sub-array, swapping indices, or reversing the direction of the indices) can be performed very efficiently by manipulating the dope vector.<sup>[2]</sup>

## Compact layouts

Often the coefficients are chosen so that the elements occupy a contiguous area of memory. However, that is not necessary. Even if arrays are always created with contiguous elements, some array slicing operations may create non-contiguous sub-arrays from them.

There are two systematic compact layouts for a two-dimensional array. For example, consider the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

In the row-major order layout (adopted by C for statically declared arrays), the elements in each row are stored in consecutive positions and all of the elements of a row have a lower address than any of the elements of a consecutive

row:

1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---

In column-major order (traditionally used by Fortran), the elements in each column are consecutive in memory and all of the elements of a columns have a lower address than any of the elements of a consecutive column:

1	4	7	2	5	8	3	6	9
---	---	---	---	---	---	---	---	---

For arrays with three or more indices, "row major order" puts in consecutive positions any two elements whose index tuples differ only by one in the *last* index. "Column major order" is analogous with respect to the *first* index.

In systems which use processor cache or virtual memory, scanning an array is much faster if successive elements are stored in consecutive positions in memory, rather than sparsely scattered. Many algorithms that use multidimensional arrays will scan them in a predictable order. A programmer (or a sophisticated compiler) may use this information to choose between row- or column-major layout for each array. For example, when computing the product  $A \cdot B$  of two matrices, it would be best to have  $A$  stored in row-major order, and  $B$  in column-major order.

## Array resizing

Static arrays have a size that is fixed when they are created and consequently do not allow elements to be inserted or removed. However, by allocating a new array and copying the contents of the old array to it, it is possible to effectively implement a *dynamic* version of an array; see dynamic array. If this operation is done infrequently, insertions at the end of the array require only amortized constant time.

Some array data structures do not reallocate storage, but do store a count of the number of elements of the array in use, called the count or size. This effectively makes the array a dynamic array with a fixed maximum size or capacity; Pascal strings are examples of this.

## Non-linear formulas

More complicated (non-linear) formulas are occasionally used. For a compact two-dimensional triangular array, for instance, the addressing formula is a polynomial of degree 2.

## Efficiency

Both *store* and *select* take (deterministic worst case) constant time. Arrays take linear ( $O(n)$ ) space in the number of elements  $n$  that they hold.

In an array with element size  $k$  and on a machine with a cache line size of  $B$  bytes, iterating through an array of  $n$  elements requires the minimum of  $\lceil nk/B \rceil$  cache misses, because its elements occupy contiguous memory locations. This is roughly a factor of  $B/k$  better than the number of cache misses needed to access  $n$  elements at random memory locations. As a consequence, sequential iteration over an array is noticeably faster in practice than iteration over many other data structures, a property called locality of reference (this does *not* mean however, that using a perfect hash or trivial hash within the same (local) array, will not be even faster - and achievable in constant time). Libraries provide low-level optimized facilities for copying ranges of memory (such as `memcpy`) which can be used to move contiguous blocks of array elements significantly faster than can be achieved through individual element access. The speedup of such optimized routines varies by array element size, architecture, and implementation.

Memory-wise, arrays are compact data structures with no per-element overhead. There may be a per-array overhead, e.g. to store index bounds, but this is language-dependent. It can also happen that elements stored in an array require *less* memory than the same elements stored in individual variables, because several array elements can be stored in a single word; such arrays are often called *packed* arrays. An extreme (but commonly used) case is the bit array, where

every bit represents a single element. A single octet can thus hold up to 256 different combinations of up to 8 different conditions, in the most compact form.

Array accesses with statically predictable access patterns are a major source of data parallelism.

## Efficiency comparison with other data structures

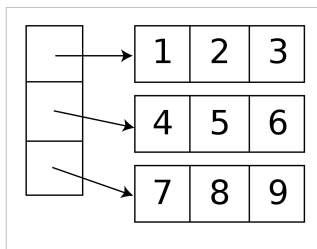
	Linked list	Array	Dynamic array	Balanced tree	Random access list
Indexing	$\Theta(n)$	$\Theta(1)$	$\Theta(1)$	$\Theta(\log n)$	$\Theta(\log n)$
Insert/delete at beginning	$\Theta(1)$	N/A	$\Theta(n)$	$\Theta(\log n)$	$\Theta(1)$
Insert/delete at end	$\Theta(1)$	N/A	$\Theta(1)$ amortized	$\Theta(\log n)$	$\Theta(\log n)$ updating
Insert/delete in middle	search time + $\Theta(1)$ <sup>[9][10][11]</sup>	N/A	$\Theta(n)$	$\Theta(\log n)$	$\Theta(\log n)$ updating
Wasted space (average)	$\Theta(n)$	0	$\Theta(n)$ <sup>[12]</sup>	$\Theta(n)$	$\Theta(n)$

Growable arrays are similar to arrays but add the ability to insert and delete elements; adding and deleting at the end is particularly efficient. However, they reserve linear ( $\Theta(n)$ ) additional storage, whereas arrays do not reserve additional storage.

Associative arrays provide a mechanism for array-like functionality without huge storage overheads when the index values are sparse. For example, an array that contains values only at indexes 1 and 2 billion may benefit from using such a structure. Specialized associative arrays with integer keys include Patricia tries, Judy arrays, and van Emde Boas trees.

Balanced trees require  $O(\log n)$  time for indexed access, but also permit inserting or deleting elements in  $O(\log n)$  time,<sup>[13]</sup> whereas growable arrays require linear ( $\Theta(n)$ ) time to insert or delete elements at an arbitrary position.

Linked lists allow constant time removal and insertion in the middle but take linear time for indexed access. Their memory use is typically worse than arrays, but is still linear.



An Iliffe vector is an alternative to a multidimensional array structure. It uses a one-dimensional array of references to arrays of one dimension less. For two dimensions, in particular, this alternative structure would be a vector of pointers to vectors, one for each row. Thus an element in row  $i$  and column  $j$  of an array  $A$  would be accessed by double indexing ( $A[i][j]$  in typical notation). This alternative structure allows *ragged* or *jagged* arrays, where each row may have a different size — or, in general, where the valid range of each index depends on the values of all

preceding indices. It also saves one multiplication (by the column address increment) replacing it by a bit shift (to index the vector of row pointers) and one extra memory access (fetching the row address), which may be worthwhile in some architectures.

## Meaning of dimension

The dimension of an array is the number of indices needed to select an element. Thus, if the array is seen as a function on a set of possible index combinations, it is the dimension of the space of which its domain is a discrete subset. Thus a one-dimensional array is a list of data, a two-dimensional array a rectangle of data, a three-dimensional array a block of data, etc.

This should not be confused with the dimension of the set of all matrices with a given domain, that is, the number of elements in the array. For example, an array with 5 rows and 4 columns is two-dimensional, but such matrices form a 20-dimensional space. Similarly, a three-dimensional vector can be represented by a one-dimensional array of size three.

## References

- [1] Black, Paul E. (13 November 2008). "array" (<http://www.nist.gov/dads/HTML/array.html>). *Dictionary of Algorithms and Data Structures*. National Institute of Standards and Technology. . Retrieved 2010-08-22.
- [2] Bjoern Andres; Ullrich Koethe; Thorben Kroeger; Hamprecht (2010). "Runtime-Flexible Multi-dimensional Arrays and Views for C++98 and C++0x". arXiv:1008.2909 [cs.DS].
- [3] Garcia, Ronald; Lumsdaine, Andrew (2005). "MultiArray: a C++ library for generic programming with arrays". *Software: Practice and Experience* **35** (2): 159–188. doi:10.1002/spe.630. ISSN 0038-0644.
- [4] David R. Richardson (2002), *The Book on Data Structures*. iUniverse, 112 pages. ISBN 0-595-24039-9, ISBN 978-0-595-24039-5.
- [5] T. Veldhuizen. Arrays in Blitz++. In Proc. of the 2nd Int. Conf. on Scientific Computing in Object-Oriented Parallel Environments (ISCOPE), LNCS 1505, pages 223-220. Springer, 1998.
- [6] Donald Knuth, *The Art of Computer Programming*, vol. 3. Addison-Wesley
- [7] "Array Code Examples - PHP Array Functions - PHP code" (<http://www.configure-all.com/arrays.php>). <http://www.configure-all.com/>: Computer Programming Web programming Tips. . Retrieved 2011-04-08. "In most computer languages array index (counting) starts from 0, not from 1. Index of the first element of the array is 0, index of the second element of the array is 1, and so on. In array of names below you can see indexes and values."
- [8] "Chapter 6 - Arrays, Types, and Constants" (<http://www.modula2.org/tutor/chapter6.php>). *Modula-2 Tutorial*. <http://www.modula2.org/tutor/index.php>. . Retrieved 2011-04-08. "The names of the twelve variables are given by Automobiles[1], Automobiles[2], ... Automobiles[12]. The variable name is "Automobiles" and the array subscripts are the numbers 1 through 12. [i.e. in Modula-2, the index starts by one!]"
- [9] Gerald Kruse. CS 240 Lecture Notes (<http://www.juniata.edu/faculty/kruse/cs240/syllabus.htm>): Linked Lists Plus: Complexity Trade-offs (<http://www.juniata.edu/faculty/kruse/cs240/linkedlist2.htm>). Juniata College. Spring 2008.
- [10] *Day 1 Keynote - Bjarne Stroustrup: C++11 Style* (<http://channel9.msdn.com/Events/GoingNative/GoingNative-2012/Keynote-Bjarne-Stroustrup-Cpp11-Style>) at *GoingNative 2012* on *channel9.msdn.com* from minute 45 or foil 44
- [11] *Number crunching: Why you should never, ever, EVER use linked-list in your code again* (<http://kjellkod.wordpress.com/2012/02/25/why-you-should-never-ever-ever-use-linked-list-in-your-code-again/>) at *kjellkod.wordpress.com*
- [12] Brodnik, Andrej; Carlsson, Svante; Sedgewick, Robert; Munro, JI; Demaine, ED (Technical Report CS-99-09), *Resizable Arrays in Optimal Time and Space* (<http://www.cs.uwaterloo.ca/research/tr/1999/09/CS-99-09.pdf>), Department of Computer Science, University of Waterloo,
- [13] Counted B-Tree (<http://www.chiark.greenend.org.uk/~sgtatham/algorithms/cbtree.html>)

# Row-major order

In computing, **row-major order** and **column-major order** describe methods for storing multidimensional arrays in linear memory. Following standard matrix notation, rows are numbered by the first index of a two-dimensional array and columns by the second index. Array layout is critical for correctly passing arrays between programs written in different languages. It is also important for performance when traversing an array because accessing array elements that are contiguous in memory is usually faster than accessing elements which are not, due to caching.

Row-major order is used in C, PL/I; column-major order is used in Fortran, MATLAB, GNU Octave, and Scilab.

## Row-major order

In row-major storage, a multidimensional array in linear memory is accessed such that rows are stored one after the other. It is the approach used by the C programming language and the statistical modelling language WinBUGS.<sup>[1]</sup>

When using row-major order, the difference between addresses of array cells in increasing rows is larger than differences between addresses of cells in increasing columns. For example, consider this 2×3 array:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

An array declared in C as

```
int A[2][3] = { {1, 2, 3}, {4, 5, 6} };
```

would be laid out contiguously in linear memory as:

```
1  2  3  4  5  6
```

To traverse this array in the order in which it is laid out in memory, one would use the following nested loop:

```
for (i = 0; i < 2; i++)
    for (j = 0; j < 3; j++)
        printf("%d\n", A[i][j]);
```

The difference in offset from one column to the next is 1 and from one row to the next is 3. The linear offset from the beginning of the array to any given element  $A[\text{row}][\text{column}]$  can then be computed as:

$$\text{offset} = \text{row} * \text{NUMCOLS} + \text{column}$$

where NUMCOLS is the number of columns in the array.

The above formula only works when using the C convention of labeling the first element 0. In other words, row 1, column 2 in matrix A, would be represented as  $A[0][1]$ .

Note that this technique generalizes, so a 2×3×4 array looks like:

```
int A[2][3][4] = { {{1,2,3,4}, {5,6,7,8}, {9,10,11,12}}, {{13,14,15,16},
{17,18,19,20}, {21,22,23,24}}};
```

and the array would be laid out in linear memory as:

```
1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
```

## Column-major order

**Column-major order** is a similar method of flattening arrays onto linear memory, but the columns are listed in sequence. The scientific programming language Fortran, the matrix-oriented languages MATLAB<sup>[2]</sup>, Octave and Scilab, the statistical languages S-Plus<sup>[1]</sup> and R<sup>[3]</sup> and the shading languages GLSL and HLSL use column-major ordering. The array

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

if stored contiguously in linear memory with column-major order would look like the following:

```
1  4  2  5  3  6
```

The memory offset could then be computed as:

$$\text{offset} = \text{row} + \text{column} * \text{NUMROWS}$$

where NUMROWS represents the number of rows in the array—in this case, 2.

Treating a row-major array as a column-major array is the same as transposing it. Because performing a transpose requires data movement, and is quite difficult to do in-place for non-square matrices, such transpositions are rarely performed explicitly. For example, software libraries for linear algebra, such as the BLAS, typically provide options to specify that certain matrices are to be interpreted in transposed order to avoid the necessity of data movement.

## Generalization to higher dimensions

It is possible to generalize both of these concepts to arrays with greater than two dimensions. For higher-dimensional arrays, the ordering determines which dimensions of the array are more consecutive in memory. Any of the dimensions could be consecutive, just as a two-dimensional array could be listed column-first or row-first. The difference in offset between listings of that dimension would then be determined by a product of other dimensions. It is uncommon, however, to have any variation except ordering dimensions first to last or last to first. These two variations correspond to row-major and column-major, respectively.

More explicitly, consider a  $d$ -dimensional  $N_1 \times N_2 \times \dots \times N_d$  array with dimensions  $N_k$  ( $k=1\dots d$ ). A given element of this array is specified by a tuple  $(n_1, n_2, \dots, n_d)$  of  $d$  (zero-based) indices  $n_k \in [0, N_k - 1]$ .

In **row-major order**, the *last* dimension is contiguous, so that the memory-offset of this element is given by:

$$n_d + N_d \cdot (n_{d-1} + N_{d-1} \cdot (n_{d-2} + N_{d-2} \cdot (\dots + N_2 n_1) \dots)) = \sum_{k=1}^d \left( \prod_{\ell=k+1}^d N_\ell \right) n_k$$

In **column-major order**, the *first* dimension is contiguous, so that the memory-offset of this element is given by:

$$n_1 + N_1 \cdot (n_2 + N_2 \cdot (n_3 + N_3 \cdot (\dots + N_{d-1} n_d) \dots)) = \sum_{k=1}^d \left( \prod_{\ell=1}^{k-1} N_\ell \right) n_k$$

Note that the difference between row-major and column-major order is simply that the order of the dimensions is reversed. Equivalently, in row-major order the rightmost indices vary faster as one steps through consecutive memory locations, while in column-major order the leftmost indices vary faster.

## References

- Donald E. Knuth, *The Art of Computer Programming Volume 1: Fundamental Algorithms*, third edition, section 2.2.6 (Addison-Wesley: New York, 1997).
- [1] Spiegelhalter et al. (2003, p. 17): Spiegelhalter, David; Thomas, Andrew; Best, Nicky; Lunn, Dave (January 2003), "Formatting of data: S-Plus format" (<http://www.mrc-bsu.cam.ac.uk/bugs>) (html), *WinBUGS User Manual* (Version 1.4 ed.), Robinson Way, Cambridge CB2 2SR, UK: MRC Biostatistics Unit, Institute of Public Health, PDF document (<http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/manual14.pdf>),
- [2] MATLAB documentation, `mxCalcSingleSubscript` function (<http://www.mathworks.com/access/helpdesk/help/techdoc/apiref/mxcalcingsubscript.html>) (retrieved from Mathworks.com, March 2010).
- [3] *An Introduction to R*, Section 5.1: Arrays (<http://cran.r-project.org/doc/manuals/R-intro.html#Arrays>) (retrieved March 2010).

## Dope vector

---

In computer programming, a **dope vector** is a data structure used to hold information about a data object,<sup>[1]</sup> e.g. an array, especially its memory layout.

A dope vector typically contains information about the type of array element, rank of an array, the extents of an array, and the stride of an array as well as a pointer to the block in memory containing the array elements.

It is often used in compilers to pass entire arrays between procedures in a high level language like Fortran.

Prior to the invention of the linked-list a dope vector was used internally in the internal structure of computer systems.

The dope vector included an identifier, a length, a parent address, and a next child address. The identifier was an assigned name and was mostly useless, but the length was the amount of allocated storage to this vector from the end of the dope vector that contained data of use to the internal processes of the computer. This length by many was called the offset, span of vector length. The parent and child references were absolute core references, or register and offset settings to the parent or child depending on the type of computer.

Dope vectors were managed internally by the operating system and allowed the processor to allocate and de-allocate storage in specific segments as needed.

Later dope vectors had a status bit that told the system if they were active; if it was not active it would be reallocated when needed. Using this technology the computer could perform a more granular memory management.

- [1] Pratt T. and M. Zelkowitz, *Programming Languages: Design and Implementation* (Third Edition), Prentice Hall, Upper Saddle River, NJ, (1996) pp 114
-



---

# Iliffe vector

---

In computer programming, an **Iliffe vector**, also known as a **display**, is a data structure used to implement multi-dimensional arrays. An Iliffe vector for an  $n$ -dimensional array (where  $n > 2$ ) consists of a vector (or 1-dimensional array) of pointers to an  $(n - 1)$ -dimensional array. They are often used to avoid the need for expensive multiplication operations when performing address calculation on an array element. They can also be used to implement triangular arrays, or other kinds of irregularly shaped arrays. The data structure is named after John K. Iliffe,

Their disadvantages include the need for multiple chained pointer indirections to access an element, and the extra work required to determine the next row in an  $n$ -dimensional array to allow an optimising compiler to prefetch it. Both of these are a source of delays on systems where the CPU is significantly faster than main memory.

The Iliffe vector for a 2-dimensional array is simply a vector of pointers to vectors of data, i.e., the Iliffe vector represents the columns of an array where each column element is a pointer to a row vector.

Multidimensional arrays in languages such as Java, Python (multidimensional lists), Ruby, Perl, PHP, JavaScript, Objective-C, and Atlas Autocode are implemented as Iliffe vectors.

Iliffe vectors are contrasted with dope vectors in languages such as Fortran, which contain the stride factors and offset values for the subscripts in each dimension.

## Notes

## References

- John K. Iliffe (1961). "The Use of The Genie System in Numerical Calculations". *Annual Review in Automatic Programming* **2**: 25. doi:10.1016/S0066-4138(61)80002-5.

# Dynamic array

In computer science, a **dynamic array**, **growable array**, **resizable array**, **dynamic table**, **mutable array**, or **array list** is a random access, variable-size list data structure that allows elements to be added or removed. It is supplied with standard libraries in many modern mainstream programming languages.

A dynamic array is not the same thing as a dynamically allocated array, which is a fixed-size array whose size is fixed when the array is allocated, although a dynamic array may use such a fixed-size array as a back end.<sup>[1]</sup>

## Bounded-size dynamic arrays and capacity

The simplest dynamic array is constructed by allocating a fixed-size array and then dividing it into two parts: the first stores the elements of the dynamic array and the second is reserved, or unused. We can then add or remove elements at the end of the dynamic array in constant time by using the reserved space, until this space is completely consumed. The number of elements used by the dynamic array contents is its *logical size* or *size*, while the size of the underlying array is called the dynamic array's *capacity*, which is the maximum possible size without relocating data.

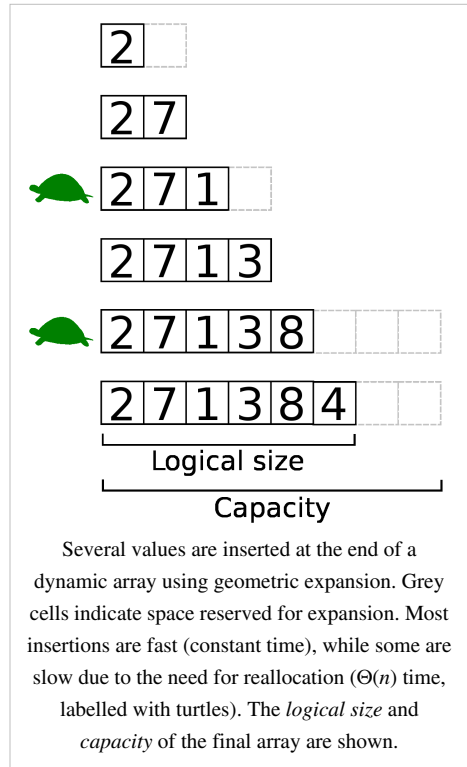
In applications where the logical size is bounded, the fixed-size data structure suffices. This may be short-sighted, when problems with the array filling up turn up later. It is best to put resize code into any array, to respond to new conditions. Then choosing initial capacity is optimization, not getting the program to run. Resizing the underlying array is an expensive task, typically involving copying the entire contents of the array.

## Geometric expansion and amortized cost

To avoid incurring the cost of resizing many times, dynamic arrays resize by a large amount, such as doubling in size, and use the reserved space for future expansion. The operation of adding an element to the end might work as follows:

```
function insertEnd(dynarray a, element e)
    if (a.size = a.capacity)
        // resize a to twice its current capacity:
        a.capacity ← a.capacity * 2
        // (copy the contents to the new memory location here)
    a[a.size] ← e
    a.size ← a.size + 1
```

As  $n$  elements are inserted, the capacities form a geometric progression. Expanding the array by any constant proportion ensures that inserting  $n$  elements takes  $O(n)$  time overall, meaning that each insertion takes amortized constant time. The value of this proportion  $a$  leads to a time-space tradeoff: the average time per insertion operation is about  $a/(a-1)$ , while the number of wasted cells is bounded above by  $(a-1)n$ . The choice of  $a$  depends on the library or application: some textbooks use  $a = 2$ ,<sup>[2][3]</sup> but Java's ArrayList implementation uses  $a = 3/2$ <sup>[1]</sup> and the C



implementation of Python's list data structure uses  $a = 9/8$ .<sup>[4]</sup>

Many dynamic arrays also deallocate some of the underlying storage if its size drops below a certain threshold, such as 30% of the capacity. This threshold must be strictly smaller than  $1/a$  in order to support mixed sequences of insertions and removals with amortized constant cost.

Dynamic arrays are a common example when teaching amortized analysis.<sup>[2][3]</sup>

## Performance

	Linked list	Array	Dynamic array	Balanced tree	Random access list
Indexing	$\Theta(n)$	$\Theta(1)$	$\Theta(1)$	$\Theta(\log n)$	$\Theta(\log n)$
Insert/delete at beginning	$\Theta(1)$	N/A	$\Theta(n)$	$\Theta(\log n)$	$\Theta(1)$
Insert/delete at end	$\Theta(1)$	N/A	$\Theta(1)$ amortized	$\Theta(\log n)$	$\Theta(\log n)$ updating
Insert/delete in middle	search time $\Theta(1)$ <sup>[5][6][7]</sup>	N/A	$\Theta(n)$	$\Theta(\log n)$	$\Theta(\log n)$ updating
Wasted space (average)	$\Theta(n)$	0	$\Theta(n)$ <sup>[8]</sup>	$\Theta(n)$	$\Theta(n)$

The dynamic array has performance similar to an array, with the addition of new operations to add and remove elements from the end:

- Getting or setting the value at a particular index (constant time)
- Iterating over the elements in order (linear time, good cache performance)
- Inserting or deleting an element in the middle of the array (linear time)
- Inserting or deleting an element at the end of the array (constant amortized time)

Dynamic arrays benefit from many of the advantages of arrays, including good locality of reference and data cache utilization, compactness (low memory use), and random access. They usually have only a small fixed additional overhead for storing information about the size and capacity. This makes dynamic arrays an attractive tool for building cache-friendly data structures.

Compared to linked lists, dynamic arrays have faster indexing (constant time versus linear time) and typically faster iteration due to improved locality of reference; however, dynamic arrays require linear time to insert or delete at an arbitrary location, since all following elements must be moved, while linked lists can do this in constant time. This disadvantage is mitigated by the gap buffer and *tiered vector* variants discussed under *Variants* below. Also, in a highly fragmented memory region, it may be expensive or impossible to find contiguous space for a large dynamic array, whereas linked lists do not require the whole data structure to be stored contiguously.

A balanced tree can store a list while providing all operations of both dynamic arrays and linked lists reasonably efficiently, but both insertion at the end and iteration over the list are slower than for a dynamic array, in theory and in practice, due to non-contiguous storage and tree traversal/manipulation overhead.

## Variants

Gap buffers are similar to dynamic arrays but allow efficient insertion and deletion operations clustered near the same arbitrary location. Some deque implementations use array deques, which allow amortized constant time insertion/removal at both ends, instead of just one end.

Goodrich<sup>[9]</sup> presented a dynamic array algorithm called *Tiered Vectors* that provided  $O(n^{1/2})$  performance for order preserving insertions or deletions from the middle of the array.

Hashed Array Tree (HAT) is a dynamic array algorithm published by Sitarski in 1996.<sup>[10]</sup> Hashed Array Tree wastes order  $n^{1/2}$  amount of storage space, where  $n$  is the number of elements in the array. The algorithm has  $O(1)$  amortized performance when appending a series of objects to the end of a Hashed Array Tree.

In a 1999 paper,<sup>[8]</sup> Brodnik et al. describe a tiered dynamic array data structure, which wastes only  $n^{1/2}$  space for  $n$  elements at any point in time, and they prove a lower bound showing that any dynamic array must waste this much space if the operations are to remain amortized constant time. Additionally, they present a variant where growing and shrinking the buffer has not only amortized but worst-case constant time.

Bagwell (2002)<sup>[11]</sup> presented the VList algorithm, which can be adapted to implement a dynamic array.

## Language support

C++'s `std::vector` is an implementation of dynamic arrays, as are the `ArrayList`<sup>[12]</sup> classes supplied with the Java API and the .NET Framework. The generic `List<>` class supplied with version 2.0 of the .NET Framework is also implemented with dynamic arrays. Smalltalk's `OrderedCollection` is a dynamic array with dynamic start and end-index, making the removal of the first element also  $O(1)$ . Python's `list` datatype implementation is a dynamic array. Delphi and D implement dynamic arrays at the language's core. Ada's `Ada.Containers.Vectors` generic package provides dynamic array implementation for a given subtype. Many scripting languages such as Perl and Ruby offer dynamic arrays as a built-in primitive data type. Several cross-platform frameworks provide dynamic array implementations for C: `CArray` and `CFMutableArray` in Core Foundation; `GArray` and `GPtrArray` in GLib.

## References

- [1] See, for example, the source code of `java.util.ArrayList` class from OpenJDK 6 (<http://hg.openjdk.java.net/jdk6/jdk6/jdk/file/e0e25ac28560/src/share/classes/java/util/ArrayList.java>).
- [2] Goodrich, Michael T.; Tamassia, Roberto (2002), "1.5.2 Analyzing an Extendable Array Implementation", *Algorithm Design: Foundations, Analysis and Internet Examples*, Wiley, pp. 39–41.
- [3] Cormen, Thomas H.; Leiserson, Charles E.; Rivest, Ronald L.; Stein, Clifford (2001) [1990]. "17.4 Dynamic tables". *Introduction to Algorithms* (2nd ed.). MIT Press and McGraw-Hill. pp. 416–424. ISBN 0-262-03293-7.
- [4] List object implementation (<http://svn.python.org/projects/python/trunk/Objects/listobject.c>) from python.org, retrieved 2011-09-27.
- [5] Gerald Kruse. CS 240 Lecture Notes (<http://www.juniata.edu/faculty/kruse/cs240/syllabus.htm>): Linked Lists Plus: Complexity Trade-offs (<http://www.juniata.edu/faculty/kruse/cs240/linkedlist2.htm>). Juniata College. Spring 2008.
- [6] *Day 1 Keynote - Bjarne Stroustrup: C++11 Style* (<http://channel9.msdn.com/Events/GoingNative/GoingNative-2012/Keynote-Bjarne-Stroustrup-Cpp11-Style>) at *GoingNative 2012* on [channel9.msdn.com](http://channel9.msdn.com) from minute 45 or foil 44
- [7] *Number crunching: Why you should never, ever, EVER use linked-list in your code again* (<http://kjellkod.wordpress.com/2012/02/25/why-you-should-never-ever-ever-use-linked-list-in-your-code-again/>) at [kjellkod.wordpress.com](http://kjellkod.wordpress.com)
- [8] Brodnik, Andrej; Carlsson, Svante; Sedgewick, Robert; Munro, JI; Demaine, ED (Technical Report CS-99-09), *Resizable Arrays in Optimal Time and Space* (<http://www.cs.uwaterloo.ca/research/tr/1999/09/CS-99-09.pdf>), Department of Computer Science, University of Waterloo,
- [9] Goodrich, Michael T.; Kloss II, John G. (1999), "Tiered Vectors: Efficient Dynamic Arrays for Rank-Based Sequences" (<http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.17.7503>), *Workshop on Algorithms and Data Structures* **1663**: 205–216, doi:10.1007/3-540-48447-7\_21,
- [10] Sitarski, Edward (September 1996), *Algorithm Alley* (<http://www.ddj.com/architect/184409965?pgno=5>), "HATs: Hashed array trees", *Dr. Dobbs's Journal* **21** (11),
- [11] Bagwell, Phil (2002), *Fast Functional Lists, Hash-Lists, Deques and Variable Length Arrays* (<http://citeseer.ist.psu.edu/bagwell02fast.html>), EPFL,

[12] Javadoc on `ArrayList`

## External links

- NIST Dictionary of Algorithms and Data Structures: Dynamic array (<http://www.nist.gov/dads/HTML/dynamicarray.html>)
- VPOOL (<http://www.bsdua.org/libbsdua.html#vpool>) - C language implementation of dynamic array.
- CollectionSpy (<http://www.collectionspy.com>) — A Java profiler with explicit support for debugging `ArrayList`- and `Vector`-related issues.
- Open Data Structures - Chapter 2 - Array-Based Lists ([http://opendatastructures.org/versions/edition-0.1e/ods-java/2\\_Array\\_Based\\_Lists.html](http://opendatastructures.org/versions/edition-0.1e/ods-java/2_Array_Based_Lists.html))

## Hashed array tree

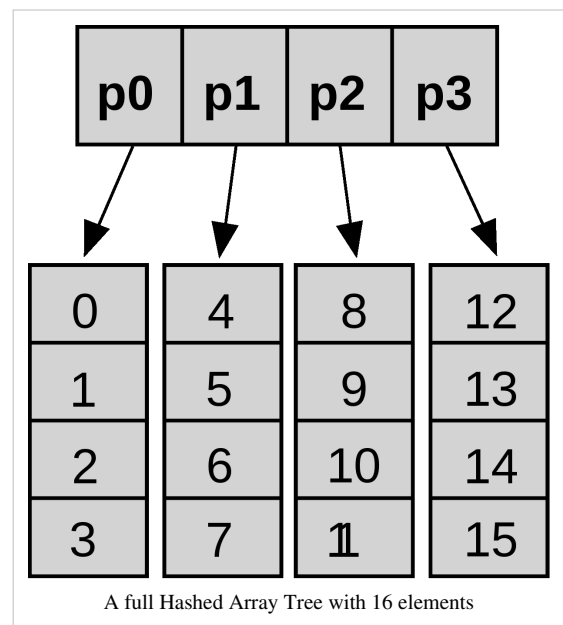
In computer science, a **hashed array tree** (HAT) is a dynamic array data-structure published by Edward Sitarski in 1996,<sup>[1]</sup> maintaining an array of separate memory fragments (or "leaves") to store the data elements, unlike simple dynamic arrays which maintain their data in one contiguous memory area. Its primary objective is to reduce the amount of element copying due to automatic array resizing operations, and to improve memory usage patterns.

Whereas simple dynamic arrays based on geometric expansion waste linear ( $\Omega(n)$ ) space, where  $n$  is the number of elements in the array, hashed array trees waste only order  $O(n^{1/2})$  storage space. An optimization of the algorithm allows to eliminate data copying completely, at a cost of increasing the wasted space.

It can perform access in constant ( $O(1)$ ) time, though slightly slower than simple dynamic arrays. The algorithm has  $O(1)$  amortized performance when appending a series of objects to the end of a hashed array tree. Contrary to its name, it does not use hash functions.

### Definitions

As defined by Sitarski, a hashed array tree has a top-level directory containing a power of two number of leaf arrays. All leaf arrays are the same size as the top-level directory. This structure superficially resembles a hash table with array-based collision chains, which is the basis for the name *hashed array tree*. A full hashed array tree can hold  $m^2$  elements, where  $m$  is the size of the top-level directory.<sup>[1]</sup> The use of powers of two enables faster physical addressing through bit operations instead of arithmetic operations of quotient and remainder<sup>[1]</sup> and ensures the  $O(1)$  amortized performance of append operation in the presence of occasional global array copy while expanding.



### Expansions and size reductions

In a usual dynamic array geometric expansion scheme, the array is reallocated as a whole sequential chunk of memory with the new size a double of its current size (and the whole data is then moved to the new location). This ensures  $O(1)$  amortized operations at a cost of  $O(n)$  wasted space, as the enlarged array is filled to the half of its new capacity.

When a hashed array tree is full, its directory and leaves must be restructured to twice their prior size to accommodate additional append operations. The data held in old structure is then moved into the new locations. Only one new leaf is then allocated and added into the top array which thus becomes filled only to a quarter of its new capacity. All the extra leaves are not allocated yet, and will only be allocated when needed, thus wasting only  $O(n^{1/2})$  of storage.

There are multiple alternatives for reducing size: when a Hashed Array Tree is one eighth full, it can be restructured to a smaller, half-full hashed array tree; another option is only freeing unused leaf arrays, without resizing the leaves. Further optimizations include adding new leaves without resizing, growing the directory array as needed, possibly through geometric expansion. This would eliminate the need for data copying completely, at the cost of making the wasted space  $O(n)$ , with a small coefficient, and only performing restructuring when a set threshold overhead is reached.<sup>[1]</sup>

## Related data structures

Brodnik et al.<sup>[2]</sup> presented a dynamic array algorithm with a similar space wastage profile to hashed array trees. Brodnik's implementation retains previously allocated leaf arrays, with a more complicated address calculation function as compared to hashed array trees.

## References

- [1] Sitarski, Edward (September 1996), *Algorithm Alley* (<http://www.ddj.com/architect/184409965?pgno=5>), "HATs: Hashed array trees", *Dr. Dobbs's Journal* **21** (11),
- [2] Brodnik, Andrej; Carlsson, Svante; Sedgewick, Robert; Munro, JI; Demaine, ED (Technical Report CS-99-09), *Resizable Arrays in Optimal Time and Space* (<http://www.cs.uwaterloo.ca/research/tr/1999/09/CS-99-09.pdf>), Department of Computer Science, University of Waterloo,

# Gap buffer

---

A **gap buffer** in computer science is a dynamic array that allows efficient insertion and deletion operations clustered near the same location. Gap buffers are especially common in text editors, where most changes to the text occur at or near the current location of the cursor. The text is stored in a large buffer in two contiguous segments, with a gap between them for inserting new text. Moving the cursor involves copying text from one side of the gap to the other (sometimes copying is delayed until the next operation that changes the text). Insertion adds new text at the end of the first segment. Deletion increases the size of the gap.

Text in a gap buffer is represented as two literal strings, which take very little extra space and which can be searched and displayed very quickly, compared to more sophisticated data structures such as linked lists. However, operations at different locations in the text and ones that fill the gap (requiring a new gap to be created) may require copying most of the text, which is especially inefficient for large files. The use of gap buffers is based on the assumption that such recopying occurs rarely enough that its cost can be amortized over the more common cheap operations.

## Example

Below are some examples of operations with buffer gaps. The gap is represented pictorially by the empty space between the square brackets. This representation is a bit misleading: in a typical implementation, the endpoints of the gap are tracked using pointers or array indices, and the contents of the gap are ignored; this allows, for example, deletions to be done by adjusting a pointer without changing the text in the buffer. It is a common programming practice to use a semi-open interval for the gap pointers, i.e. the start-of-gap points to the invalid character following the last character in the first buffer, and the end-of-gap points to the first valid character in the second buffer (or

---

equivalently, the pointers are considered to point "between" characters).

Initial state:

```
This is the way [          ]out.
```

User inserts some new text:

```
This is the way the world started [  ]out.
```

User moves the cursor before "started"; system moves "started " from the first buffer to the second buffer.

```
This is the way the world [  ]started out.
```

User adds text filling the gap; system creates new gap:

```
This is the way the world as we know it [          ]started out.
```

## External links

- Overview and implementation in .NET/C# <sup>[1]</sup>
- Brief overview and sample C++ code <sup>[2]</sup>
- Implementation of a cyclic sorted gap buffer in .NET/C# <sup>[3]</sup>
- Use of gap buffer in early editor. <sup>[4]</sup> (First written somewhere between 1969 and 1971)
- emacs gap buffer info <sup>[5]</sup> (Emacs gap buffer reference)
- Text Editing <sup>[6]</sup>

## References

- [1] <http://www.codeproject.com/KB/recipes/GenericGapBuffer.aspx>
- [2] <http://www.lazyhacker.com/gapbuffer/gapbuffer.htm>
- [3] <http://www.codeproject.com/KB/recipes/SplitArrayDictionary.aspx>
- [4] <http://history.dcs.ed.ac.uk/archive/apps/ecce/hmd/e915.imp.html>
- [5] [http://www.gnu.org/software/emacs/elisp/html\\_node/Buffer-Gap.html](http://www.gnu.org/software/emacs/elisp/html_node/Buffer-Gap.html)
- [6] <http://www.common-lisp.net/project/flexichain/download/StrandhVilleneuveMoore.pdf>

# Circular buffer

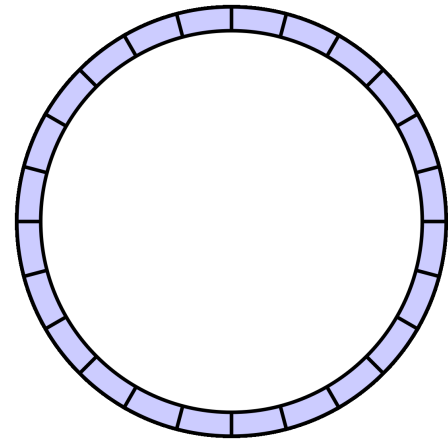
A **circular buffer**, **cyclic buffer** or **ring buffer** is a data structure that uses a single, fixed-size buffer as if it were connected end-to-end. This structure lends itself easily to buffering data streams.

## Uses

An example that could possibly use an overwriting circular buffer is with multimedia. If the buffer is used as the bounded buffer in the producer-consumer problem then it is probably desired for the producer (e.g., an audio generator) to overwrite old data if the consumer (e.g., the sound card) is unable to momentarily keep up. Another example is the digital waveguide synthesis method which uses circular buffers to efficiently simulate the sound of vibrating strings or wind instruments.

The "prized" attribute of a circular buffer is that it does not need to have its elements shuffled around when one is consumed. (If a non-circular buffer were used then it would be necessary to shift all elements when one is consumed.) In other words, the circular buffer is well suited as a FIFO buffer while a standard, non-circular buffer is well suited as a LIFO buffer.

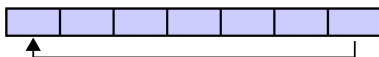
Circular buffering makes a good implementation strategy for a queue that has fixed maximum size. Should a maximum size be adopted for a queue, then a circular buffer is a completely ideal implementation; all queue operations are constant time. However, expanding a circular buffer requires shifting memory, which is comparatively costly. For arbitrarily expanding queues, a Linked list approach may be preferred instead.



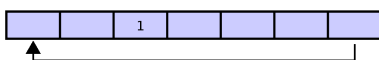
A ring showing, conceptually, a circular buffer. This visually shows that the buffer has no real end and it can loop around the buffer. However, since memory is never physically created as a ring, a linear representation is generally used as is done below.

## How it works

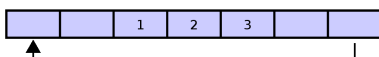
A circular buffer first starts empty and of some predefined length. For example, this is a 7-element buffer:



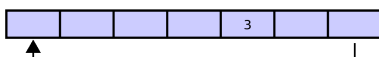
Assume that a 1 is written into the middle of the buffer (exact starting location does not matter in a circular buffer):



Then assume that two more elements are added — 2 & 3 — which get appended after the 1:

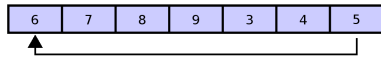


If two elements are then removed from the buffer, the oldest values inside the buffer are removed. The two elements removed, in this case, are 1 & 2 leaving the buffer with just a 3:

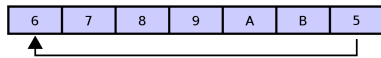


If the buffer has 7 elements then it is completely full:



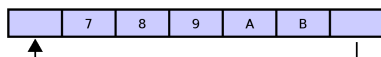


A consequence of the circular buffer is that when it is full and a subsequent write is performed, then it starts overwriting the oldest data. In this case, two more elements — A & B — are added and they *overwrite* the 3 & 4:



Alternatively, the routines that manage the buffer could prevent overwriting the data and return an error or raise an exception. Whether or not data is overwritten is up to the semantics of the buffer routines or the application using the circular buffer.

Finally, if two elements are now removed then what would be returned is **not** 3 & 4 but 5 & 6 because A & B overwrote the 3 & the 4 yielding the buffer with:



## Circular buffer mechanics

What is not shown in the example above is the mechanics of how the circular buffer is managed.

### Start / End Pointers

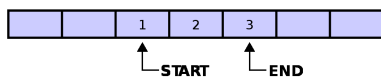
Generally, a circular buffer requires four pointers:

- one to the actual buffer in memory
- one to the buffer end in memory (or alternately: the size of the buffer)
- one to point to the start of valid data (or alternately: amount of data written to the buffer)
- one to point to the end of valid data (or alternately: amount of data read from the buffer)

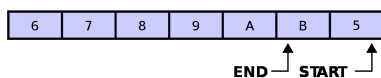
Alternatively, a fixed-length buffer with two integers to keep track of indices can be used in languages that do not have pointers.

Taking a couple of examples from above. (While there are numerous ways to label the pointers and exact semantics can vary, this is one way to do it.)

This image shows a partially full buffer:



This image shows a full buffer with two elements having been overwritten:

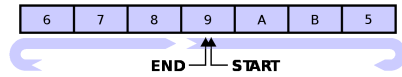


What to note about the second one is that after each element is overwritten then the start pointer is incremented as well.

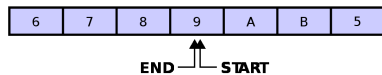
## Difficulties

### Full / Empty Buffer Distinction

A small disadvantage of relying on pointers or relative indices of the start and end of data is, that in the case the buffer is entirely full, both pointers point to the same element:



This is exactly the same situation as when the buffer is empty:



To solve this confusion there are a number of solutions:

- Always keep one slot open.
- Use a fill count to distinguish the two cases.
- Use read and write counts to get the fill count from.
- Use absolute indices.

### Always Keep One Slot Open

This design always keeps one slot unallocated. A full buffer has at most  $(\text{size} - 1)$  slots. If both pointers refer to the same slot, the buffer is empty. If the end (write) pointer refers to the slot preceding the one referred to by the start (read) pointer, the buffer is full. This is a simple, robust, approach that only requires two pointers, at the expense of one buffer slot.

### Example implementation, 'C' language

```
/* Circular buffer example, keeps one slot open */

#include <stdio.h>
#include <malloc.h>

/* Opaque buffer element type. This would be defined by the
application. */
typedef struct { int value; } ElemType;

/* Circular buffer object */
typedef struct {
    int      size;      /* maximum number of elements */
    int      start;     /* index of oldest element */
    int      end;       /* index at which to write new element */
    ElemType *elems;    /* vector of elements */
} CircularBuffer;

void cbInit(CircularBuffer *cb, int size) {
    cb->size = size + 1; /* include empty elem */
    cb->start = 0;
    cb->end = 0;
    cb->elems = (ElemType *)calloc(cb->size, sizeof(ElemType));
}
```

```
void cbFree(CircularBuffer *cb) {
    free(cb->elems); /* OK if null */ }

int cbIsFull(CircularBuffer *cb) {
    return (cb->end + 1) % cb->size == cb->start; }

int cbIsEmpty(CircularBuffer *cb) {
    return cb->end == cb->start; }

/* Write an element, overwriting oldest element if buffer is full. App
can
choose to avoid the overwrite by checking cbIsFull(). */
void cbWrite(CircularBuffer *cb, ElemType *elem) {
    cb->elems[cb->end] = *elem;
    cb->end = (cb->end + 1) % cb->size;
    if (cb->end == cb->start)
        cb->start = (cb->start + 1) % cb->size; /* full, overwrite */
}

/* Read oldest element. App must ensure !cbIsEmpty() first. */
void cbRead(CircularBuffer *cb, ElemType *elem) {
    *elem = cb->elems[cb->start];
    cb->start = (cb->start + 1) % cb->size;
}

int main(int argc, char **argv) {
    CircularBuffer cb;
    ElemType elem = {0};

    int testBufferSize = 10; /* arbitrary size */
    cbInit(&cb, testBufferSize);

    /* Fill buffer with test elements 3 times */
    for (elem.value = 0; elem.value < 3 * testBufferSize; ++ elem.value)
        cbWrite(&cb, &elem);

    /* Remove and print all elements */
    while (!cbIsEmpty(&cb)) {
        cbRead(&cb, &elem);
        printf("%d\n", elem.value);
    }

    cbFree(&cb);
    return 0;
}
```

### Use a Fill Count

This approach replaces the end pointer with a counter that tracks the number of readable items in the buffer. This unambiguously indicates when the buffer is empty or full and allows use of all buffer slots.

The performance impact should be negligible, since this approach adds the costs of maintaining the counter and computing the tail slot on writes but eliminates the need to maintain the end pointer and simplifies the fullness test.

- Note: When using semaphores in a Producer-consumer model, the semaphores act as a fill count.

### Differences from previous example

```
/* This approach replaces the CircularBuffer 'end' field with the  
   'count' field and changes these functions: */  
  
void cbInit(CircularBuffer *cb, int size) {  
    cb->size = size;  
    cb->start = 0;  
    cb->count = 0;  
    cb->elems = (ElemType *)calloc(cb->size, sizeof(ElemType));  
}  
  
int cbIsFull(CircularBuffer *cb) {  
    return cb->count == cb->size; }  
  
int cbIsEmpty(CircularBuffer *cb) {  
    return cb->count == 0; }  
  
void cbWrite(CircularBuffer *cb, ElemType *elem) {  
    int end = (cb->start + cb->count) % cb->size;  
    cb->elems[end] = *elem;  
    if (cb->count == cb->size)  
        cb->start = (cb->start + 1) % cb->size; /* full, overwrite */  
    else  
        ++ cb->count;  
}  
  
void cbRead(CircularBuffer *cb, ElemType *elem) {  
    *elem = cb->elems[cb->start];  
    cb->start = (cb->start + 1) % cb->size;  
    -- cb->count;  
}
```

### Read / Write Counts

Another solution is to keep counts of the number of items written to and read from the circular buffer. Both counts are stored in signed integer variables with numerical limits larger than the number of items that can be stored and are allowed to wrap freely.

The unsigned difference ( $\text{write\_count} - \text{read\_count}$ ) always yields the number of items placed in the buffer and not yet retrieved. This can indicate that the buffer is empty, partially full, completely full (without waste of a storage location) or in a state of overrun.

The advantage is:

- The source and sink of data can implement independent policies for dealing with a full buffer and overrun while adhering to the rule that only the source of data modifies the write count and only the sink of data modifies the read count. This can result in elegant and robust circular buffer implementations even in multi-threaded environments.

The disadvantage is:

- You need two additional variables.

### Record last operation

Another solution is to keep a flag indicating whether the most recent operation was a read or a write. If the two pointers are equal, then the flag will show whether the buffer is full or empty: if the most recent operation was a write, the buffer must be full, and conversely if it was a read, it must be empty.

The advantages are:

- Only a single bit needs to be stored (which may be particularly useful if the algorithm is implemented in hardware)
- The test for full/empty is simple

The disadvantage is:

- You need an extra variable

### Absolute indices

If indices are used instead of pointers, indices can store read/write counts instead of the offset from start of the buffer. This is similar to the above solution, except that there are no separate variables, and relative indices are obtained on the fly by division modulo the buffer's length.

The advantage is:

- No extra variables are needed.

The disadvantages are:

- Every access needs an additional *modulo* operation.
- If counter wrap is possible, complex logic can be needed if the buffer's length is not a divisor of the counter's capacity.

On binary computers, both of these disadvantages disappear if the buffer's length is a power of two—at the cost of a constraint on possible buffers lengths.

## Multiple Read Pointers

A little bit more complex are multiple read pointers on the same circular buffer. This is useful if you have  $n$  threads, which are reading from the same buffer, but *one* thread writing to the buffer.

## Chunked Buffer

Much more complex are different chunks of data in the same circular buffer. The writer is not only writing elements to the buffer, it also assigns these elements to chunks .

The reader should not only be able to read from the buffer, it should also get informed about the chunk borders.

Example: The writer is reading data from small files, writing them into the same circular buffer. The reader is reading the data, but needs to know when and which file is starting at a given position.

## Optimization

A circular-buffer implementation may be optimized by mapping the underlying buffer to two contiguous regions of virtual memory. (Naturally, the underlying buffer's length must then equal some multiple of the system's page size.) Reading from and writing to the circular buffer may then be carried out with greater efficiency by means of direct memory access; those accesses which fall beyond the end of the first virtual-memory region will automatically wrap around to the beginning of the underlying buffer. When the read offset is advanced into the second virtual-memory region, both offsets—read and write—are decremented by the length of the underlying buffer.

## Optimized POSIX Implementation

```
#include <sys/mman.h>
#include <stdlib.h>
#include <unistd.h>

#define report_exceptional_condition() abort ()

struct ring_buffer
{
    void *address;

    unsigned long count_bytes;
    unsigned long write_offset_bytes;
    unsigned long read_offset_bytes;
};

//Warning order should be at least 12 for Linux
void
ring_buffer_create (struct ring_buffer *buffer, unsigned long order)
{
    char path[] = "/dev/shm/ring-buffer-XXXXXX";
    int file_descriptor;
    void *address;
    int status;

    file_descriptor = mkstemp (path);
```

```
if (file_descriptor < 0)
    report_exceptional_condition ();

status = unlink (path);
if (status)
    report_exceptional_condition ();

buffer->count_bytes = 1UL << order;
buffer->write_offset_bytes = 0;
buffer->read_offset_bytes = 0;

status = ftruncate (file_descriptor, buffer->count_bytes);
if (status)
    report_exceptional_condition ();

buffer->address = mmap (NULL, buffer->count_bytes << 1, PROT_NONE,
                        MAP_ANONYMOUS | MAP_PRIVATE, -1, 0);

if (buffer->address == MAP_FAILED)
    report_exceptional_condition ();

address =
    mmap (buffer->address, buffer->count_bytes, PROT_READ | PROT_WRITE,
          MAP_FIXED | MAP_SHARED, file_descriptor, 0);

if (address != buffer->address)
    report_exceptional_condition ();

address = mmap (buffer->address + buffer->count_bytes,
                buffer->count_bytes, PROT_READ | PROT_WRITE,
                MAP_FIXED | MAP_SHARED, file_descriptor, 0);

if (address != buffer->address + buffer->count_bytes)
    report_exceptional_condition ();

status = close (file_descriptor);
if (status)
    report_exceptional_condition ();
}

void
ring_buffer_free (struct ring_buffer *buffer)
{
    int status;

    status = munmap (buffer->address, buffer->count_bytes << 1);
    if (status)
```

```
    report_exceptional_condition ();
}

void *
ring_buffer_write_address (struct ring_buffer *buffer)
{
    /** void pointer arithmetic is a constraint violation. **/
    return buffer->address + buffer->write_offset_bytes;
}

void
ring_buffer_write_advance (struct ring_buffer *buffer,
                          unsigned long count_bytes)
{
    buffer->write_offset_bytes += count_bytes;
}

void *
ring_buffer_read_address (struct ring_buffer *buffer)
{
    return buffer->address + buffer->read_offset_bytes;
}

void
ring_buffer_read_advance (struct ring_buffer *buffer,
                         unsigned long count_bytes)
{
    buffer->read_offset_bytes += count_bytes;

    if (buffer->read_offset_bytes >= buffer->count_bytes)
    {
        buffer->read_offset_bytes -= buffer->count_bytes;
        buffer->write_offset_bytes -= buffer->count_bytes;
    }
}

unsigned long
ring_buffer_count_bytes (struct ring_buffer *buffer)
{
    return buffer->write_offset_bytes - buffer->read_offset_bytes;
}

unsigned long
ring_buffer_count_free_bytes (struct ring_buffer *buffer)
{
    return buffer->count_bytes - ring_buffer_count_bytes (buffer);
}
```



```
void
ring_buffer_clear (struct ring_buffer *buffer)
{
    buffer->write_offset_bytes = 0;
    buffer->read_offset_bytes = 0;
}

/*Note, that initial anonymous mmap() can be avoided - after initial
mmap() for descriptor fd,
you can try mmap() with hinted address as (buffer->address +
buffer->count_bytes) and if it fails -
another one with hinted address as (buffer->address -
buffer->count_bytes).
Make sure MAP_FIXED is not used in such case, as under certain
situations it could end with segfault.
The advantage of such approach is, that it avoids requirement to map
twice the amount you need initially
(especially useful e.g. if you want to use hugetlbfs and the allowed
amount is limited)
and in context of gcc/glibc - you can avoid certain feature macros
(MAP_ANONYMOUS usually requires one of: _BSD_SOURCE, _SVID_SOURCE or
_GNU_SOURCE).*/
```

## External links

- CircularBuffer at the Portland Pattern Repository
- Boost: Templated Circular Buffer Container<sup>[1]</sup>
- <http://www.dspguide.com/ch28/2.htm>

## References

- [1] [http://www.boost.org/doc/libs/1\\_39\\_0/libs/circular\\_buffer/doc/circular\\_buffer.html](http://www.boost.org/doc/libs/1_39_0/libs/circular_buffer/doc/circular_buffer.html)

# Sparse array

---

In computer science, a **sparse array** is an array in which most of the elements have the same value (known as the default value—usually 0 or null). The occurrence of zero elements in a large array is inefficient for both computation and storage. An array in which there are a large number of zero elements is referred to as being sparse.

In the case of sparse arrays, we can ask for a value from an "empty" array position. If we do this, then for array of numbers, it should return zero, and for array of objects, it should return null.

A naive implementation of an array may allocate space for the entire array, but in the case where there are few non-default values, this implementation is inefficient. Typically the algorithm used instead of an ordinary array is determined by other known features (or statistical features) of the array; for instance, if the sparsity is known in advance, or if the elements are arranged according to some function (e.g. occur in blocks).

A heap memory allocator in a program might choose to store regions of blank space in a linked list rather than storing all of the allocated regions in, say a bit array.

## Representation

Sparse Array can be represented as

`Sparse_Array[{pos1 -> val1, pos2 -> val2,...}]` or  
`Sparse_Array[{pos1, pos2,...} -> {val1, val2,...}]`

which yields a sparse array in which values  $val_i$  appear at positions  $pos_i$ .

## Sparse Array as Linked List

An obvious question that might be asked is why we need a linked list to represent a sparse array if we can represent it easily using a normal array. The answer to this question lies in the fact that while representing a sparse array as a normal array, a lot of space is allocated for zero or null elements. For example, consider the following array declaration:

```
double arr[1000][1000];
```

When we define this array as enough space of 1,000,000 doubles is allocated. As each double requires 8 bytes of memory, this array will require 8 million bytes of memory. Because this is a sparse array, most of its elements will have a value of zero (or null). Hence, defining this array will soak up all this space which will result in wastage of memory. An effective way to overcome this problem is to represent the array using a linked list which requires less memory as only elements having non-zero value are stored. Also, when a linked list is used, the array elements can be accessed through fewer iterations than in a normal array.

A sparse array as a linked list contains nodes linked to each other. In a one-dimensional sparse array, each node consist of an "index" (position) of the non-zero element and the "value" at that position and a node pointer "next"(for linking to the next node), nodes are linked in order as per the index. In the case of a two-dimensional sparse array, each node contains a row index, a column index (which together gives us its position), a value at that position and a pointer to the next node.

---

## External links

- Boost sparse vector class <sup>[1]</sup>
- Rodolphe Buda, "Two Dimensional Aggregation Procedure: An Alternative to the Matrix Algebraic Algorithm", *Computational Economics*, 31(4), May, pp.397–408, 2008. <sup>[2]</sup>

## References

[1] [http://boost.org/libs/numeric/ublas/doc/vector\\_sparse.htm](http://boost.org/libs/numeric/ublas/doc/vector_sparse.htm)

[2] <http://portal.acm.org/citation.cfm?id=1363086&jmp=cit&coll=GUIDE&dl=GUIDE>

# Bit array

A **bit array** (also known as **bitmap**, **bitset**, **bit string**, or **bit vector**) is an array data structure that compactly stores bits. It can be used to implement a simple set data structure. A bit array is effective at exploiting bit-level parallelism in hardware to perform operations quickly. A typical bit array stores  $kw$  bits, where  $w$  is the number of bits in the unit of storage, such as a byte or word, and  $k$  is some nonnegative integer. If  $w$  does not divide the number of bits to be stored, some space is wasted due to internal fragmentation.

## Definition

A bit array is a mapping from some domain (almost always a range of integers) to values in the set  $\{0, 1\}$ . The values can be interpreted as dark/light, absent/present, locked/unlocked, valid/invalid, et cetera. The point is that there are only two possible values, so they can be stored in one bit. The array can be viewed as a subset of the domain (e.g.  $\{0, 1, 2, \dots, n-1\}$ ), where a 1 bit indicates a number in the set and a 0 bit a number not in the set. This set data structure uses about  $n/w$  words of space, where  $w$  is the number of bits in each machine word. Whether the least significant bit or the most significant bit indicates the smallest-index number is largely irrelevant, but the former tends to be preferred.

## Basic operations

Although most machines are not able to address individual bits in memory, nor have instructions to manipulate single bits, each bit in a word can be singled out and manipulated using bitwise operations. In particular:

- OR can be used to set a bit to one:  $11101010 \text{ OR } 00000100 = 11101110$
- AND can be used to set a bit to zero:  $11101010 \text{ AND } 11111101 = 11101000$
- AND together with zero-testing can be used to determine if a bit is set:

$$11101010 \text{ AND } 00000001 = 00000000 = 0$$

$$11101010 \text{ AND } 00000010 = 00000010 \neq 0$$

- XOR can be used to invert or toggle a bit:

$$11101010 \text{ XOR } 00000100 = 11101110$$

$$11101110 \text{ XOR } 00000100 = 11101010$$

To obtain the bit mask needed for these operations, we can use a bit shift operator to shift the number 1 to the left by the appropriate number of places, as well as bitwise negation if necessary.

Given two bit arrays of the same size representing sets, we can compute their union, intersection, and set-theoretic difference using  $n/w$  simple bit operations each ( $2n/w$  for difference), as well as the complement of either:

```
for i from 0 to n/w-1
    complement_a[i] := not a[i]
```

```

union[i]      := a[i] or b[i]
intersection[i] := a[i] and b[i]
difference[i]  := a[i] and (not b[i])

```

If we wish to iterate through the bits of a bit array, we can do this efficiently using a doubly nested loop that loops through each word, one at a time. Only  $n/w$  memory accesses are required:

```

for i from 0 to n/w-1
  index := 0      // if needed
  word := a[i]
  for b from 0 to w-1
    value := word and 1 ≠ 0
    word := word shift right 1
    // do something with value
    index := index + 1    // if needed

```

Both of these code samples exhibit ideal locality of reference, and so get a large performance boost from a data cache. If a cache line is  $k$  words, only about  $n/wk$  cache misses will occur.

## More complex operations

### Population / Hamming weight

If we wish to find the number of 1 bits in a bit array, sometimes called the population count or Hamming weight, there are efficient branch-free algorithms that can compute the number of bits in a word using a series of simple bit operations. We simply run such an algorithm on each word and keep a running total. Counting zeros is similar. See the Hamming weight article for examples of an efficient implementation.

### Sorting

Similarly, sorting a bit array is trivial to do in  $O(n)$  time using counting sort — we count the number of ones  $k$ , fill the last  $k/w$  words with ones, set only the low  $k \bmod w$  bits of the next word, and set the rest to zero.

### Inversion

Vertical flipping of a one-bit-per-pixel image, or some FFT algorithms, require to flip the bits of individual words (so  $b_{31} \ b_{30} \ \dots \ b_0$  becomes  $b_0 \ \dots \ b_{30} \ b_{31}$ ). When this operation is not available on the processor, it's still possible to proceed by successive passes, in this example on 32 bits:

```

exchange two 16bit halfwords
exchange bytes by pairs (0xddccbbaa -> 0xccddaabb)
...
swap bits by pairs
swap bits (b31 b30 ... b1 b0 -> b30 b31 ... b0 b1)

```

The last operation can be written  $((x \& 0x55555555) \ll 1) \mid (x \& 0xaaaaaaaa) \gg 1$ .

## Find first one

The find first one or *find first set* operation identifies the index or position of the least significant *one* bit in a word, and has widespread hardware support and efficient algorithms for its computation. When a priority queue is stored in a bit array, find first one can be used to identify the highest priority element in the queue. To expand a word-size *find first one* to longer arrays, one can find the first nonzero word and then run *find first one* on that word. The related operations *find first zero*, *count leading zeros*, *count leading ones*, *count trailing zeros*, *count trailing ones*, and *log base 2* (see find first set) can also be extended to a bit array in a straightforward manner.

## Compression

Large bit arrays tend to have long streams of zeroes or ones. This phenomenon wastes storage and processing time. Run-length encoding is commonly used to compress these long streams. However, by compressing bit arrays too aggressively we run the risk of losing the benefits due to bit-level parallelism (vectorization). Thus, instead of compressing bit arrays as streams of bits, we might compress them as streams bytes or words (see Bitmap index (compression)).

Examples:

- compressedbitset<sup>[1]</sup>: WAH Compressed BitSet for Java
- javaewah<sup>[2]</sup>: A compressed alternative to the Java BitSet class (using Enhanced WAH)
- CONCISE<sup>[3]</sup>: COmpressed 'N' Composable Integer Set, another bitmap compression scheme for Java
- EWAHBoolArray<sup>[4]</sup>: A compressed bitmap/bitset class in C++
- CSharpEWAH<sup>[5]</sup>: compressed bitset class in C#

## Advantages and disadvantages

Bit arrays, despite their simplicity, have a number of marked advantages over other data structures for the same problems:

- They are extremely compact; few other data structures can store  $n$  independent pieces of data in  $n/w$  words.
- They allow small arrays of bits to be stored and manipulated in the register set for long periods of time with no memory accesses.
- Because of their ability to exploit bit-level parallelism, limit memory access, and maximally use the data cache, they often outperform many other data structures on practical data sets, even those that are more asymptotically efficient.

However, bit arrays aren't the solution to everything. In particular:

- Without compression, they are wasteful set data structures for sparse sets (those with few elements compared to their range) in both time and space. For such applications, compressed bit arrays, Judy arrays, tries, or even Bloom filters should be considered instead.
- Accessing individual elements can be expensive and difficult to express in some languages. If random access is more common than sequential and the array is relatively small, a byte array may be preferable on a machine with byte addressing. A word array, however, is probably not justified due to the huge space overhead and additional cache misses it causes, unless the machine only has word addressing.

## Applications

Because of their compactness, bit arrays have a number of applications in areas where space or efficiency is at a premium. Most commonly, they are used to represent a simple group of boolean flags or an ordered sequence of boolean values.

Bit arrays are used for priority queues, where the bit at index  $k$  is set if and only if  $k$  is in the queue; this data structure is used, for example, by the Linux kernel, and benefits strongly from a find-first-zero operation in hardware.

Bit arrays can be used for the allocation of memory pages, inodes, disk sectors, etc. In such cases, the term *bitmap* may be used. However, this term is frequently used to refer to raster images, which may use multiple bits per pixel.

Another application of bit arrays is the Bloom filter, a probabilistic set data structure that can store large sets in a small space in exchange for a small probability of error. It is also possible to build probabilistic hash tables based on bit arrays that accept either false positives or false negatives.

Bit arrays and the operations on them are also important for constructing succinct data structures, which use close to the minimum possible space. In this context, operations like finding the  $n$ th 1 bit or counting the number of 1 bits up to a certain position become important.

Bit arrays are also a useful abstraction for examining streams of compressed data, which often contain elements that occupy portions of bytes or are not byte-aligned. For example, the compressed Huffman coding representation of a single 8-bit character can be anywhere from 1 to 255 bits long.

In information retrieval, bit arrays are a good representation for the posting lists of very frequent terms. If we compute the gaps between adjacent values in a list of strictly increasing integers and encode them using unary coding, the result is a bit array with a 1 bit in the  $n$ th position if and only if  $n$  is in the list. The implied probability of a gap of  $n$  is  $1/2^n$ . This is also the special case of Golomb coding where the parameter  $M$  is 1; this parameter is only normally selected when  $-\log(2-p)/\log(1-p) \leq 1$ , or roughly the term occurs in at least 38% of documents.

## Language support

The C programming language's *bitfields*, pseudo-objects found in structs with size equal to some number of bits, are in fact small bit arrays; they are limited in that they cannot span words. Although they give a convenient syntax, the bits are still accessed using bitwise operators on most machines, and they can only be defined statically (like C's static arrays, their sizes are fixed at compile-time). It is also a common idiom for C programmers to use words as small bit arrays and access bits of them using bit operators. A widely available header file included in the X11 system, `xtrapbits.h`, is "a portable way for systems to define bit field manipulation of arrays of bits.". A more explanatory description of aforementioned approach can be found in the `comp.lang.c` faq <sup>[6]</sup>.

In C++, although individual `bool`s typically occupy the same space as a byte or an integer, the STL type `vector<bool>` is a partial template specialization in which bits are packed as a space efficiency optimization. Since bytes (and not bits) are the smallest addressable unit in C++, the `[]` operator does *not* return a reference to an element, but instead returns a proxy reference. This might seem a minor point, but it means that `vector<bool>` is *not* a standard STL container, which is why the use of `vector<bool>` is generally discouraged. Another unique STL class, `bitset`,<sup>[7]</sup> creates a vector of bits fixed at a particular size at compile-time, and in its interface and syntax more resembles the idiomatic use of words as bit sets by C programmers. It also has some additional power, such as the ability to efficiently count the number of bits that are set. The Boost C++ Libraries provide a `dynamic_bitset` class<sup>[8]</sup> whose size is specified at run-time.

The D programming language provides bit arrays in both of its competing standard libraries. In Phobos, they are provided in `std.bitmanip`, and in Tango, they are provided in `tango.core.BitArray`. As in C++, the `[]` operator does not return a reference, since individual bits are not directly addressable on most hardware, but instead returns a `bool`.

In Java, the class `BitSet` creates a bit array that is then manipulated with functions named after bitwise operators familiar to C programmers. Unlike the `bitset` in C++, the Java `BitSet` does not have a "size" state (it has an effectively infinite size, initialized with 0 bits); a bit can be set or tested at any index. In addition, there is a class `EnumSet`, which represents a Set of values of an enumerated type internally as a bit vector, as a safer alternative to bitfields.

The .NET Framework supplies a `BitArray` collection class. It stores boolean values, supports random access and bitwise operators, can be iterated over, and its `Length` property can be changed to grow or truncate it.

Although Standard ML has no support for bit arrays, Standard ML of New Jersey has an extension, the `BitArray` structure, in its SML/NJ Library. It is not fixed in size and supports set operations and bit operations, including, unusually, shift operations.

Haskell likewise currently lacks standard support for bitwise operations, but both GHC and Hugs provide a `Data.Bits` module with assorted bitwise functions and operators, including shift and rotate operations and an "unboxed" array over boolean values may be used to model a Bit array, although this lacks support from the former module.

In Perl, strings can be used as expandable bit arrays. They can be manipulated using the usual bitwise operators (`~`, `|`, `&`, `^`),<sup>[9]</sup> and individual bits can be tested and set using the `vec` function.<sup>[10]</sup>

In Ruby, you can access (but not set) a bit of an integer (`Fixnum` or `Bignum`) using the bracket operator (`[]`), as if it were an array of bits.

Apple's Core Foundation library contains `CFBitVector`<sup>[11]</sup> and `CFMutableBitVector`<sup>[12]</sup> structures.

PL/I supports arrays of *bit strings* of arbitrary length, which may be either fixed-length or varying. The array elements may be *aligned*— each element begins on a byte or word boundary— or *unaligned*— elements immediately follow each other with no padding,

## References

- [1] <http://code.google.com/p/compressedbitset/>
- [2] <http://code.google.com/p/javaewah/>
- [3] <http://ricerca.mat.uniroma3.it/users/colanton/concise.html>
- [4] <http://github.com/lemire/EWAHBoolArray>
- [5] <http://code.google.com/p/csharpewah/>
- [6] <http://c-faq.com/misc/bitsets.html>
- [7] `std::bitset` (<http://www.sgi.com/tech/stl/bitset.html>)
- [8] `boost::dynamic_bitset` ([http://www.boost.org/libs/dynamic\\_bitset/dynamic\\_bitset.html](http://www.boost.org/libs/dynamic_bitset/dynamic_bitset.html))
- [9] <http://perldoc.perl.org/perlop.html#Bitwise-String-Operators>
- [10] <http://perldoc.perl.org/functions/vec.html>
- [11] <http://developer.apple.com/library/mac/#documentation/CoreFoundation/Reference/CFBitVectorRef/Reference/reference.html>
- [12] [http://developer.apple.com/library/mac/#documentation/CoreFoundation/Reference/CFMutableBitVectorRef/Reference/reference.html#/apple\\_ref/doc/uid/20001500](http://developer.apple.com/library/mac/#documentation/CoreFoundation/Reference/CFMutableBitVectorRef/Reference/reference.html#/apple_ref/doc/uid/20001500)

## External links

- mathematical bases (<http://www-cs-faculty.stanford.edu/~knuth/fasc1a.ps.gz>) by Pr. D.E.Knuth
- bitarray module (<http://pypi.python.org/pypi/bitarray>) for Python
- vector<bool> Is Nonconforming, and Forces Optimization Choice (<http://www.gotw.ca/publications/N1185.pdf>)
- vector<bool>: More Problems, Better Solutions (<http://www.gotw.ca/publications/N1211.pdf>)

## Bitboard

---

A **bitboard** is a data structure commonly used in computer systems that play board games.

A bitboard, often used for boardgames such as chess, checkers and othello, is a specialization of the bitset data structure, where each bit represents a game position or state, designed for optimization of speed and/or memory or disk use in mass calculations. Bits in the same bitboard relate to each other in the rules of the game often forming a game position when taken together. Other bitboards are commonly used as masks to transform or answer queries about positions. The "game" may be any game-like system where information is tightly packed in a structured form with "rules" affecting how the individual units or pieces relate.

### Short description

Bitboards are used in many of the world's best chess playing programs. They help the programs analyze chess positions with few CPU instructions and hold a massive number of positions in memory efficiently.

Bitboards are interesting because they allow the computer to answer some questions about game state with one logical operation. For example, if a chess program wants to know if the white player has any pawns in the center of the board (center four squares) it can just compare a bitboard for the player's pawns with one for the center of the board using a logical AND operation. If there are no center pawns then the result will be zero.

Query results can also be represented using bitboards. For example, the query "What are the squares between X and Y?" can be represented as a bitboard. These query results are generally pre-calculated, so that a program can simply retrieve a query result with one memory load.

However, as a result of the massive compression and encoding, bitboard programs are not easy for software developers to either write or debug.

### History

The bitboard method for holding a board game appears to have been invented in the mid-1950s, by Arthur Samuel and was used in his checkers program. The method was published in 1959 as "Some Studies in Machine Learning Using the Game of Checkers" in the IBM Journal of Research and Development.

For the more complicated game of chess, it appears the method was independently rediscovered later by the Kaissa team in the Soviet Union in the late 1960s, although not publicly documented, and again by the authors of the U.S. Northwestern University program "Chess" in the early 1970s, and documented in 1977 in "Chess Skill in Man and Machine".

---



## Description for all games or applications

A bitboard or bit field is a format that stuffs a whole group of related boolean variables into the same integer, typically representing positions on a board game. Each bit is a position, and when the bit is positive, a property of that position is true. In chess, for example, there would be a bitboard for black knights. There would be 64-bits where each bit represents a chess square. Another bitboard might be a constant representing the center four squares of the board. By comparing the two numbers with a bitwise logical AND instruction, we get a third bitboard which represents the black knights on the center four squares, if any. This format is generally more CPU and memory friendly than others.

## General technical advantages and disadvantages

### Processor use

#### Pros

The advantage of the bitboard representation is that it takes advantage of the essential logical bitwise operations available on nearly all CPUs that complete in one cycle and are full pipelined and cached etc. Nearly all CPUs have AND, OR, NOR, and XOR. Many CPUs have additional bit instructions, such as finding the "first" bit, that make bitboard operations even more efficient. If they do not have instructions well known algorithms can perform some "magic" transformations that do these quickly.

Furthermore, modern CPUs have instruction pipelines that queue instructions for execution. A processor with multiple execution units can perform more than one instruction per cycle if more than one instruction is available in the pipeline. Branching (the use of conditionals like if) makes it harder for the processor to fill its pipeline(s) because the CPU can't tell what it needs to do in advance. Too much branching makes the pipeline less effective and potentially reduces the number of instructions the processor can execute per cycle. Many bitboard operations require fewer conditionals and therefore increase pipelining and make effective use of multiple execution units on many CPUs.

CPUs have a bit width which they are designed toward and can carry out bitwise operations in one cycle in this width. So, on a 64-bit or more CPU, 64-bit operations can occur in one instruction. There may be support for higher or lower width instructions. Many 32-bit CPUs may have some 64-bit instructions and those may take more than one cycle or otherwise be handicapped compared to their 32-bit instructions.

If the bitboard is larger than the width of the instruction set, then a performance hit will be the result. So a program using 64-bit bitboards would run faster on a real 64-bit processor than on a 32-bit processor.

#### Cons

Some queries are going to take longer than they would with perhaps arrays, so bitboards are generally used in conjunction with array boards in chess programs.

### Memory use

#### Pros

Bitboards are extremely compact. Since only a very small amount of memory is required to represent a position or a mask, more positions can find their way into registers, full speed cache, Level 2 cache, etc. In this way, compactness translates into better performance (on most machines). Also on some machines this might mean that more positions can be stored in main memory before going to disk.

## Cons

For some games writing a suitable bitboard engine requires a fair amount of source code that will be longer than the straight forward implementation. For limited devices (like cell phones) with a limited number of registers or processor instruction cache, this can cause a problem. For full-sized computers it may cause cache misses between level one and level two cache. This is a potential problem—not a major drawback. Most machines will have enough instruction cache so that this isn't an issue.

## Chess bitboards

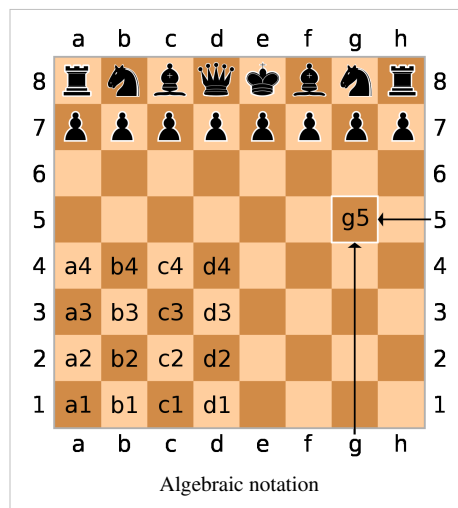
### Standard

The first bit usually represents the square a1 (the lower left square), and the 64th bit represents the square h8 (the diagonally opposite square).

There are twelve types of pieces, and each type gets its own bitboard. Black pawns get a board, white pawns, etc. Together these twelve boards can represent a position. Some trivial information also needs to be tracked elsewhere; the programmer may use boolean variables for whether each side is in check, can castle, etc.

Constants are likely available, such as `WHITE_SQUARES`, `BLACK_SQUARES`, `FILE_A`, `RANK_4` etc. More interesting ones might include `CENTER`, `CORNERS`, `CASTLE_SQUARES`, etc.

Examples of variables would be `WHITE_ATTACKING`, `ATTACKED_BY_PAWN`, `WHITE_PASSED_PAWN`, etc.



### Rotated

"Rotated" bitboards are usually used in programs that use bitboards. Rotated bitboards make certain operations more efficient. While engines are simply referred to as "rotated bitboard engines," this is a misnomer as rotated boards are used in *addition* to normal boards making these hybrid standard/rotated bitboard engines.

These bitboards rotate the bitboard positions by 90 degrees, 45 degrees, and/or 315 degrees. A typical bitboard will have one byte per rank of the chess board. With this bitboard it's easy to determine rook attacks across a rank, using a table indexed by the occupied square and the occupied positions in the rank (because rook attacks stop at the first occupied square). By rotating the bitboard 90 degrees, rook attacks across a file can be examined the same way. Adding bitboards rotated 45 degrees and 315 degrees produces bitboards in which the diagonals are easy to examine. The queen can be examined by combining rook and bishop attacks. Rotated bitboards appear to have been developed separately and (essentially) simultaneously by the developers of the DarkThought and Crafty programs.

### Magics

Magic move bitboard generation is a new and fast alternative to rotated move bitboard generators. These are also more versatile than rotated move bitboard generators because the generator can be used independently from any position. The basic idea is that you can use a multiply, right-shift hashing function to index a move database, which can be as small as 1.5K. A speedup is gained because no rotated bitboards need to be updated, and because the lookups are more cache-friendly.

## Other bitboards

Many other games besides chess benefit from bitboards.

- In Connect Four, they allow for very efficient testing for four consecutive discs, by just two shift+and operations per direction.
- In the Conway's Game of Life, they are a possible alternative to arrays.
- Othello/Reversi (see the Reversi article).

## External links

### Checkers

- Checkers Bitboard Tutorial <sup>[1]</sup> by Jonathan Kreuzer

### Chess

#### Articles

- Programming area of the Beowulf project <sup>[2]</sup>
- Heinz, Ernst A. How DarkThought plays chess. *ICCA Journal*, Vol. 20(3), pp. 166-176, Sept. 1997 <sup>[3]</sup>
- Laramee, Francois-Dominic. Chess Programming Part 2: Data Structures. <sup>[4]</sup>
- Verhelst, Paul. Chess Board Representations <sup>[5]</sup>
- Hyatt, Robert. Chess program board representations <sup>[6]</sup>
- Hyatt, Robert. Rotated bitmaps, a new twist on an old idea <sup>[7]</sup>
- Frayn, Colin. How to implement bitboards in a chess engine (chess programming theory) <sup>[8]</sup>
- Pepicelli, Glen. *Bitfields, Bitboards, and Beyond* <sup>[9]</sup> - (Example of bitboards in the Java Language and a discussion of why this optimization works with the Java Virtual Machine (www.OnJava.com publisher: O'Reilly 2005))
- Magic Move-Bitboard Generation in Computer Chess. Pradyumna Kannan <sup>[10]</sup>

#### Code examples

- <sup>[11]</sup> The author of the Frenzee engine had posted some source examples.

link not working please update

- [12] A 155 line java Connect-4 program demonstrating the use of bitboards.

#### Implementations

##### Open source

- Beowulf <sup>[13]</sup> Unix, Linux, Windows. Rotated bitboards.
  - Crafty See the Crafty article. Written in straight C. Rotated bitboards in the old versions, now uses magic bitboards. Strong.
  - GNU Chess See the GNU Chess Article.
  - Stockfish UCI chess engine ranking second in Elo as of 2010
  - Gray Matter <sup>[14]</sup> C++, rotated bitboards.
  - KnightCap GPL. ELO of 2300.
  - Pepito <sup>[15]</sup> C. Bitboard, by Carlos del Cacho. Windows and Linux binaries as well as source available.
  - Simontacci <sup>[16]</sup> Rotated bitboards.
-

### Closed source

- DarkThought Home Page <sup>[17]</sup>

### Othello

- A complete discussion <sup>[18]</sup> of Othello (Reversi) engines with some source code including an Othello bitboard in C and assembly.

## References

- [1] <http://www.3dkingdoms.com/checkers/bitboards.htm>
- [2] <http://www.frayn.net/beowulf/theory.html>
- [3] <http://supertech.lcs.mit.edu/~heinz/dt/node2.html>
- [4] <http://www.gamedev.net/reference/programming/features/chess2/page3.asp>
- [5] <http://chess.verhelst.org/1997/03/10/representations/>
- [6] <http://www.cis.uab.edu/info/faculty/hyatt/boardrep.html>
- [7] <http://www.cis.uab.edu/info/faculty/hyatt/bitmaps.html>
- [8] <http://www.frayn.net/beowulf/theory.html#bitboards>
- [9] <http://www.onjava.com/pub/a/onjava/2005/02/02/bitsets.html>
- [10] [http://www.pradu.us/old/Nov27\\_2008/Buzz/research/magic/Bitboards.pdf](http://www.pradu.us/old/Nov27_2008/Buzz/research/magic/Bitboards.pdf)
- [11] <http://web.archive.org/web/20061204195709/http://www.geocities.com/ruleren/sources.html>
- [12] <http://www.cwi.nl/~tromp/c4/Connect4.java>
- [13] <http://www.frayn.net/beowulf/index.html>
- [14] <http://code.google.com/p/gray-matter/>
- [15] <http://www.quarkchess.de/pepito/>
- [16] <http://simontacchi.sourceforge.net/>
- [17] <http://supertech.lcs.mit.edu/~heinz/dt/>
- [18] <http://www.radagast.se/othello/>

## Parallel array

In computing, a **parallel array** is a data structure for representing arrays of records. It keeps a separate, homogeneous array for each field of the record, each having the same number of elements. Then, objects located at the same index in each array are implicitly the fields of a single record. Pointers from one object to another are replaced by array indices. This contrasts with the normal approach of storing all fields of each record together in memory. For example, one might declare an array of 100 names, each a string, and 100 ages, each an integer, associating each name with the age that has the same index.

An example in C using parallel arrays:

```
int  ages[]    = {0,          17,          2,          52,          25};
char *names[] = {"None",     "Mike",      "Billy",     "Tom",
"Stan"};
int  parent[] = {0 /*None*/, 3 /*Tom*/,  1 /*Mike*/, 0 /*None*/, 3
/*Tom*/};

for(i = 1; i <= 4; i++) {
    printf("Name: %s, Age: %d, Parent: %s \n",
           names[i], ages[i], names[parent[i]]);
}
```

in Perl (using a hash of arrays to hold references to each array):

```

my %data = (
    first_name   => ['Joe',   'Bob',   'Frank', 'Hans'   ],
    last_name    => ['Smith', 'Seger', 'Sinatra', 'Schultze'],
    height_in_cm => [169,     158,     201,     199     ]);

for $i (0..${#{$data{first_name}}}) {
    printf "Name: %s %s\n", $data{first_name}[$i], $data{last_name}[$i];
    printf "Height in CM: %i\n", $data{height_in_cm}[$i];
}

```

Or, in Python:

```

firstName = ['Joe', 'Bob', 'Frank', 'Hans' ]
lastName  = ['Smith', 'Seger', 'Sinatra', 'Schultze']
heightInCM = [169, 158, 201, 199 ]

for i in xrange(len(firstName)):
    print "Name: %s %s" % (firstName[i], lastName[i])
    print "Height in CM: %s" % heightInCM[i]

```

Parallel arrays have a number of practical advantages over the normal approach:

- They can be used in languages which support only arrays of primitive types and not of records (or perhaps don't support records at all).
- Parallel arrays are simple to understand and use, and are often used where declaring a record is more trouble than it's worth.
- They can save a substantial amount of space in some cases by avoiding alignment issues. For example, one of the fields of the record can be a single bit, and its array would only need to reserve one bit for each record, whereas in the normal approach many more bits would "pad" the field so that it consumes an entire byte or a word.
- If the number of items is small, array indices can occupy significantly less space than full pointers, particularly on architectures with large words.
- Sequentially examining a single field of each record in the array is very fast on modern machines, since this amounts to a linear traversal of a single array, exhibiting ideal locality of reference and cache behavior.

However, parallel arrays also have several strong disadvantages, which serves to explain why they are not generally preferred:

- They have significantly worse locality of reference when visiting the records sequentially and examining multiple fields of each record, which is the norm.
- They obscure the relationship between fields of a single record.
- They have little direct language support (the language and its syntax typically express no relationship between the arrays in the parallel array).
- They are expensive to grow or shrink, since each of several arrays must be reallocated. Multi-level arrays can ameliorate this problem, but impacts performance due to the additional indirection needed to find the desired elements.

The bad locality of reference is the worst issue. However, a compromise can be made in some cases: if a structure can be divided into groups of fields that are generally accessed together, an array can be constructed for each group, and its elements are records containing only these subsets of the larger structure's fields. This is a valuable way of speeding up access to very large structures with many members, while keeping the portions of the structure tied together. An alternative to tying them together using array indexes is to use references to tie the portions together, but this can be less efficient in time and space. Another alternative is to mock up a record structure in a

single-dimensional array by declaring an array of  $n \times m$  size and referring to the  $r$ -th field in record  $i$  as element as  $\text{array}(m \times i + r)$ . Some compiler optimizations, particularly for vector processors, are able to perform this transformation automatically when arrays of structures are created in the program.

## References

- Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7. Page 209 of section 10.3: Implementing pointers and objects.

## Lookup table

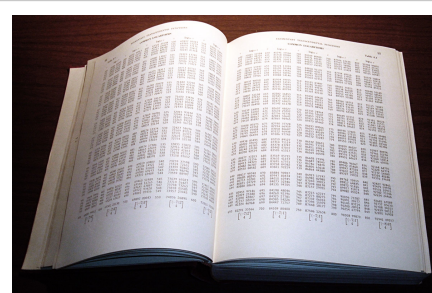
In computer science, a **lookup table** is a data structure, usually an array or associative array, often used to replace a runtime computation with a simpler array indexing operation. The savings in terms of processing time can be significant, since retrieving a value from memory is often faster than undergoing an 'expensive' computation or input/output operation.<sup>[1]</sup> The tables may be precalculated and stored in static program storage, calculated (or "pre-fetched") as part of a programs initialization phase (memoization), or even store in hardware in application-specific platforms. Lookup tables are also used extensively to validate input values by matching against a list of valid (or invalid) items in an array and, in some programming languages, may include pointer functions (or offsets to labels) to process the matching input.

## History

Before the advent of computers, lookup tables of values were used by people to speed up hand calculations of complex functions, such as in trigonometry, logarithms, and statistical density functions<sup>[2]</sup>

In ancient India, Aryabhata created one of the first sine tables, which he encoded in a Sanskrit-letter-based number system. In 493 A.D., Victorius of Aquitaine wrote a 98-column multiplication table which gave (in Roman numerals) the product of every number from 2 to 50 times and the rows were "a list of numbers starting with one thousand, descending by hundreds to one hundred, then descending by tens to ten, then by ones to one, and then the fractions down to 1/144"<sup>[3]</sup> Modern school children are often taught to memorize "times tables" to avoid calculations of the most commonly used numbers (up to  $9 \times 9$  or  $12 \times 12$ ).

Early in the history of computers, input/output operations were particularly slow – even in comparison to processor speeds of the time. It made sense to reduce expensive read operations by a form of manual caching by creating either static lookup tables (embedded in the program) or dynamic prefetched arrays to contain only the most commonly occurring data items. Despite the introduction of systemwide caching that now automates this process, application level lookup tables can still improve performance for data items that rarely, if ever, change.



Part of a 20th century table of common logarithms in the reference book Abramowitz and Stegun.

## Examples

### Simple lookup in an array, an associative array or a linked list (unsorted list)

This is known as a linear search or brute-force search, each element being checked for equality in turn and the associated value, if any, used as a result of the search. This is often the slowest search method unless frequently occurring values occur early in the list. For a one dimensional array or linked list, the lookup is usually to determine whether or not there is a match with an 'input' data value.

#### Linked lists vs. arrays

Linked lists have some advantages over arrays:

- Insertion or deletion of an element at a specific point of a list is a constant time operation. (While one can "delete" an element from an array in constant time by somehow marking its slot as "vacant", an algorithm that iterates over the elements may have to skip a large number of vacant slots).
- arbitrarily many elements may be inserted into a linked list, limited only by the total memory available; while an array will eventually fill up, and then have to be resized — an expensive operation, that may not even be possible if memory is fragmented. Similarly, an array from which many elements are removed, may have to be resized in order to avoid wasting too much space.

On the other hand:

- arrays allow random access, while linked lists allow only sequential access to elements. Singly linked lists, in fact, can only be traversed in one direction. This makes linked lists unsuitable for applications where it's useful to quickly look up an element by its index, such as heapsort. See also trivial hash function below.
- Sequential access on arrays is also faster than on linked lists on many machines, because they have greater locality of reference and thus benefit more from processor caching.
- linked lists require extra storage needed for references, that often makes them impractical for lists of small data items such as characters or boolean values. It can also be slow, and with a naïve allocator, wasteful, to allocate memory separately for each new element, a problem generally solved using memory pools.

Some hybrid solutions try to combine the advantages of the two representations. Unrolled linked lists store several elements in each list node, increasing cache performance while decreasing memory overhead for references. CDR coding does both these as well, by replacing references with the actual data referenced, which extends off the end of the referencing record.

### Binary search in an array or an associative array (sorted list)

An example of a "divide and conquer algorithm", binary search involves each element being found by determining which half of the table a match may be found in and repeating until either success or failure. Only possible if the list is sorted but gives good performance even if the list is lengthy.

#### Trivial hash function

For a trivial hash function lookup, the unsigned raw data value is used *directly* as an index to a one dimensional table to extract a result. For small ranges, this can be amongst the fastest lookup, even exceeding binary search speed with zero branches and executing in constant time.

---

### Counting '1' bits in a series of bytes

One discrete problem that is expensive to solve on many computers, is that of counting the number of bits which are set to 1 in a (binary) number, sometimes called the *population function*. For example, the decimal number "37" is "00100101" in binary, so it contains three bits that are set to binary "1".

A simple example of C code, designed to count the 1 bits in a *int*, might look like this:

```
int count_ones(unsigned int x) {
    int result = 0;
    while (x != 0)
        result++, x = x & (x-1);
    return result;
}
```

This apparently simple algorithm can take potentially hundreds of cycles even on a modern architecture, because it makes many branches in the loop - and branching is slow. This can be ameliorated using loop unrolling and some other compiler optimizations. There is however a simple and much faster algorithmic solution - using a trivial hash function table lookup.

Simply construct a static table, *bits\_set*, with 256 entries giving the number of one bits set in each possible byte value (e.g. 0x00 = 0, 0x01 = 1, 0x02 = 1, and so on). Then use this table to find the number of ones in each byte of the integer using a trivial hash function lookup on each byte in turn, and sum them. This requires no branches, and just four indexed memory accesses, considerably faster than the earlier code.

```
/* (this code assumes that 'int' is 32-bits wide) */
int count_ones(unsigned int x) {
    return bits_set[ x          & 255] + bits_set[ (x >> 8) & 255]
        + bits_set[ (x >> 16) & 255] + bits_set[ (x >> 24) & 255];
}
```

The above source can be improved easily, (avoiding AND'ing, and shifting) by 'recasting' 'x' as a 4 byte unsigned char array and, preferably, coded in-line as a single statement instead of being a function. Note that even this simple algorithm can be too slow now, because the original code might run faster from the cache of modern processors, and (large) lookup tables do not fit well in caches and can cause a slower access to memory (in addition, in the above example, it requires computing addresses within a table, to perform the four lookups needed).

### Lookup tables in image processing

In data analysis applications, such as image processing, a lookup table (LUT) is used to transform the input data into a more desirable output format. For example, a grayscale picture of the planet Saturn will be transformed into a color image to emphasize the differences in its rings.

A classic example of reducing run-time computations using lookup tables is to obtain the result of a trigonometry calculation, such as the sine of a value. Calculating trigonometric functions can substantially slow a computing application. The same application can finish much sooner when it first precalculates the sine of a number of values, for example for each whole number of degrees (The table can be defined as static variables at compile time, reducing repeated run time costs). When the program requires the sine of a value, it can use the lookup table to retrieve the closest sine value from a memory address, and may also take the step of interpolating to the sine of the desired value, instead of calculating by mathematical formula. Lookup tables are thus used by mathematics co-processors in computer systems. An error in a lookup table was responsible for Intel's infamous floating-point divide bug.

Functions of a single variable (such as sine and cosine) may be implemented by a simple array. Functions involving two or more variables require multidimensional array indexing techniques. The latter case may thus employ a



two-dimensional array of **power[x][y]** to replace a function to calculate  $x^y$  for a limited range of  $x$  and  $y$  values. Functions that have more than one result may be implemented with lookup tables that are arrays of structures.

As mentioned, there are intermediate solutions that use tables in combination with a small amount of computation, often using interpolation. Pre-calculation combined with interpolation can produce higher accuracy for values that fall between two precomputed values. This technique requires slightly more time to be performed but can greatly enhance accuracy in applications that require the higher accuracy. Depending on the values being precomputed, pre-computation with interpolation can also be used to shrink the lookup table size while maintaining accuracy.

In image processing, lookup tables are often called **LUTs** and give an output value for each of a range of index values. One common LUT, called the *colormap* or *palette*, is used to determine the colors and intensity values with which a particular image will be displayed. In computed tomography, "windowing" refers to a related concept for determining how to display the intensity of measured radiation..

While often effective, employing a lookup table may nevertheless result in a severe penalty if the computation that the LUT replaces is relatively simple. Memory retrieval time and the complexity of memory requirements can increase application operation time and system complexity relative to what would be required by straight formula computation. The possibility of polluting the cache may also become a problem. Table accesses for large tables will almost certainly cause a cache miss. This phenomenon is increasingly becoming an issue as processors outpace memory. A similar issue appears in rematerialization, a compiler optimization. In some environments, such as the Java programming language, table lookups can be even more expensive due to mandatory bounds-checking involving an additional comparison and branch for each lookup.

There are two fundamental limitations on when it is possible to construct a lookup table for a required operation. One is the amount of memory that is available: one cannot construct a lookup table larger than the space available for the table, although it is possible to construct disk-based lookup tables at the expense of lookup time. The other is the time required to compute the table values in the first instance; although this usually needs to be done only once, if it takes a prohibitively long time, it may make the use of a lookup table an inappropriate solution. As previously stated however, tables can be statically defined in many cases.

## Computing sines

Most computers, which only perform basic arithmetic operations, cannot directly calculate the sine of a given value. Instead, they use the CORDIC algorithm or a complex formula such as the following Taylor series to compute the value of sine to a high degree of precision:

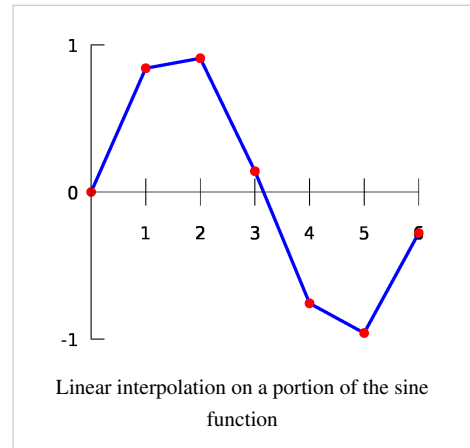
$$\sin(x) \approx x - \frac{x^3}{6} + \frac{x^5}{120} - \frac{x^7}{5040} \text{ (for } x \text{ close to 0)}$$

However, this can be expensive to compute, especially on slow processors, and there are many applications, particularly in traditional computer graphics, that need to compute many thousands of sine values every second. A common solution is to initially compute the sine of many evenly distributed values, and then to find the sine of  $x$  we choose the sine of the value closest to  $x$ . This will be close to the correct value because sine is a continuous function with a bounded rate of change. For example:

```
real array sine_table[-1000..1000]
for x from -1000 to 1000
    sine_table[x] := sine(pi * x / 1000)

function lookup_sine(x)
    return sine_table[round(1000 * x / pi)]
```

Unfortunately, the table requires quite a bit of space: if IEEE double-precision floating-point numbers are used, over 16,000 bytes would be required. We can use fewer samples, but then our precision will significantly worsen. One good solution is linear interpolation, which draws a line between the two points in the table on either side of the value and locates the answer on that line. This is still quick to compute, and much more accurate for smooth functions such as the sine function. Here is our example using linear interpolation:



```
function lookup_sine(x)
    x1 := floor(x*1000/pi)
    y1 := sine_table[x1]
    y2 := sine_table[x1+1]
    return y1 + (y2-y1)*(x*1000/pi-x1)
```

Another solution that uses a quarter of the space but takes a bit longer to compute would be to take into account the relationships between sine and cosine along with their symmetry rules. In this case, the lookup table is calculated by using the sine function for the first quadrant (i.e.  $\sin(0..pi/2)$ ). When we need a value, we assign a variable to be the angle wrapped to the first quadrant. We then wrap the angle to the four quadrants (not needed if values are always between 0 and  $2\pi$ ) and return the correct value (i.e. first quadrant is a straight return, second quadrant is read from  $\pi/2-x$ , third and fourth are negatives of the first and second respectively). For cosine, we only have to return the angle shifted by  $\pi/2$  (i.e.  $x+\pi/2$ ). For tangent, we divide the sine by the cosine (divide-by-zero handling may be needed depending on implementation):

```
function init_sine()
    for x from 0 to (360/4)+1
        sine_table[x] := sine(2*pi * x / 360)

function lookup_sine(x)
    x = wrap x from 0 to 360
    y := mod (x, 90)

    if (x < 90) return sine_table[ y]
    if (x < 180) return sine_table[90-y]
    if (x < 270) return -sine_table[ y]
    return -sine_table[90-y]

function lookup_cosine(x)
    return lookup_sine(x + 90)

function lookup_tan(x)
    return (lookup_sine(x) / lookup_cosine(x))
```

When using interpolation, the size of the lookup table can be reduced by using *nonuniform sampling*, which means that where the function is close to straight, we use few sample points, while where it changes value quickly we use

more sample points to keep the approximation close to the real curve. For more information, see interpolation.

## Other usage of lookup tables

### Caches

Storage caches (including disk caches for files, or processor caches for either code or data) work also like a lookup table. The table is built with very fast memory instead of being stored on slower external memory, and maintains two pieces of data for a subrange of bits composing an external memory (or disk) address (notably the lowest bits of any possible external address):

- one piece (the tag) contains the value of the remaining bits of the address; if these bits match with those from the memory address to read or write, then the other piece contains the cached value for this address.
- the other piece maintains the data associated to that address.

A single (fast) lookup is performed to read the tag in the lookup table at the index specified by the lowest bits of the desired external storage address, and to determine if the memory address is hit by the cache. When a hit is found, no access to external memory is needed (except for write operations, where the cached value may need to be updated asynchronously to the slower memory after some time, or if the position in the cache must be replaced to cache another address).

### Hardware LUTs

In digital logic, an  $n$ -bit lookup table can be implemented with a multiplexer whose select lines are the inputs of the LUT and whose inputs are constants. An  $n$ -bit LUT can encode any  $n$ -input Boolean function by modeling such functions as truth tables. This is an efficient way of encoding Boolean logic functions, and LUTs with 4-6 bits of input are in fact the key component of modern field-programmable gate arrays (FPGAs).

## References

- [1] <http://apl.jhu.edu/~paulmac/c++-memoization.html>
- [2] Campbell-Kelly, Martin; Croarken, Mary; Robson, Eleanor, eds. (October 2, 2003) [2003]. *The History of Mathematical Tables From Sumer to Spreadsheets* (1st ed.). New York, USA: Oxford University Press. ISBN 978-0-19-850841-0.
- [3] Maher, David. W. J. and John F. Makowski. "Literary Evidence for Roman Arithmetic With Fractions", 'Classical Philology' (2001) Vol. 96 No. 4 (2001) pp. 376–399. (See page p.383.)

## External links

- Fast table lookup using input character as index for branch table ([http://en.wikibooks.org/wiki/360\\_Assembly/Branch\\_Instructions](http://en.wikibooks.org/wiki/360_Assembly/Branch_Instructions))
- Art of Assembly: Calculation via Table Lookups (<http://webster.cs.ucr.edu/AoA/Windows/HTML/TableLookups.html>)
- "Bit Twiddling Hacks" (includes lookup tables) (<http://graphics.stanford.edu/~seander/bithacks.html#CountBitsSetTable>) By Sean Eron Anderson of Stanford university
- Memoization in C++ (<http://apl.jhu.edu/~paulmac/c++-memoization.html>) by Paul McNamee, Johns Hopkins University showing savings
- "The Quest for an Accelerated Population Count" ([http://books.google.co.uk/books?id=gJrmszNHQV4C&lpg=PT169&dq=beautiful code "population count"&pg=PT169#v=onepage&q=beautiful code "population count"&f=false](http://books.google.co.uk/books?id=gJrmszNHQV4C&lpg=PT169&dq=beautiful+code+\)) by Henry S. Warren, Jr.

---

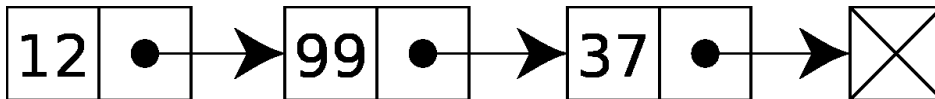
# Lists

---

## Linked list

---

In computer science, a **linked list** is a data structure consisting of a group of nodes which together represent a sequence. Under the simplest form, each node is composed of a datum and a reference (in other words, a *link*) to the next node in the sequence; more complex variants add additional links. This structure allows for efficient insertion or removal of elements from any position in the sequence.



*A linked list whose nodes contain two fields: an integer value and a link to the next node. The last node is linked to a terminator used to signify the end of the list.*

Linked lists are among the simplest and most common data structures. They can be used to implement several other common abstract data types, including stacks, queues, associative arrays, and symbolic expressions, though it is not uncommon to implement the other data structures directly without using a list as the basis of implementation.

The principal benefit of a linked list over a conventional array is that the list elements can easily be inserted or removed without reallocation or reorganization of the entire structure because the data items need not be stored contiguously in memory or on disk. Linked lists allow insertion and removal of nodes at any point in the list, and can do so with a constant number of operations if the link previous to the link being added or removed is maintained during list traversal.

On the other hand, simple linked lists by themselves do not allow random access to the data, or any form of efficient indexing. Thus, many basic operations — such as obtaining the last node of the list (assuming that the last node is not maintained as separate node reference in the list structure), or finding a node that contains a given datum, or locating the place where a new node should be inserted — may require scanning most or all of the list elements.

## History

Linked lists were developed in 1955-56 by Allen Newell, Cliff Shaw and Herbert A. Simon at RAND Corporation as the primary data structure for their Information Processing Language. IPL was used by the authors to develop several early artificial intelligence programs, including the Logic Theory Machine, the General Problem Solver, and a computer chess program. Reports on their work appeared in IRE Transactions on Information Theory in 1956, and several conference proceedings from 1957 to 1959, including Proceedings of the Western Joint Computer Conference in 1957 and 1958, and Information Processing (Proceedings of the first UNESCO International Conference on Information Processing) in 1959. The now-classic diagram consisting of blocks representing list nodes with arrows pointing to successive list nodes appears in "Programming the Logic Theory Machine" by Newell and Shaw in Proc. WJCC, February 1957. Newell and Simon were recognized with the ACM Turing Award in 1975 for having "made basic contributions to artificial intelligence, the psychology of human cognition, and list processing". The problem of machine translation for natural language processing led Victor Yngve at Massachusetts Institute of Technology (MIT) to use linked lists as data structures in his COMIT programming language for computer research in the field of linguistics. A report on this language entitled "A programming language for mechanical translation" appeared in Mechanical Translation in 1958.

LISP, standing for list processor, was created by John McCarthy in 1958 while he was at MIT and in 1960 he published its design in a paper in the Communications of the ACM, entitled "Recursive Functions of Symbolic

---

Expressions and Their Computation by Machine, Part I". One of LISP's major data structures is the linked list. By the early 1960s, the utility of both linked lists and languages which use these structures as their primary data representation was well established. Bert Green of the MIT Lincoln Laboratory published a review article entitled "Computer languages for symbol manipulation" in IRE Transactions on Human Factors in Electronics in March 1961 which summarized the advantages of the linked list approach. A later review article, "A Comparison of list-processing computer languages" by Bobrow and Raphael, appeared in Communications of the ACM in April 1964.

Several operating systems developed by Technical Systems Consultants (originally of West Lafayette Indiana, and later of Chapel Hill, North Carolina) used singly linked lists as file structures. A directory entry pointed to the first sector of a file, and succeeding portions of the file were located by traversing pointers. Systems using this technique included Flex (for the Motorola 6800 CPU), mini-Flex (same CPU), and Flex9 (for the Motorola 6809 CPU). A variant developed by TSC for and marketed by Smoke Signal Broadcasting in California, used doubly linked lists in the same manner.

The TSS/360 operating system, developed by IBM for the System 360/370 machines, used a double linked list for their file system catalog. The directory structure was similar to Unix, where a directory could contain files and/or other directories and extend to any depth. A utility flea was created to fix file system problems after a crash, since modified portions of the file catalog were sometimes in memory when a crash occurred. Problems were detected by comparing the forward and backward links for consistency. If a forward link was corrupt, then if a backward link to the infected node was found, the forward link was set to the node with the backward link. A humorous comment in the source code where this utility was invoked stated "Everyone knows a flea collar gets rid of bugs in cats".

## Basic concepts and nomenclature

Each record of a linked list is often called an **element** or **node**.

The field of each node that contains the address of the next node is usually called the **next link** or **next pointer**. The remaining fields are known as the **data**, **information**, **value**, **cargo**, or **payload** fields.

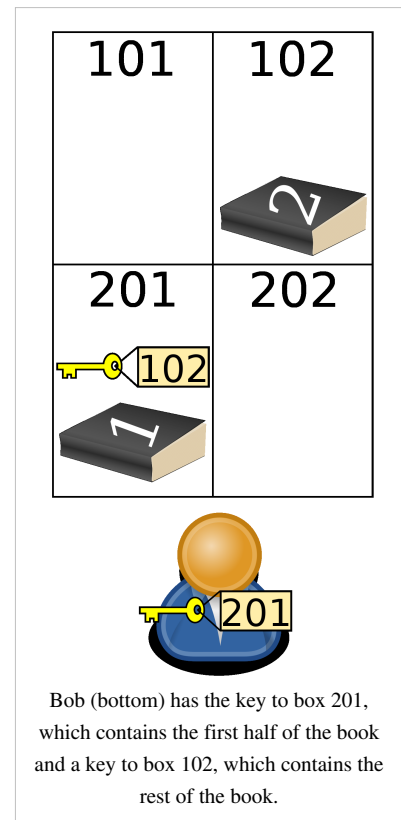
The **head** of a list is its first node. The **tail** of a list may refer either to the rest of the list after the head, or to the last node in the list. In Lisp and some derived languages, the next node may be called the **cdr** (pronounced *could-er*) of the list, while the payload of the head node may be called the **car**.

---

## Post office box analogy

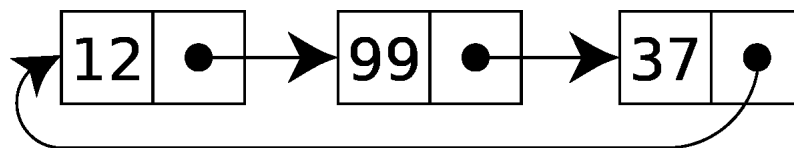
The concept of a linked list can be explained by a simple analogy to real-world post office boxes. Suppose Alice is a spy who wishes to give a codebook to Bob by putting it in a post office box and then giving him the key. However, the book is too thick to fit in a single post office box, so instead she divides the book into two halves and purchases two post office boxes. In the first box, she puts the first half of the book and a key to the second box, and in the second box she puts the second half of the book. She then gives Bob a key to the first box. No matter how large the book is, this scheme can be extended to any number of boxes by always putting the key to the next box in the previous box.

In this analogy, the boxes correspond to *elements* or *nodes*, the keys correspond to *pointers*, and the book itself is the *data*. The key given to Bob is the *head* pointer, while those stored in the boxes are *next* pointers. The scheme as described above is a *singly linked list* (see below).



## Linear and circular lists

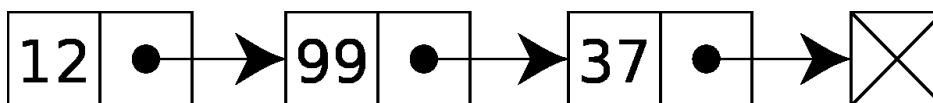
In the last node of a list, the link field often contains a **null** reference, a special value used to indicate the lack of further nodes. A less common convention is to make it point to the first node of the list; in that case the list is said to be **circular** or **circularly linked**; otherwise it is said to be **open** or **linear**.



A circular linked list

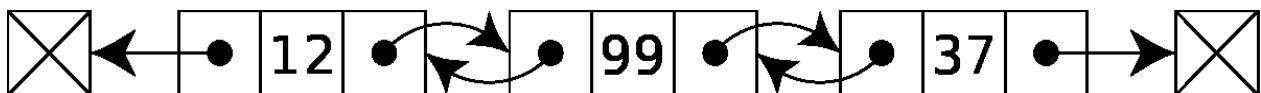
## Singly, doubly, and multiply linked lists

Singly linked lists contain nodes which have a data field as well as a *next* field, which points to the next node in the linked list.



A singly linked list whose nodes contain two fields: an integer value and a link to the next node

In a **doubly linked list**, each node contains, besides the next-node link, a second link field pointing to the *previous* node in the sequence. The two links may be called **forward(s)** and **backwards**, or **next** and **prev(ious)**.



A doubly linked list whose nodes contain three fields: an integer value, the link forward to the next node, and the link backward to the previous node

A technique known as XOR-linking allows a doubly linked list to be implemented using a single link field in each node. However, this technique requires the ability to do bit operations on addresses, and therefore may not be available in some high-level languages.

In a **multiply linked list**, each node contains two or more link fields, each field being used to connect the same set of data records in a different order (e.g., by name, by department, by date of birth, etc.). (While doubly linked lists can be seen as special cases of multiply linked list, the fact that the two orders are opposite to each other leads to simpler and more efficient algorithms, so they are usually treated as a separate case.)

In the case of a circular doubly linked list, the only change that occurs is that end, or "tail", of the said list is linked back to the front, or "head", of the list and vice versa.

### Sentinel nodes

In some implementations, an extra **sentinel** or **dummy** node may be added before the first data record and/or after the last one. This convention simplifies and accelerates some list-handling algorithms, by ensuring that all links can be safely dereferenced and that every list (even one that contains no data elements) always has a "first" and "last" node.

### Empty lists

An *empty list* is a list that contains no data records. This is usually the same as saying that it has zero nodes. If sentinel nodes are being used, the list is usually said to be empty when it has only sentinel nodes.

### Hash linking

The link fields need not be physically part of the nodes. If the data records are stored in an array and referenced by their indices, the link field may be stored in a separate array with the same indices as the data records.

### List handles

Since a reference to the first node gives access to the whole list, that reference is often called the **address**, **pointer**, or **handle** of the list. Algorithms that manipulate linked lists usually get such handles to the input lists and return the handles to the resulting lists. In fact, in the context of such algorithms, the word "list" often means "list handle". In some situations, however, it may be convenient to refer to a list by a handle that consists of two links, pointing to its first and last nodes.

### Combining alternatives

The alternatives listed above may be arbitrarily combined in almost every way, so one may have circular doubly linked lists without sentinels, circular singly linked lists with sentinels, etc.

### Tradeoffs

As with most choices in computer programming and design, no method is well suited to all circumstances. A linked list data structure might work well in one case, but cause problems in another. This is a list of some of the common tradeoffs involving linked list structures.

### Linked lists vs. dynamic arrays

	Linked list	Array	Dynamic array	Balanced tree	Random access list
Indexing	$\Theta(n)$	$\Theta(1)$	$\Theta(1)$	$\Theta(\log n)$	$\Theta(\log n)$
Insert/delete at beginning	$\Theta(1)$	N/A	$\Theta(n)$	$\Theta(\log n)$	$\Theta(1)$
Insert/delete at end	$\Theta(1)$	N/A	$\Theta(1)$ amortized	$\Theta(\log n)$	$\Theta(\log n)$ updating
Insert/delete in middle	search time $\Theta(1)^{[1][2][3]}$	N/A	$\Theta(n)$	$\Theta(\log n)$	$\Theta(\log n)$ updating
Wasted space (average)	$\Theta(n)$	0	$\Theta(n)^{[4]}$	$\Theta(n)$	$\Theta(n)$

A *dynamic array* is a data structure that allocates all elements contiguously in memory, and keeps a count of the current number of elements. If the space reserved for the dynamic array is exceeded, it is reallocated and (possibly) copied, an expensive operation.

Linked lists have several advantages over dynamic arrays. Insertion or deletion of an element at a specific point of a list, assuming that we have a pointer to the node (before the one to be removed, or before the insertion point) already, is a constant-time operation, whereas insertion in a dynamic array at random locations will require moving half of the elements on average, and all the elements in the worst case. While one can "delete" an element from an array in constant time by somehow marking its slot as "vacant", this causes fragmentation that impedes the performance of iteration.

Moreover, arbitrarily many elements may be inserted into a linked list, limited only by the total memory available; while a dynamic array will eventually fill up its underlying array data structure and will have to reallocate — an expensive operation, one that may not even be possible if memory is fragmented, although the cost of reallocation can be averaged over insertions, and the cost of an insertion due to reallocation would still be amortized  $O(1)$ . This helps with appending elements at the array's end, but inserting into (or removing from) middle positions still carries prohibitive costs due to data moving to maintain contiguity. An array from which many elements are removed may also have to be resized in order to avoid wasting too much space.

On the other hand, dynamic arrays (as well as fixed-size array data structures) allow constant-time random access, while linked lists allow only sequential access to elements. Singly linked lists, in fact, can only be traversed in one direction. This makes linked lists unsuitable for applications where it's useful to look up an element by its index quickly, such as heapsort. Sequential access on arrays and dynamic arrays is also faster than on linked lists on many machines, because they have optimal locality of reference and thus make good use of data caching.

Another disadvantage of linked lists is the extra storage needed for references, which often makes them impractical for lists of small data items such as characters or boolean values, because the storage overhead for the links may exceed by a factor of two or more the size of the data. In contrast, a dynamic array requires only the space for the data itself (and a very small amount of control data).<sup>[5]</sup> It can also be slow, and with a naïve allocator, wasteful, to allocate memory separately for each new element, a problem generally solved using memory pools.

Some hybrid solutions try to combine the advantages of the two representations. Unrolled linked lists store several elements in each list node, increasing cache performance while decreasing memory overhead for references. CDR coding does both these as well, by replacing references with the actual data referenced, which extends off the end of the referencing record.

A good example that highlights the pros and cons of using dynamic arrays vs. linked lists is by implementing a program that resolves the Josephus problem. The Josephus problem is an election method that works by having a group of people stand in a circle. Starting at a predetermined person, you count around the circle  $n$  times. Once you reach the  $n$ th person, take them out of the circle and have the members close the circle. Then count around the circle the same  $n$  times and repeat the process, until only one person is left. That person wins the election. This shows the



strengths and weaknesses of a linked list vs. a dynamic array, because if you view the people as connected nodes in a circular linked list then it shows how easily the linked list is able to delete nodes (as it only has to rearrange the links to the different nodes). However, the linked list will be poor at finding the next person to remove and will need to search through the list until it finds that person. A dynamic array, on the other hand, will be poor at deleting nodes (or elements) as it cannot remove one node without individually shifting all the elements up the list by one. However, it is exceptionally easy to find the  $n$ th person in the circle by directly referencing them by their position in the array.

The list ranking problem concerns the efficient conversion of a linked list representation into an array. Although trivial for a conventional computer, solving this problem by a parallel algorithm is complicated and has been the subject of much research.

A balanced tree has similar memory access patterns and space overhead to a linked list while permitting much more efficient indexing, taking  $O(\log n)$  time instead of  $O(n)$  for a random access. However, insertion and deletion operations are more expensive due to the overhead of tree manipulations to maintain balance. Efficient schemes exist for trees to automatically maintain themselves in almost-balanced state, like AVL trees or red-black trees.

### Singly linked linear lists vs. other lists

While doubly linked and/or circular lists have advantages over singly linked linear lists, linear lists offer some advantages that make them preferable in some situations.

For one thing, a singly linked linear list is a recursive data structure, because it contains a pointer to a *smaller* object of the same type. For that reason, many operations on singly linked linear lists (such as merging two lists, or enumerating the elements in reverse order) often have very simple recursive algorithms, much simpler than any solution using iterative commands. While one can adapt those recursive solutions for doubly linked and circularly linked lists, the procedures generally need extra arguments and more complicated base cases.

Linear singly linked lists also allow tail-sharing, the use of a common final portion of sub-list as the terminal portion of two different lists. In particular, if a new node is added at the beginning of a list, the former list remains available as the tail of the new one — a simple example of a persistent data structure. Again, this is not true with the other variants: a node may never belong to two different circular or doubly linked lists.

In particular, end-sentinel nodes can be shared among singly linked non-circular lists. One may even use the same end-sentinel node for *every* such list. In Lisp, for example, every proper list ends with a link to a special node, denoted by `nil` or `()`, whose `CAR` and `CDR` links point to itself. Thus a Lisp procedure can safely take the `CAR` or `CDR` of *any* list.

Indeed, the advantages of the fancy variants are often limited to the complexity of the algorithms, not in their efficiency. A circular list, in particular, can usually be emulated by a linear list together with two variables that point to the first and last nodes, at no extra cost.

### Doubly linked vs. singly linked

Double-linked lists require more space per node (unless one uses XOR-linking), and their elementary operations are more expensive; but they are often easier to manipulate because they allow sequential access to the list in both directions. In a doubly linked list, one can insert or delete a node in a constant number of operations given only that node's address. To do the same in a singly linked list, one must have the *address of the pointer* to that node, which is either the handle for the whole list (in case of the first node) or the link field in the *previous* node. Some algorithms require access in both directions. On the other hand, doubly linked lists do not allow tail-sharing and cannot be used as persistent data structures.

---

## Circularly linked vs. linearly linked

A circularly linked list may be a natural option to represent arrays that are naturally circular, e.g. the corners of a polygon, a pool of buffers that are used and released in FIFO order, or a set of processes that should be time-shared in round-robin order. In these applications, a pointer to any node serves as a handle to the whole list.

With a circular list, a pointer to the last node gives easy access also to the first node, by following one link. Thus, in applications that require access to both ends of the list (e.g., in the implementation of a queue), a circular structure allows one to handle the structure by a single pointer, instead of two.

A circular list can be split into two circular lists, in constant time, by giving the addresses of the last node of each piece. The operation consists in swapping the contents of the link fields of those two nodes. Applying the same operation to any two nodes in two distinct lists joins the two list into one. This property greatly simplifies some algorithms and data structures, such as the quad-edge and face-edge.

The simplest representation for an empty *circular* list (when such a thing makes sense) is a null pointer, indicating that the list has no nodes. Without this choice, many algorithms have to test for this special case, and handle it separately. By contrast, the use of null to denote an empty *linear* list is more natural and often creates fewer special cases.

## Using sentinel nodes

Sentinel node may simplify certain list operations, by ensuring that the next and/or previous nodes exist for every element, and that even empty lists have at least one node. One may also use a sentinel node at the end of the list, with an appropriate data field, to eliminate some end-of-list tests. For example, when scanning the list looking for a node with a given value  $x$ , setting the sentinel's data field to  $x$  makes it unnecessary to test for end-of-list inside the loop. Another example is the merging two sorted lists: if their sentinels have data fields set to  $+\infty$ , the choice of the next output node does not need special handling for empty lists.

However, sentinel nodes use up extra space (especially in applications that use many short lists), and they may complicate other operations (such as the creation of a new empty list).

However, if the circular list is used merely to simulate a linear list, one may avoid some of this complexity by adding a single sentinel node to every list, between the last and the first data nodes. With this convention, an empty list consists of the sentinel node alone, pointing to itself via the next-node link. The list handle should then be a pointer to the last data node, before the sentinel, if the list is not empty; or to the sentinel itself, if the list is empty.

The same trick can be used to simplify the handling of a doubly linked linear list, by turning it into a circular doubly linked list with a single sentinel node. However, in this case, the handle should be a single pointer to the dummy node itself.<sup>[6]</sup>

## Linked list operations

When manipulating linked lists in-place, care must be taken to not use values that you have invalidated in previous assignments. This makes algorithms for inserting or deleting linked list nodes somewhat subtle. This section gives pseudocode for adding or removing nodes from singly, doubly, and circularly linked lists in-place. Throughout we will use *null* to refer to an end-of-list marker or sentinel, which may be implemented in a number of ways.

## Linearly linked lists

### Singly linked lists

Our node data structure will have two fields. We also keep a variable *firstNode* which always points to the first node in the list, or is *null* for an empty list.

---

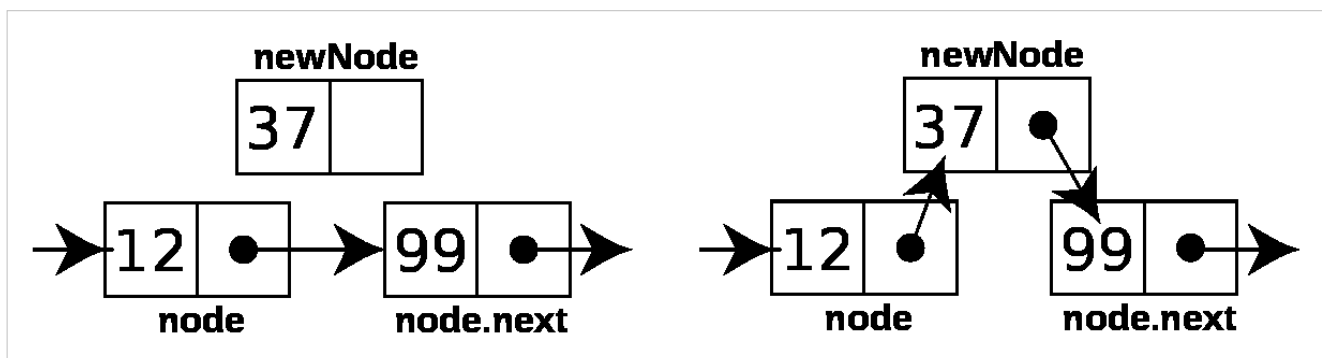
```
record Node {
    data; // The data being stored in the node
    Node next // A reference to the next node, null for last node
}
```

```
record List {
    Node firstNode // points to first node of list; null for empty list
}
```

Traversal of a singly linked list is simple, beginning at the first node and following each *next* link until we come to the end:

```
node := list.firstNode
while node not null
    (do something with node.data)
    node := node.next
```

The following code inserts a node after an existing node in a singly linked list. The diagram shows how it works. Inserting a node before an existing one cannot be done directly; instead, one must keep track of the previous node and insert a node after it.

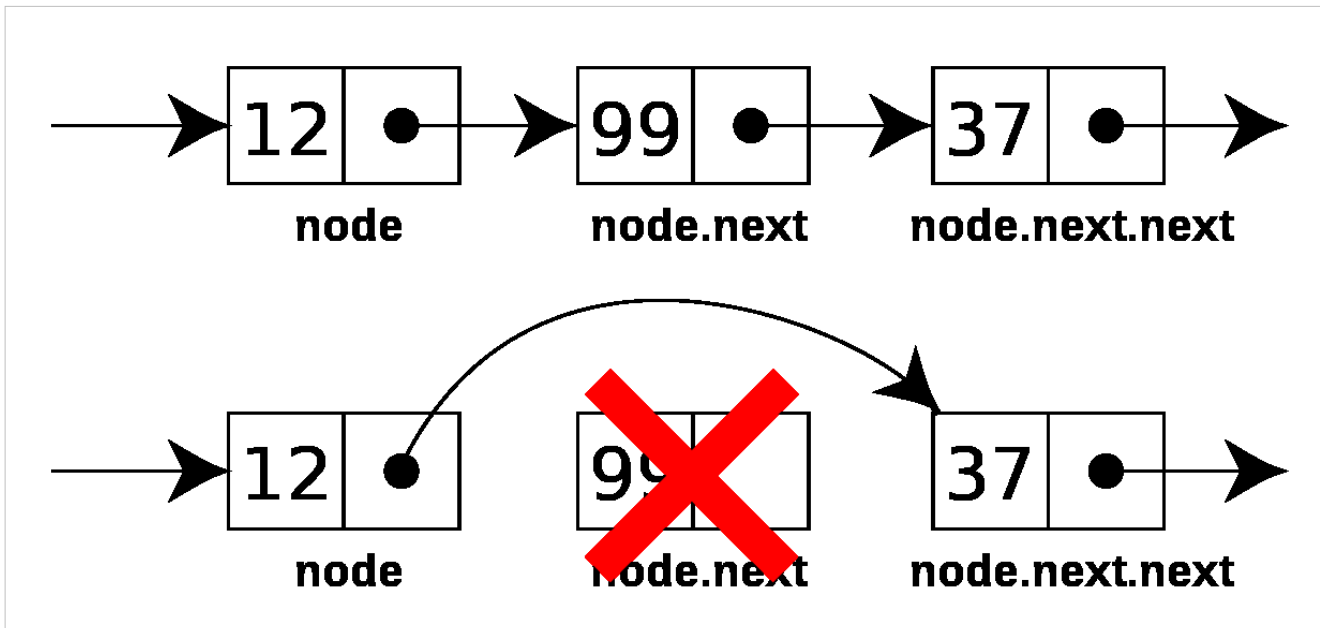


```
function insertAfter(Node node, Node newNode) // insert newNode after node
    newNode.next := node.next
    node.next    := newNode
```

Inserting at the beginning of the list requires a separate function. This requires updating *firstNode*.

```
function insertBeginning(List list, Node newNode) // insert node before current first node
    newNode.next := list.firstNode
    list.firstNode := newNode
```

Similarly, we have functions for removing the node *after* a given node, and for removing a node from the beginning of the list. The diagram demonstrates the former. To find and remove a particular node, one must again keep track of the previous element.



```
function removeAfter(node node) // remove node past this one
  obsoleteNode := node.next
  node.next := node.next.next
  destroy obsoleteNode
```

```
function removeBeginning(List list) // remove first node
  obsoleteNode := list.firstNode
  list.firstNode := list.firstNode.next // point past deleted node
  destroy obsoleteNode
```

Notice that `removeBeginning()` sets `list.firstNode` to null when removing the last node in the list.

Since we can't iterate backwards, efficient `insertBefore` or `removeBefore` operations are not possible.

Appending one linked list to another can be inefficient unless a reference to the tail is kept as part of the `List` structure, because we must traverse the entire first list in order to find the tail, and then append the second list to this. Thus, if two linearly linked lists are each of length  $n$ , list appending has asymptotic time complexity of  $O(n)$ . In the Lisp family of languages, list appending is provided by the `append` procedure.

Many of the special cases of linked list operations can be eliminated by including a dummy element at the front of the list. This ensures that there are no special cases for the beginning of the list and renders both `insertBeginning()` and `removeBeginning()` unnecessary. In this case, the first useful data in the list will be found at `list.firstNode.next`.

## Circularly linked list

In a circularly linked list, all nodes are linked in a continuous circle, without using *null*. For lists with a front and a back (such as a queue), one stores a reference to the last node in the list. The *next* node after the last node is the first node. Elements can be added to the back of the list and removed from the front in constant time.

Circularly linked lists can be either singly or doubly linked.

Both types of circularly linked lists benefit from the ability to traverse the full list beginning at any given node. This often allows us to avoid storing *firstNode* and *lastNode*, although if the list may be empty we need a special representation for the empty list, such as a *lastNode* variable which points to some node in the list or is *null* if it's empty; we use such a *lastNode* here. This representation significantly simplifies adding and removing nodes with a non-empty list, but empty lists are then a special case.

### Algorithms

Assuming that *someNode* is some node in a non-empty circular singly linked list, this code iterates through that list starting with *someNode*:

```
function iterate(someNode)
  if someNode ≠ null
    node := someNode
    do
      do something with node.value
      node := node.next
    while node ≠ someNode
```

Notice that the test "**while** node ≠ someNode" must be at the end of the loop. If the test was moved to the beginning of the loop, the procedure would fail whenever the list had only one node.

This function inserts a node "newNode" into a circular linked list after a given node "node". If "node" is null, it assumes that the list is empty.

```
function insertAfter(Node node, Node newNode)
  if node = null
    newNode.next := newNode
  else
    newNode.next := node.next
    node.next := newNode
```

Suppose that "L" is a variable pointing to the last node of a circular linked list (or null if the list is empty). To append "newNode" to the *end* of the list, one may do

```
insertAfter(L, newNode)
L := newNode
```

To insert "newNode" at the *beginning* of the list, one may do

```
insertAfter(L, newNode)
if L = null
  L := newNode
```

## Linked lists using arrays of nodes

Languages that do not support any type of reference can still create links by replacing pointers with array indices. The approach is to keep an array of records, where each record has integer fields indicating the index of the next (and possibly previous) node in the array. Not all nodes in the array need be used. If records are also not supported, parallel arrays can often be used instead.

As an example, consider the following linked list record that uses arrays instead of pointers:

```
record Entry {
    integer next; // index of next entry in array
    integer prev; // previous entry (if double-linked)
    string name;
    real balance;
}
```

By creating an array of these structures, and an integer variable to store the index of the first element, a linked list can be built:

```
integer listHead
Entry Records[1000]
```

Links between elements are formed by placing the array index of the next (or previous) cell into the Next or Prev field within a given element. For example:

Index	Next	Prev	Name	Balance
0	1	4	Jones, John	123.45
1	-1	0	Smith, Joseph	234.56
2 (listHead)	4	-1	Adams, Adam	0.00
3			Ignore, Ignatius	999.99
4	0	2	Another, Anita	876.54
5				
6				
7				

In the above example, `ListHead` would be set to 2, the location of the first entry in the list. Notice that entry 3 and 5 through 7 are not part of the list. These cells are available for any additions to the list. By creating a `ListFree` integer variable, a free list could be created to keep track of what cells are available. If all entries are in use, the size of the array would have to be increased or some elements would have to be deleted before new entries could be stored in the list.

The following code would traverse the list and display names and account balance:

```
i := listHead
while i ≥ 0 // loop through the list
    print i, Records[i].name, Records[i].balance // print entry
    i := Records[i].next
```

When faced with a choice, the advantages of this approach include:

- The linked list is relocatable, meaning it can be moved about in memory at will, and it can also be quickly and directly serialized for storage on disk or transfer over a network.

- Especially for a small list, array indexes can occupy significantly less space than a full pointer on many architectures.
- Locality of reference can be improved by keeping the nodes together in memory and by periodically rearranging them, although this can also be done in a general store.
- Naïve dynamic memory allocators can produce an excessive amount of overhead storage for each node allocated; almost no allocation overhead is incurred per node in this approach.
- Seizing an entry from a pre-allocated array is faster than using dynamic memory allocation for each node, since dynamic memory allocation typically requires a search for a free memory block of the desired size.

This approach has one main disadvantage, however: it creates and manages a private memory space for its nodes. This leads to the following issues:

- It increase complexity of the implementation.
- Growing a large array when it is full may be difficult or impossible, whereas finding space for a new linked list node in a large, general memory pool may be easier.
- Adding elements to a dynamic array will occasionally (when it is full) unexpectedly take linear ( $O(n)$ ) instead of constant time (although it's still an amortized constant).
- Using a general memory pool leaves more memory for other data if the list is smaller than expected or if many nodes are freed.

For these reasons, this approach is mainly used for languages that do not support dynamic memory allocation. These disadvantages are also mitigated if the maximum size of the list is known at the time the array is created.

## Language support

Many programming languages such as Lisp and Scheme have singly linked lists built in. In many functional languages, these lists are constructed from nodes, each called a *cons* or *cons cell*. The cons has two fields: the *car*, a reference to the data for that node, and the *cdr*, a reference to the next node. Although cons cells can be used to build other data structures, this is their primary purpose.

In languages that support abstract data types or templates, linked list ADTs or templates are available for building linked lists. In other languages, linked lists are typically built using references together with records.

## Internal and external storage

When constructing a linked list, one is faced with the choice of whether to store the data of the list directly in the linked list nodes, called *internal storage*, or merely to store a reference to the data, called *external storage*. Internal storage has the advantage of making access to the data more efficient, requiring less storage overall, having better locality of reference, and simplifying memory management for the list (its data is allocated and deallocated at the same time as the list nodes).

External storage, on the other hand, has the advantage of being more generic, in that the same data structure and machine code can be used for a linked list no matter what the size of the data is. It also makes it easy to place the same data in multiple linked lists. Although with internal storage the same data can be placed in multiple lists by including multiple *next* references in the node data structure, it would then be necessary to create separate routines to add or delete cells based on each field. It is possible to create additional linked lists of elements that use internal storage by using external storage, and having the cells of the additional linked lists store references to the nodes of the linked list containing the data.

In general, if a set of data structures needs to be included in multiple linked lists, external storage is the best approach. If a set of data structures need to be included in only one linked list, then internal storage is slightly better, unless a generic linked list package using external storage is available. Likewise, if different sets of data that can be stored in the same data structure are to be included in a single linked list, then internal storage would be fine.

---

Another approach that can be used with some languages involves having different data structures, but all have the initial fields, including the *next* (and *prev* if double linked list) references in the same location. After defining separate structures for each type of data, a generic structure can be defined that contains the minimum amount of data shared by all the other structures and contained at the top (beginning) of the structures. Then generic routines can be created that use the minimal structure to perform linked list type operations, but separate routines can then handle the specific data. This approach is often used in message parsing routines, where several types of messages are received, but all start with the same set of fields, usually including a field for message type. The generic routines are used to add new messages to a queue when they are received, and remove them from the queue in order to process the message. The message type field is then used to call the correct routine to process the specific type of message.

### Example of internal and external storage

Suppose you wanted to create a linked list of families and their members. Using internal storage, the structure might look like the following:

```
record member { // member of a family
    member next;
    string firstName;
    integer age;
}
record family { // the family itself
    family next;
    string lastName;
    string address;
    member members // head of list of members of this family
}
```

To print a complete list of families and their members using internal storage, we could write:

```
aFamily := Families // start at head of families list
while aFamily ≠ null // loop through list of families
    print information about family
    aMember := aFamily.members // get head of list of this family's members
    while aMember ≠ null // loop through list of members
        print information about member
        aMember := aMember.next

aFamily := aFamily.next
```

Using external storage, we would create the following structures:

```
record node { // generic link structure
    node next;
    pointer data // generic pointer for data at node
}
record member { // structure for family member
    string firstName;
    integer age
}
record family { // structure for family
```



```

    string lastName;
    string address;
    node members // head of list of members of this family
}

```

To print a complete list of families and their members using external storage, we could write:

```

famNode := Families // start at head of families list
while famNode ≠ null // loop through list of families
    aFamily := (family) famNode.data // extract family from node
    print information about family
    memNode := aFamily.members // get list of family members
    while memNode ≠ null // loop through list of members
        aMember := (member) memNode.data // extract member from node
        print information about member
        memNode := memNode.next

famNode := famNode.next

```

Notice that when using external storage, an extra step is needed to extract the record from the node and cast it into the proper data type. This is because both the list of families and the list of members within the family are stored in two linked lists using the same data structure (*node*), and this language does not have parametric types.

As long as the number of families that a member can belong to is known at compile time, internal storage works fine. If, however, a member needed to be included in an arbitrary number of families, with the specific number known only at run time, external storage would be necessary.

## Speeding up search

Finding a specific element in a linked list, even if it is sorted, normally requires  $O(n)$  time (linear search). This is one of the primary disadvantages of linked lists over other data structures. In addition to the variants discussed above, below are two simple ways to improve search time.

In an unordered list, one simple heuristic for decreasing average search time is the *move-to-front heuristic*, which simply moves an element to the beginning of the list once it is found. This scheme, handy for creating simple caches, ensures that the most recently used items are also the quickest to find again.

Another common approach is to "index" a linked list using a more efficient external data structure. For example, one can build a red-black tree or hash table whose elements are references to the linked list nodes. Multiple such indexes can be built on a single list. The disadvantage is that these indexes may need to be updated each time a node is added or removed (or at least, before that index is used again).

## Random access lists

A random access list is a list with support for fast random access to read or modify any element in the list.<sup>[7]</sup> One possible implementation is a skew binary random access list using the skew binary number system, which involves a list of trees with special properties; this allows worst-case constant time head/cons operations, and worst-case logarithmic time random access to an element by index).<sup>[7]</sup> Random access lists can be implemented as persistent data structures.<sup>[7]</sup>

Random access lists can be viewed as immutable linked lists in that they likewise support the same  $O(1)$  head and tail operations.<sup>[7]</sup>

A simple extension to random access lists is the min-list, which provides an additional operation that yields the minimum element in the entire list in constant time (without mutation complexities).<sup>[7]</sup>

## Related data structures

Both stacks and queues are often implemented using linked lists, and simply restrict the type of operations which are supported.

The skip list is a linked list augmented with layers of pointers for quickly jumping over large numbers of elements, and then descending to the next layer. This process continues down to the bottom layer, which is the actual list.

A binary tree can be seen as a type of linked list where the elements are themselves linked lists of the same nature. The result is that each node may include a reference to the first node of one or two other linked lists, which, together with their contents, form the subtrees below that node.

An unrolled linked list is a linked list in which each node contains an array of data values. This leads to improved cache performance, since more list elements are contiguous in memory, and reduced memory overhead, because less metadata needs to be stored for each element of the list.

A hash table may use linked lists to store the chains of items that hash to the same position in the hash table.

A heap shares some of the ordering properties of a linked list, but is almost always implemented using an array. Instead of references from node to node, the next and previous data indexes are calculated using the current data's index.

A self-organizing list rearranges its nodes based on some heuristic which reduces search times for data retrieval by keeping commonly accessed nodes at the head of the list.

## Notes

- [1] Gerald Kruse. CS 240 Lecture Notes (<http://www.juniata.edu/faculty/kruse/cs240/syllabus.htm>): Linked Lists Plus: Complexity Trade-offs (<http://www.juniata.edu/faculty/kruse/cs240/linkedlist2.htm>). Juniata College. Spring 2008.
- [2] *Day 1 Keynote - Bjarne Stroustrup: C++11 Style* (<http://channel9.msdn.com/Events/GoingNative/GoingNative-2012/Keynote-Bjarne-Stroustrup-Cpp11-Style>) at *GoingNative 2012* on *channel9.msdn.com* from minute 45 or foil 44
- [3] *Number crunching: Why you should never, ever, EVER use linked-list in your code again* (<http://kjellkod.wordpress.com/2012/02/25/why-you-should-never-ever-ever-use-linked-list-in-your-code-again/>) at *kjellkod.wordpress.com*
- [4] Brodnik, Andrej; Carlsson, Svante; Sedgewick, Robert; Munro, JI; Demaine, ED (Technical Report CS-99-09), *Resizable Arrays in Optimal Time and Space* (<http://www.cs.uwaterloo.ca/research/tr/1999/09/CS-99-09.pdf>), Department of Computer Science, University of Waterloo,
- [5] The amount of control data required for a dynamic array is usually of the form  $K + B * n$ , where  $K$  is a per-array constant,  $B$  is a per-dimension constant, and  $n$  is the number of dimensions.  $K$  and  $B$  are typically on the order of 10 bytes.
- [6] Ford, William and Topp, William *Data Structures with C++ using STL Second Edition* (2002). Prentice-Hall. ISBN 0-13-085850-1, pp. 466-467
- [7] C Okasaki, "Purely Functional Random-Access Lists (<http://cs.oberlin.edu/~jwalker/refs/fpca95.ps>)"

## Footnotes

## References

- Juan, Angel (2006). "Ch20 –Data Structures; ID06 - PROGRAMMING with JAVA (slide part of the book "Big Java", by CayS. Horstmann)" (<http://www.uoc.edu/in3/emath/docs/java/ch20.pdf>) (PDF). p. 3
- "Definition of a linked list" (<http://nist.gov/dads/HTML/LinkedList.html>). National Institute of Standards and Technology. 2004-08-16. Retrieved 2004-12-14.
- Antonakos, James L.; Mansfield, Kenneth C., Jr. (1999). *Practical Data Structures Using C/C++*. Prentice-Hall. pp. 165–190. ISBN 0-13-280843-9.
- Collins, William J. (2005) [2002]. *Data Structures and the Java Collections Framework*. New York: McGraw Hill. pp. 239–303. ISBN 0-07-282379-8.
- Cormen, Thomas H.; Charles E. Leiserson; Ronald L. Rivest; Clifford Stein (2003). *Introduction to Algorithms*. MIT Press. pp. 205–213 & 501–505. ISBN 0-262-03293-7.
- Cormen, Thomas H.; Charles E. Leiserson; Ronald L. Rivest; Clifford Stein (2001). "10.2: Linked lists". *Introduction to Algorithms* (2nd ed.). MIT Press. pp. 204–209. ISBN 0-262-03293-7.
- Green, Bert F. Jr. (1961). "Computer Languages for Symbol Manipulation". *IRE Transactions on Human Factors in Electronics* (2): 3–8.
- McCarthy, John (1960). "Recursive Functions of Symbolic Expressions and Their Computation by Machine, Part I" (<http://www-formal.stanford.edu/jmc/recursive.html>). *Communications of the ACM*.
- Knuth, Donald (1997). "2.2.3-2.2.5". *Fundamental Algorithms* (3rd ed.). Addison-Wesley. pp. 254–298. ISBN 0-201-89683-4.
- Newell, Allen; Shaw, F. C. (1957). "Programming the Logic Theory Machine". *Proceedings of the Western Joint Computer Conference*: 230–240.
- Parlante, Nick (2001). "Linked list basics" (<http://cslibrary.stanford.edu/103/LinkedListBasics.pdf>). Stanford University. Retrieved 2009-09-21.
- Sedgewick, Robert (1998). *Algorithms in C*. Addison Wesley. pp. 90–109. ISBN 0-201-31452-5.
- Shaffer, Clifford A. (1998). *A Practical Introduction to Data Structures and Algorithm Analysis*. New Jersey: Prentice Hall. pp. 77–102. ISBN 0-13-660911-2.
- Wilkes, Maurice Vincent (1964). "An Experiment with a Self-compiling Compiler for a Simple List-Processing Language". *Annual Review in Automatic Programming* (Pergamon Press) **4** (1).
- Wilkes, Maurice Vincent (1964). "Lists and Why They are Useful". *Proceeds of the ACM National Conference, Philadelphia 1964* (ACM) (P-64): F1–1.
- Shanmugasundaram, Kulesh (2005-04-04). "Linux Kernel Linked List Explained" (<http://isis.poly.edu/kulesh/stuff/src/klist/>). Retrieved 2009-09-21.

## External links

- Description (<http://nist.gov/dads/HTML/LinkedList.html>) from the Dictionary of Algorithms and Data Structures
- Introduction to Linked Lists (<http://cslibrary.stanford.edu/103/>), Stanford University Computer Science Library
- Linked List Problems (<http://cslibrary.stanford.edu/105/>), Stanford University Computer Science Library
- Open Data Structures - Chapter 3 - Linked Lists ([http://opendatastructures.org/versions/edition-0.1e/ods-java/3\\_Linked\\_Lists.html](http://opendatastructures.org/versions/edition-0.1e/ods-java/3_Linked_Lists.html))
- Patent for the idea of having nodes which are in several linked lists simultaneously (<http://www.google.com/patents?vid=USPAT7028023>) (note that this technique was widely used for many decades before the patent was granted)

# XOR linked list

An **XOR linked list** is a data structure used in computer programming. It takes advantage of the bitwise exclusive disjunction (XOR) operation, here denoted by  $\oplus$ , to decrease storage requirements for doubly linked lists. An ordinary doubly linked list stores addresses of the previous and next list items in each list node, requiring two address fields:

```
...  A      B      C      D      E  ...
      -> next -> next -> next ->
      <- prev <- prev <- prev <-
```

An XOR linked list compresses the same information into *one* address field by storing the bitwise XOR of the address for *previous* and the address for *next* in one field:

```
...  A      B      C      D      E  ...
      <-> A⊕C <-> B⊕D <-> C⊕E <->
```

When you traverse the list from left to right: supposing you are at C, you can take the address of the previous item, B, and XOR it with the value in the link field ( $B \oplus D$ ). You will then have the address for D and you can continue traversing the list. The same pattern applies in the other direction.

To start traversing the list in either direction from some point, you need the address of two consecutive items, not just one. If the addresses of the two consecutive items are reversed, you will end up traversing the list in the opposite direction.

This form of linked list may be inadvisable:

- General-purpose debugging tools cannot follow the XOR chain, making debugging more difficult; <sup>[1]</sup>
- The price for the decrease in memory usage is an increase in code complexity, making maintenance more expensive;
- Most garbage collection schemes do not work with data structures that do not contain literal pointers;
- XOR of pointers is not defined in some contexts (e.g., the C language), although many languages provide some kind of type conversion between pointers and integers;
- The pointers will be unreadable if one isn't traversing the list — for example, if the pointer to a list item was contained in another data structure;
- While traversing the list you need to remember the address of the previously accessed node in order to calculate the next node's address.

Computer systems have increasingly cheap and plentiful memory, and storage overhead is not generally an overriding issue outside specialized embedded systems. Where it is still desirable to reduce the overhead of a linked list, unrolling provides a more practical approach (as well as other advantages, such as increasing cache performance and speeding random access).

## Features

- Given only one list item, one cannot immediately obtain the addresses of the other elements of the list.
- Two XOR operations suffice to do the traversal from one item to the next, the same instructions sufficing in both cases. Consider a list with items  $\{...B \ C \ D...\}$  and with R1 and R2 being registers containing, respectively, the address of the current (say C) list item and a work register containing the XOR of the current address with the previous address (say  $C \oplus D$ ). Cast as System/360 instructions:

```
X  R2,Link      R2 <- C⊕D ⊕ B⊕D (i.e. B⊕C, "Link" being the link field
                                in the current record, containing B⊕D)
```

```
XR R1,R2      R1 <- C ⊕ B⊕C      (i.e. B, voilà: the next record)
```

- End of list is signified by imagining a list item at address zero placed adjacent to an end point, as in {0 A B C...}. The link field at A would be 0⊕B. An additional instruction is needed in the above sequence after the two XOR operations to detect a zero result in developing the address of the current item,
- A list end point can be made reflective by making the link pointer be zero. A zero pointer is a *mirror*. (The XOR of the left and right neighbor addresses, being the same, is zero.)

## Why does it work?

The key is the first operation, and the properties of XOR:

- $X \oplus X = 0$
- $X \oplus 0 = X$
- $X \oplus Y = Y \oplus X$
- $(X \oplus Y) \oplus Z = X \oplus (Y \oplus Z)$

The R2 register always contains the XOR of the address of current item C with the address of the predecessor item P:  $C \oplus P$ . The Link fields in the records contain the XOR of the left and right successor addresses, say  $L \oplus R$ . XOR of R2 ( $C \oplus P$ ) with the current link field ( $L \oplus R$ ) yields  $C \oplus P \oplus L \oplus R$ .

- If the predecessor was L, the  $P(=L)$  and L *cancel out* leaving  $C \oplus R$ .
- If the predecessor had been R, the  $P(=R)$  and R cancel, leaving  $C \oplus L$ .

In each case, the result is the XOR of the current address with the next address. XOR of this with the current address in R1 leaves the next address. R2 is left with the requisite XOR pair of the (now) current address and the predecessor.

## Variations

The underlying principle of the XOR linked list can be applied to any reversible binary operation. Replacing XOR by addition or subtraction gives slightly different, but largely equivalent, formulations:

### Addition linked list

...	A		B		C		D		E	...
		<->	A+C	<->	B+D	<->	C+E	<->		

This kind of list has exactly the same properties as the XOR linked list, except that a zero link field is not a "mirror". The address of the next node in the list is given by subtracting the previous node's address from the current node's link field.

### Subtraction linked list

...	A		B		C		D		E	...
		<->	C-A	<->	D-B	<->	E-C	<->		

This kind of list differs from the "traditional" XOR linked list in that the instruction sequences needed to traverse the list forwards is different from the sequence needed to traverse the list in reverse. The address of the next node, going forwards, is given by *adding* the link field to the previous node's address; the address of the preceding node is given by *subtracting* the link field from the next node's address.

The subtraction linked list is also special in that the entire list can be relocated in memory without needing any patching of pointer values, since adding a constant offset to each address in the list will not require any changes to the values stored in the link fields. (See also *Serialization*.) This is an advantage over both XOR linked lists and

traditional linked lists.

*Note about implementations in C:*

The subtraction linked list also does not require casting C pointers to integers, provided the whole list structure is inside a single contiguous block of memory. In that case the subtraction of two C pointers yields an integer. Note that on most platforms the maximum size of a contiguous block of memory will be considerably smaller than the total available memory, so large lists will typically not fit into a single contiguous block of memory. This is not a problem as long as the platform provides the C99 type `uintptr_t`, because then pointers can be portably cast to `uintptr_t` and back again.

## References

[1] <http://www.iecc.com/gclist/GC-faq.html#GC,%20C,%20and%20C++>

## External links

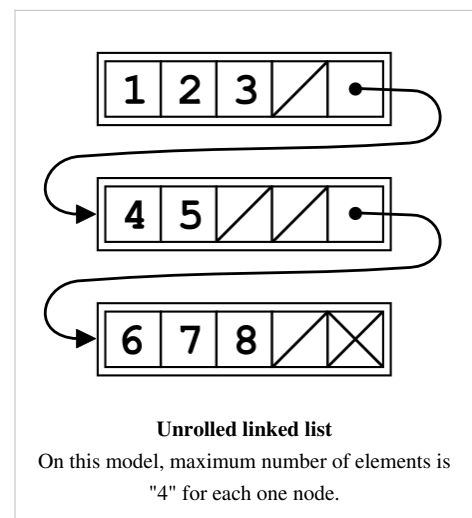
- Example implementation in C++. (<http://blog.wsensors.com/?p=177>)

# Unrolled linked list

In computer programming, an **unrolled linked list** is a variation on the linked list which stores multiple elements in each node. It can dramatically increase cache performance, while decreasing the memory overhead associated with storing list metadata such as references. It is related to the B-tree.

## Overview

A typical unrolled linked list node looks like this:



```
record node {
    node next          // reference to next node in list
    int numElements    // number of elements in this node, up to maxElements
    array elements     // an array of numElements elements,
                      // with space allocated for maxElements elements
}
```

Each node holds up to a certain maximum number of elements, typically just large enough so that the node fills a single cache line or a small multiple thereof. A position in the list is indicated by both a reference to the node and a position in the elements array. It is also possible to include a *previous* pointer for an unrolled doubly linked list.

To insert a new element, we simply find the node the element should be in and insert the element into the `elements` array, incrementing `numElements`. If the array is already full, we first insert a new node either preceding or following the current one and move half of the elements in the current node into it.

To remove an element, similarly, we simply find the node it is in and delete it from the `elements` array, decrementing `numElements`. If `numElements` falls below `maxElements ÷ 2` then we pull elements from adjacent nodes to fill it back up to this level. If both adjacent nodes are too low, we combine it with one adjacent node and then move some values into the other. This is necessary to avoid wasting space.

## Performance

One of the primary benefits of unrolled linked lists is decreased storage requirements. All nodes (except at most one) are at least half-full. If many random inserts and deletes are done, the average node will be about three-quarters full, and if inserts and deletes are only done at the beginning and end, almost all nodes will be full. Assume that:

- $m = \text{maxElements}$ , the maximum number of elements in each `elements` array;
- $v$  = the overhead per node for references and element counts;
- $s$  = the size of a single element.

Then, the space used for  $n$  elements varies between  $(v/m + s)n$  and  $(2v/m + s)n$ . For comparison, ordinary linked lists require  $(v + s)n$  space, although  $v$  may be smaller, and arrays, one of the most compact data structures, require  $sn$  space. Unrolled linked lists effectively spread the overhead  $v$  over a number of elements of the list. Thus, we see the most significant space gain when overhead is large, `maxElements` is large, or elements are small.

If the elements are particularly small, such as bits, the overhead can be as much as 64 times larger than the data on many machines. Moreover, many popular memory allocators will keep a small amount of metadata for each node allocated, increasing the effective overhead  $v$ . Both these make unrolled linked lists more attractive.

Because unrolled linked list nodes each store a count next to the `next` field, retrieving the  $k$ th element of an unrolled linked list (indexing) can be done in  $n/m + 1$  cache misses, up to a factor of  $m$  better than ordinary linked lists. Additionally, if the size of each element is small compared to the cache line size, the list can be iterated over in order with fewer cache misses than ordinary linked lists. In either case, operation time still increases linearly with the size of the list.

## External links

- Implementation written in C++<sup>[1]</sup>
- Another implementation written in Java<sup>[2]</sup>

## References

[1] [http://en.literateprograms.org/Unrolled\\_linked\\_list\\_%28C\\_Plus\\_Plus%29](http://en.literateprograms.org/Unrolled_linked_list_%28C_Plus_Plus%29)

[2] <https://github.com/megatherion/Unrolled-linked-list>

- Shao, Z.; Reppy, J. H.; Appel, A.W. (1994), "Unrolling lists" ([http://dl.acm.org/ft\\_gateway.cfm?id=182453&type=pdf&CFID=80599027&CFTOKEN=68993242](http://dl.acm.org/ft_gateway.cfm?id=182453&type=pdf&CFID=80599027&CFTOKEN=68993242)), *Conference record of the 1994 ACM Conference on Lisp and Functional Programming*: 185-191

# VList

In computer science, the **VList** is a persistent data structure designed by Phil Bagwell in 2002 that combines the fast indexing of arrays with the easy extension of cons-based (or singly linked) linked lists.<sup>[1]</sup>

Like arrays, VLists have constant-time lookup on average and are highly compact, requiring only  $O(\log n)$  storage for pointers, allowing them to take advantage of locality of reference. Like singly linked or cons-based lists, they are persistent, and elements can be added to or removed from the front in constant time. Length can also be found in  $O(\log n)$  time.

The primary operations of a VList are:

- Locate the  $k$ th element ( $O(1)$  average,  $O(\log n)$  worst-case)
- Add an element to the front of the VList ( $O(1)$  average, with an occasional allocation)
- Obtain a new array beginning at the second element of an old array ( $O(1)$ )
- Compute the length of the list ( $O(\log n)$ )

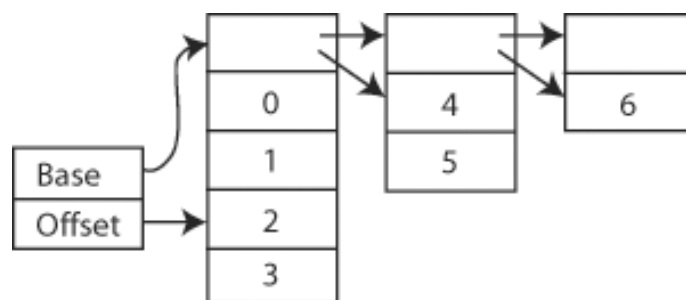
The primary advantage VLists have over arrays is that different updated versions of the VList automatically share structure. Because VLists are immutable, they are most useful in functional programming languages, where their efficiency allows a purely functional implementation of data structures traditionally thought to require mutable arrays, such as hash tables.

However, VLists also have a number of disadvantages over their competitors:

- While immutability is a benefit, it is also a drawback, making it inefficient to modify elements in the middle of the array.
- Access near the end of the list can be as expensive as  $O(\log n)$ ; it is only constant on average over all elements. This is still, however, much better than performing the same operation on cons-based lists.
- Wasted space in the first block is proportional to  $n$ . This is similar to linked lists, but there are data structures with less overhead. When used as a fully persistent data structure, the overhead may be considerably higher and this data structure may not be appropriate.

## Structure

The underlying structure of a VList can be seen as a singly linked list of arrays whose sizes decrease geometrically; in its simplest form, the first contains the first half of the elements in the list, the next the first half of the remainder, and so on. Each of these blocks stores some information such as its size and a pointer to the next.



*An array-list. The reference shown refers to the VList (2,3,4,5,6).*

The average constant-time indexing operation comes directly from this structure; given a random valid index, we simply observe the size of the blocks and follow pointers until we reach the one it should be in. The chance is  $1/2$  that it falls in the first block and we need not follow any pointers; the chance is  $1/4$  we have to follow only one, and so on, so that the expected number of pointers we have to follow is:

$$\sum_{i=1}^{\lceil \log_2 n \rceil} \frac{i-1}{2^i} < \sum_{i=1}^{\infty} \frac{i-1}{2^i} = 1.$$



Any particular reference to a VList is actually a  $\langle base, offset \rangle$  pair indicating the position of its first element in the data structure described above. The *base* part indicates which of the arrays its first element falls in, while the *offset* part indicates its index in that array. This makes it easy to "remove" an element from the front of the list; we simply increase the offset, or increase the base and set the offset to zero if the offset goes out of range. If a particular reference is the last to leave a block, the block will be garbage-collected if such facilities are available, or otherwise must be freed explicitly.

Because the lists are constructed incrementally, the first array in the array list may not contain twice as many values as the next one, although the rest do; this does not significantly impact indexing performance. We nevertheless allocate this much space for the first array, so that if we add more elements to the front of the list in the future we can simply add them to this list and update the size. If the array fills up, we create a new array, twice as large again as this one, and link it to the old first array.

The trickier case, however, is adding a new item to the front of a list, call it A, which starts somewhere in the middle of the array-list data structure. This is the operation that allows VLists to be persistent. To accomplish this, we create a new array, and we link it to the array containing the first element of A. The new array must also store the offset of the first element of A in that array. Then, we can proceed to add any number of items we like to our new array, and any references into this new array will point to VLists which share a tail of values with the old array. Note that with this operation it is possible to create VLists which degenerate into simple linked lists, thus obliterating the performance claims made at the beginning of this article.

## Variants

VList may be modified to support the implementation of a growable array. In the application of a growable array, immutability is no longer required. Instead of growing at the beginning of the list, the ordering interpretation is reversed to allow growing at the end of the array.

## References

- [1] Bagwell, Phil (2002), *Fast Functional Lists, Hash-Lists, Deques and Variable Length Arrays* (<http://infoscience.epfl.ch/record/64410/files/techlists.pdf>), EPFL,

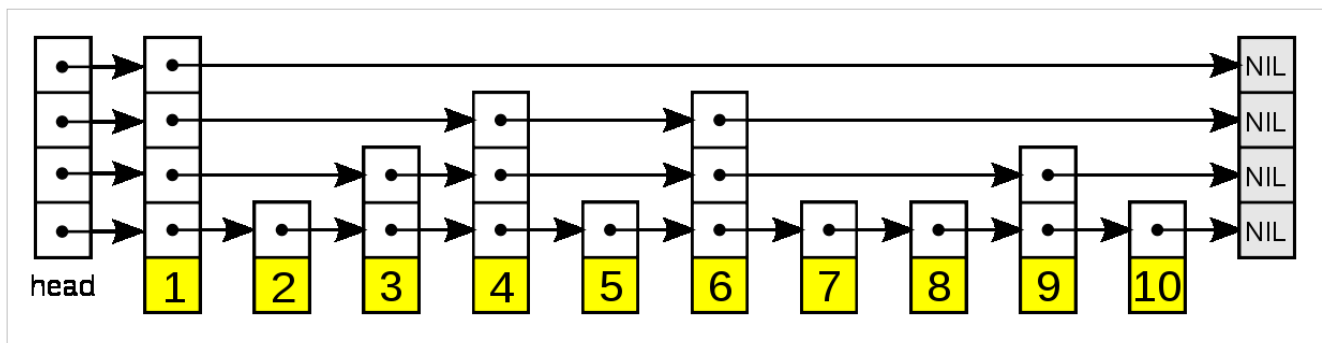
## External links

- C++ implementation of VLists (<http://www.ootl.org/doc/vlist.html>)
- C# implementation of VLists (<http://www.codeproject.com/KB/collections/vlist.aspx>)
- Scheme implementation of VLists and VList-based hash lists (<http://git.savannah.gnu.org/cgit/guile.git/tree/module/ice-9/vlist.scm>) for GNU Guile
- Scheme (Typed Racket) implementation of VLists (<http://planet.plt-scheme.org/package-source/krhari/pfds.plt/1/3/vlist.ss>) for Racket

# Skip list

Skip List		
Type	List	
Invented	1989	
Invented by	W. Pugh	
Time complexity in big O notation		
	Average	Worst case
Space	O(n)	O(nlog n) <sup>[1]</sup>
Search	O(log n)	O(n) <sup>[2]</sup>
Insert	O(log n)	O(n)
Delete	O(log n)	O(n)

A **skip list** is a data structure for storing a sorted list of items using a hierarchy of linked lists that connect increasingly sparse subsequences of the items. These auxiliary lists allow item lookup with efficiency comparable to balanced binary search trees (that is, with number of probes proportional to  $\log n$  instead of  $n$ ).



Each link of the sparser lists skips over many items of the full list in one step, hence the structure's name. These forward links may be added in a randomized way with a geometric / negative binomial distribution.<sup>[3]</sup> Insert, search and delete operations are performed in logarithmic expected time. The links may also be added in a non-probabilistic way so as to guarantee amortized (rather than merely expected) logarithmic cost.

## Description

A skip list is built in layers. The bottom layer is an ordinary ordered linked list. Each higher layer acts as an "express lane" for the lists below, where an element in layer  $i$  appears in layer  $i+1$  with some fixed probability  $p$  (two commonly used values for  $p$  are  $1/2$  or  $1/4$ ). On average, each element appears in  $1/(1-p)$  lists, and the tallest element (usually a special head element at the front of the skip list) in  $\log_{1/p} n$  lists.

A search for a target element begins at the head element in the top list, and proceeds horizontally until the current element is greater than or equal to the target. If the current element is equal to the target, it has been found. If the current element is greater than the target, or the search reaches the end of the linked list, the procedure is repeated after returning to the previous element and dropping down vertically to the next lower list. The expected number of steps in each linked list is at most  $1/p$ , which can be seen by tracing the search path backwards from the target until reaching an element that appears in the next higher list or reaching the beginning of the current list. Therefore, the total *expected* cost of a search is  $(\log_{1/p} n)/p$ , which is  $\mathcal{O}(\log n)$  when  $p$  is a constant. By choosing different

values of  $p$ , it is possible to trade search costs against storage costs.

## Implementation details

The elements used for a skip list can contain more than one pointer since they can participate in more than one list.

Insertions and deletions are implemented much like the corresponding linked-list operations, except that "tall" elements must be inserted into or deleted from more than one linked list.

$\Theta(n)$  operations, which force us to visit every node in ascending order (such as printing the entire list), provide the opportunity to perform a behind-the-scenes derandomization of the level structure of the skip-list in an optimal way, bringing the skip list to  $\mathcal{O}(\log n)$  search time. (Choose the level of the  $i$ 'th finite node to be 1 plus the number of times we can repeatedly divide  $i$  by 2 before it becomes odd. Also,  $i=0$  for the negative infinity header as we have the usual special case of choosing the highest possible level for negative and/or positive infinite nodes.) However this also allows someone to know where all of the higher-than-level 1 nodes are and delete them.

Alternatively, we could make the level structure quasi-random in the following way:

```
make all nodes level 1
j ← 1
while the number of nodes at level j > 1 do
  for each i'th node at level j do
    if i is odd
      if i is not the last node at level j
        randomly choose whether to promote it to level j+1
      else
        do not promote
      end if
    else if i is even and node i-1 was not promoted
      promote it to level j+1
    end if
  repeat
  j ← j + 1
repeat
```

Like the derandomized version, quasi-randomization is only done when there is some other reason to be running a  $\Theta(n)$  operation (which visits every node).

The advantage of this quasi-randomness is that it doesn't give away nearly as much level-structure related information to an adversarial user as the de-randomized one. This is desirable because an adversarial user who is able to tell which nodes are not at the lowest level can pessimize performance by simply deleting higher-level nodes. The search performance is still guaranteed to be logarithmic.

It would be tempting to make the following "optimization": In the part which says "Next, for each  $i$ 'th...", forget about doing a coin-flip for each even-odd pair. Just flip a coin once to decide whether to promote only the even ones or only the odd ones. Instead of  $\Theta(n \lg n)$  coin flips, there would only be  $\Theta(\lg n)$  of them. Unfortunately, this gives the adversarial user a 50/50 chance of being correct upon guessing that all of the even numbered nodes (among the ones at level 1 or higher) are higher than level one. This is despite the property that he has a very low probability of guessing that a particular node is at level  $N$  for some integer  $N$ .

The following proves these two claims concerning the advantages of quasi-randomness over the totally derandomized version. First, to prove that the search time is guaranteed to be logarithmic. Suppose a node  $n$  is searched for, where  $n$  is the position of the found node among the nodes of level 1 or higher. If  $n$  is even, then there is a 50/50 chance that it is higher than level 1. However, if it is not higher than level 1 then node  $n-1$  is guaranteed to

be higher than level 1. If  $n$  is odd, then there is a 50/50 chance that it is higher than level 1. Suppose that it is not; there is a 50/50 chance that node  $n-1$  is higher than level 1. Suppose that this is not either; we are guaranteed that node  $n-2$  is higher than level 1. The analysis can then be repeated for nodes of level 2 or higher, level 3 or higher, etc. always keeping in mind that  $n$  is the position of the node among the ones of level  $k$  or higher for integer  $k$ . So the search time is constant in the best case (if the found node is the highest possible level) and 2 times the worst case for the search time for the totally derandomized skip-list (because we have to keep moving left twice rather than keep moving left once).

Next, an examination of the probability of an adversarial user's guess of a node being level  $k$  or higher being correct. First, the adversarial user has a 50/50 chance of correctly guessing that a particular node is level 2 or higher. This event is independent of whether or not the user correctly guesses at some other node being level 2 or higher. If the user knows the positions of two consecutive nodes of level 2 or higher, and knows that the one on the left is in an odd numbered position among the nodes of level 2 or higher, the user has a 50/50 chance of correctly guessing which one is of level 3 or higher. So, the user's probability of being correct, when guessing that a node is level 3 or higher, is  $1/4$ . Inductively continuing this analysis, we see that the user's probability of guessing that a particular node is level  $k$  or higher is  $1/(2^{k-1})$ .

The above analyses only work when the number of nodes is a power of two. However, because of the third rule which says, "Finally, if  $i$  is odd and also the last node at level 1 then do not promote." (where we substitute the appropriate level number for 1) it becomes a sequence of exact-power-of-two-sized skiplists, concatenated onto each other, for which the analysis does work. In fact, the exact powers of two correspond to the binary representation for the number of nodes in the whole list.

A skip list, upon which we have not recently performed either of the above mentioned  $\Theta(n)$  operations, does not provide the same absolute worst-case performance guarantees as more traditional balanced tree data structures, because it is always possible (though with very low probability) that the coin-flips used to build the skip list will produce a badly balanced structure. However, they work well in practice, and the randomized balancing scheme has been argued to be easier to implement than the deterministic balancing schemes used in balanced binary search trees. Skip lists are also useful in parallel computing, where insertions can be done in different parts of the skip list in parallel without any global rebalancing of the data structure. Such parallelism can be especially advantageous for resource discovery in an ad-hoc Wireless network because a randomized skip list can be made robust to the loss of any single node.<sup>[4]</sup>

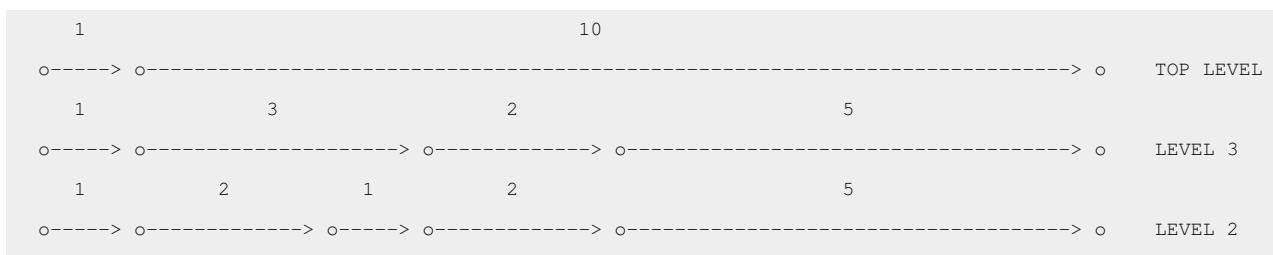
There has been some evidence that skip lists have worse real-world performance and space requirements than B trees due to memory locality and other issues.<sup>[5]</sup>

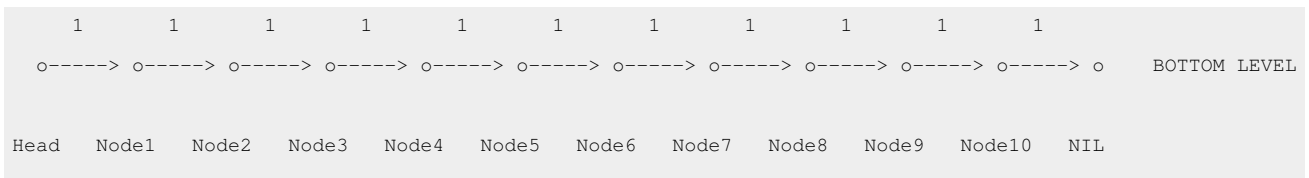
## Indexable skiplist

As described above, a skiplist is capable of fast  $\Theta(\log n)$  insertion and removal of values from a sorted sequence, but it has only slow  $\Theta(n)$  lookups of values at a given position in the sequence (i.e. return the 500th value); however, with a minor modification the speed of random access indexed lookups can be improved to  $\Theta(\log n)$ .

For every link, also store the width of the link. The width is defined as the number of bottom layer links being traversed by each of the higher layer "express lane" links.

For example, here are the widths of the links in the example at the top of the page:





Notice that the width of a higher level link is the sum of the component links below it (i.e. the width 10 link spans the links of widths 3, 2 and 5 immediately below it). Consequently, the sum of all widths is the same on every level ( $10 + 1 = 1 + 3 + 2 + 5 = 1 + 2 + 1 + 2 + 5$ ).

To index the skiplist and find the  $i$ -th value, traverse the skiplist while counting down the widths of each traversed link. Descend a level whenever the upcoming width would be too large.

For example, to find the node in the fifth position (Node 5), traverse a link of width 1 at the top level. Now four more steps are needed but the next width on this level is ten which is too large, so drop one level. Traverse one link of width 3. Since another step of width 2 would be too far, drop down to the bottom level. Now traverse the final link of width 1 to reach the target running total of 5 ( $1+3+1$ ).

```
function lookupByPositionIndex(i)
  node ← head
  i ← i + 1                                # don't count the head as a step
  for level from top to bottom do
    while i ≥ node.width[level] do         # if next step is not too far
      i ← i - node.width[level]           # subtract the current width
      node ← node.next[level]             # traverse forward at the current level
    repeat
  repeat
  return node.value
end function
```

This method of implementing indexing is detailed in Section 3.4 Linear List Operations in "A skip list cookbook" by William Pugh<sup>[6]</sup>.

Also, see Running Median using an Indexable Skiplist<sup>[7]</sup> for a complete implementation written in Python and for an example of it being used to solve a computationally intensive statistics problem. And see Regaining Lost Knowledge<sup>[8]</sup> for the history of that solution.

## History

Skip lists were first described in 1990 by William Pugh. He details how they work in *Skip lists: a probabilistic alternative to balanced trees* in Communications of the ACM, June 1990, 33(6) 668-676. See also citations<sup>[9]</sup> and downloadable documents<sup>[10]</sup>.

To quote the author:

*Skip lists are a probabilistic data structure that seem likely to supplant balanced trees as the implementation method of choice for many applications. Skip list algorithms have the same asymptotic expected time bounds as balanced trees and are simpler, faster and use less space.*

## Usages

List of applications and frameworks that use skip lists:

- Cyrus IMAP<sup>[11]</sup> offers a "skiplist" backend DB implementation (source file<sup>[12]</sup>)
- QMap<sup>[13]</sup> template class of Qt that provides a dictionary.
- Redis<sup>[14]</sup> is an ANSI-C open-source persistent key/value store for Posix systems, uses skip lists in its implementation of ordered sets.
- nessDB<sup>[15]</sup> is a very fast key-value, embedded Database Storage Engine (Using log-structured-merge (LSM) trees), uses skip lists for its memtable.
- skipdb<sup>[16]</sup> is an open-source database format using ordered key/value pairs.
- Running Median using an Indexable Skiplist<sup>[7]</sup> is a Python implementation of a skiplist augmented by link widths to make the structure indexable (giving fast access to the nth item). The indexable skiplist is used to efficiently solve the running median problem (recomputing medians and quartiles as values are added and removed from a large sliding window).
- ConcurrentSkipListSet<sup>[17]</sup> and ConcurrentSkipListMap<sup>[18]</sup> in the Java 1.6 API.

Skip lists are also used in distributed applications (where the nodes represent physical computers, and pointers represent network connections) and for implementing highly scalable concurrent priority queues with less lock contention,<sup>[19]</sup> or even without locking,<sup>[20][21][22]</sup> as well lockless concurrent dictionaries.<sup>[23]</sup> There are also several US patents for using skip lists to implement (lockless) priority queues and concurrent dictionaries.

## References

- [1] <http://www.cs.uwaterloo.ca/research/tr/1993/28/root2side.pdf>
- [2] <http://www.cs.uwaterloo.ca/research/tr/1993/28/root2side.pdf>
- [3] Pugh, William (June 1990). "Skip lists: a probabilistic alternative to balanced trees". *Communications of the ACM* **33** (6): 668–676. doi:10.1145/78973.78977.
- [4] Shah, Gauri Ph.D.; James Aspnes (December 2003) (PDF). *Distributed Data Structures for Peer-to-Peer Systems* (<http://www.cs.yale.edu/homes/shah/pubs/thesis.pdf>). . Retrieved 2008-09-23.
- [5] [http://resnet.uoregon.edu/~gurney\\_j/jmpc/skiplist.html](http://resnet.uoregon.edu/~gurney_j/jmpc/skiplist.html)
- [6] <http://cg.scs.carleton.ca/~morin/teaching/5408/refs/p90b.pdf>
- [7] <http://code.activestate.com/recipes/576930/>
- [8] <http://rhettinger.wordpress.com/2010/02/06/lost-knowledge/>
- [9] <http://citeseer.ist.psu.edu/pugh90skip.html>
- [10] <ftp://ftp.cs.umd.edu/pub/skipLists/>
- [11] <http://cyrusimap.web.cmu.edu/>
- [12] [http://git.cyrusimap.org/cyrus-imapd/tree/lib/cyrusdb\\_skiplist.c](http://git.cyrusimap.org/cyrus-imapd/tree/lib/cyrusdb_skiplist.c)
- [13] <http://doc.trolltech.com/4.5/qmap.html#details>
- [14] <http://redis.io>
- [15] <https://github.com/shuttler/nessDB>
- [16] <http://www.dekorte.com/projects/opensource/skipdb/>
- [17] <http://download.oracle.com/javase/6/docs/api/java/util/concurrent/ConcurrentSkipListSet.html>
- [18] <http://download.oracle.com/javase/6/docs/api/java/util/concurrent/ConcurrentSkipListMap.html>
- [19] Skiplist-based concurrent priority queues (<http://dx.doi.org/10.1109/IPDPS.2000.845994>)
- [20] Sundell, H.; Tsigas, P. (2003). "Fast and lock-free concurrent priority queues for multi-thread systems". *Proceedings International Parallel and Distributed Processing Symposium*. pp. 11. doi:10.1109/IPDPS.2003.1213189. ISBN 0-7695-1926-1.
- [21] Fomitchev, M.; Ruppert, E. (2004). "Lock-free linked lists and skip lists". *Proceedings of the twenty-third annual ACM symposium on Principles of distributed computing - PODC '04*. pp. 50. doi:10.1145/1011767.1011776. ISBN 1581138024.
- [22] Bajpai, R.; Dhara, K. K.; Krishnaswamy, V. (2008). "QPID: A Distributed Priority Queue with Item Locality". *2008 IEEE International Symposium on Parallel and Distributed Processing with Applications*. pp. 215. doi:10.1109/ISPA.2008.90. ISBN 978-0-7695-3471-8.
- [23] Sundell, H. K.; Tsigas, P. (2004). "Scalable and lock-free concurrent dictionaries". *Proceedings of the 2004 ACM symposium on Applied computing - SAC '04*. pp. 1438. doi:10.1145/967900.968188. ISBN 1581138121.

## External links

- Skip Lists: A Probabilistic Alternative to Balanced Trees (<ftp://ftp.cs.umd.edu/pub/skipLists/skiplists.pdf>) - William Pugh's original paper
- "Skip list" entry (<http://nist.gov/dads/HTML/skiplist.html>) in the Dictionary of Algorithms and Data Structures
- Skip Lists: A Linked List with Self-Balancing BST-Like Properties ([http://msdn.microsoft.com/en-us/library/ms379573\(VS.80\).aspx#datastructures20\\_4\\_topic4](http://msdn.microsoft.com/en-us/library/ms379573(VS.80).aspx#datastructures20_4_topic4)) on MSDN in C# 2.0
- SkipDB, a BerkeleyDB-style database implemented using skip lists. (<http://dekorte.com/projects/opensource/SkipDB/>)
- Skip Lists lecture (MIT OpenCourseWare: Introduction to Algorithm) ([http://videolectures.net/mit6046jf05\\_demaine\\_lec12/](http://videolectures.net/mit6046jf05_demaine_lec12/))
- Open Data Structures - Chapter 4 - Skiplists ([http://opendatastructures.org/versions/edition-0.1e/ods-java/4\\_Skiplists.html](http://opendatastructures.org/versions/edition-0.1e/ods-java/4_Skiplists.html))

### Demo applets

- Skip List Applet (<http://people.ksp.sk/~kuko/bak/index.html>) by Kubo Kovac
- Thomas Wenger's demo applet on skiplists (<http://iamwww.unibe.ch/~wenger/DA/SkipList/>)

### Implementations

- A generic Skip List in C++ (<http://codingplayground.blogspot.com/2009/01/generic-skip-list-skiplist.html>) by Antonio Gulli
- Algorithm::SkipList, implementation in Perl on CPAN (<http://search.cpan.org/~rrwo/Algorithm-SkipList-1.02/>)
- John Shipman's implementation in Python (<http://infohost.nmt.edu/tcc/help/lang/python/examples/pyskip/>)
- A Lua port of John Shipman's Python version ([http://love2d.org/wiki/Skip\\_list](http://love2d.org/wiki/Skip_list))
- ConcurrentSkipListSet documentation for Java 6 (<http://java.sun.com/javase/6/docs/api/java/util/concurrent/ConcurrentSkipListSet.html>) (and sourcecode (<http://www.docjar.com/html/api/java/util/concurrent/ConcurrentSkipListSet.java.html>))

# Self-organizing list

---

A **self-organizing list** is a list that reorders its elements based on some self-organising heuristic to improve average access time. The aim of a self organizing list is to improve efficiency of linear search by moving more frequently accessed items towards the head of the list. A self organizing list achieves near constant time for element access in the best case. A self organizing list uses a reorganizing algorithm to adapt to various query distributions at runtime.

## History

The concept of self organizing lists was introduced by McCabe in 1965. In a pioneering work, he introduced two heuristics- the MTF rule and the transposition rule. Further improvements were made, and algorithms suggested by Ronald Rivest, Tenenbaum and Nemes, D. Knuth and so on.

## Introduction

The simplest implementation of a Self Organizing List is as a linked list and thus while being efficient in random node inserting and memory allocation, suffers from inefficient accesses to random nodes. A self organizing list reduces the inefficiency by dynamically rearranging the nodes in the list based on access frequency.

## Inefficiency of Linked List Traversals

If a particular node is to be searched for in the list, each node in the list must be sequentially compared till the desired node is reached. In a linked list, retrieving the  $n$ th element is a  $O(n)$  operation. This is highly inefficient when compared to an array for example, where accessing the  $n^{\text{th}}$  element is a  $O(1)$  operation.

## Efficiency of Self Organizing Lists

A self organizing list rearranges the nodes keeping the most frequently accessed ones at the head of the list. Generally, in a particular query, the chances of accessing a node which has been accessed many times before are higher than the chances of accessing a node which historically has not been so frequently accessed. As a result, keeping the commonly accessed nodes at the head of the list results in reducing the number of comparisons required in an average case to reach the desired node. This leads to better efficiency and generally reduced query times.

## Implementation of a Self Organizing List

The implementation and methods of a self organizing list are identical to the those for a standard linked list. The linked list and the self organizing list differ only in terms of the organization of the nodes; the interface remains the same.

## Analysis of Running Times for Access/ Search in a List

### Average Case

It can be shown that in the average case, the time required to a search on a self organizing list of size  $n$  is

$$T_{avg} = 1 * p(1) + 2 * p(2) + 3 * p(3) + \dots + n * p(n).$$

where  $p(i)$  is the probability of accessing the  $i$ th element in the list, thus also called the access probability. If the access probability of each element is the same (i.e.  $p(1) = p(2) = p(3) = \dots = p(n) = 1/n$ ) then the ordering of the elements is irrelevant and the average time complexity is given by

$$T(n) = 1/n + 2/n + 3/n + \dots + n/n = (1 + 2 + 3 + \dots + n)/n = (n + 1)/2$$


---



and  $T(n)$  does not depend on the individual access probabilities of the elements in the list in this case. However in the case of searches on lists with non uniform record access probabilities (i.e. those lists in which the probability of accessing one element is different from another), the average time complexity can be reduced drastically by proper positioning of the elements contained in the list.

This is done by pairing smaller  $i$  with larger access probabilities so as to reduce the overall average time complexity.

This may be demonstrated as follows:

Given List: A(0.1), B(0.1), C(0.3), D(0.1), E(0.4)

Without rearranging, average search time required is:

$$T(n) = 1 * 0.1 + 2 * 0.1 + 3 * 0.3 + 4 * 0.1 + 5 * 0.4 = 3.6$$

Now suppose the nodes are rearranged so that those nodes with highest probability of access are placed closest to the front so that the rearranged list is now:

E(0.4), C(0.3), D(0.1), A(0.1), B(0.1)

Here, average search time is:

$$T(n) = 1 * 0.4 + 2 * 0.3 + 3 * 0.1 + 4 * 0.1 + 5 * 0.1 = 2.2$$

Thus the average time required for searching in an organized list is (in this case) around 40% less than the time required to search a randomly arranged list.

This is the concept of the self organized list in that the average speed of data retrieval is increased by rearranging the nodes according to access frequency.

## Worst Case

In the worst case, the element to be located is at the very end of the list be it a normal list or a self organized one and thus  $n$  comparisons must be made to reach it. Therefore the worst case running time of a linear search on the list is  $O(n)$  independent of the type of list used. Note that the expression for the average search time in the previous section is a probabilistic one. Keeping the commonly accessed elements at the head of the list simply reduces the probability of the worst case occurring but does not eliminate it completely. Even in a self organizing list, if a lowest access probability element (obviously located at the end of the list) is to be accessed, the entire list must be traversed completely to retrieve it. This is the worst case search.

## Best Case

In the best case, the node to be searched is one which has been commonly accessed and has thus been identified by the list and kept at the head. This will result in a near constant time operation. In big-oh notation, in the best case, accessing an element is an  $O(1)$  operation.

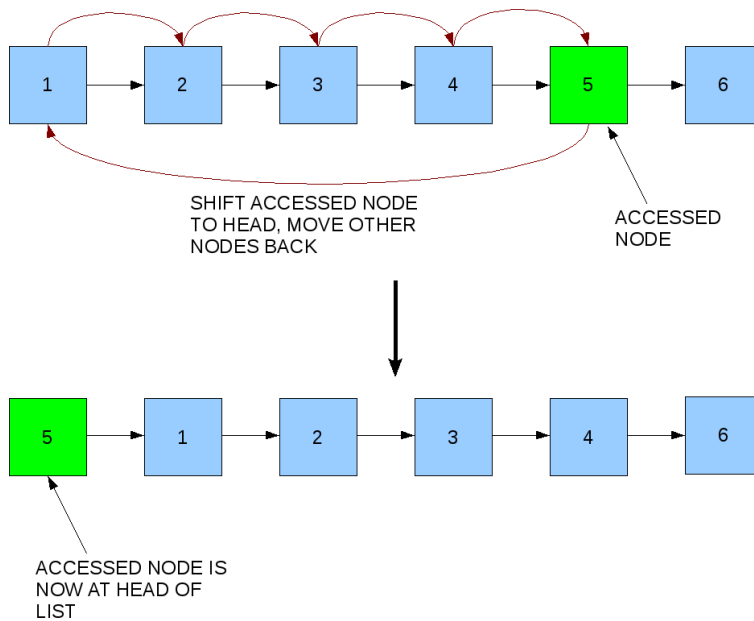
## Techniques for Rearranging Nodes

While ordering the elements in the list, the access probabilities of the elements are not generally known in advance. This has led to the development of various heuristics to approximate optimal behavior. The basic heuristics used to reorder the elements in the list are:

### Move to Front Method (MTF)

This technique moves the element which is accessed to the head of the list. This has the advantage of being easily implemented and requiring no extra memory. This heuristic also adapts quickly to rapid changes in the query distribution. On the other hand, this method may prioritize infrequently accessed nodes-for example, if an uncommon node is accessed even once, it is moved to the head of the list and given maximum priority even if it is not going to be accessed frequently in the future. These 'over rewarded' nodes destroy the optimal ordering of the list and lead to slower access times for commonly accessed elements. Another disadvantage is that this method may

become too flexible leading to access patterns that change too rapidly. This means that due to the very short memories of access patterns even an optimal arrangement of the list can be disturbed immediately by accessing an infrequent node in the list.

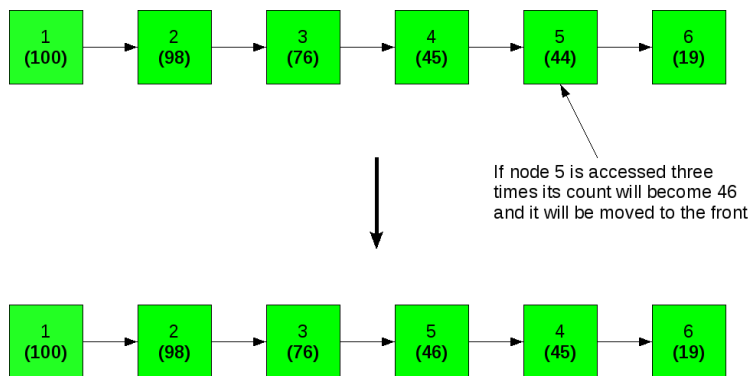


If the 5th node is selected, it is moved to the front

```
At the t-th item selection:
    if item i is selected:
        move item i to head of the list
```

## Count Method

In this technique, the number of times each node was searched for is counted i.e. every node keeps a separate counter variable which is incremented every time it is called. The nodes are then rearranged according to decreasing count. Thus, the nodes of highest count i.e. most frequently accessed are kept at the head of the list. The primary advantage of this technique is that it generally is more realistic in representing the actual access pattern. However, there is an added memory requirement, that of maintaining a counter variable for each node in the list. Also, this technique does not adapt quickly to rapid changes in the access patterns. For example: if the count of the head element say A is 100 and for any node after it say B is 40, then even if B becomes the new most commonly accessed element, it must still be accessed at least  $(100 - 40 = 60)$  times before it can become the head element and thus make the list ordering optimal.



swapped with the 4th

If the 5th node in the list is searched for twice, it will be

```
init: count(i) = 0 for each item i
At t-th item selection:
```

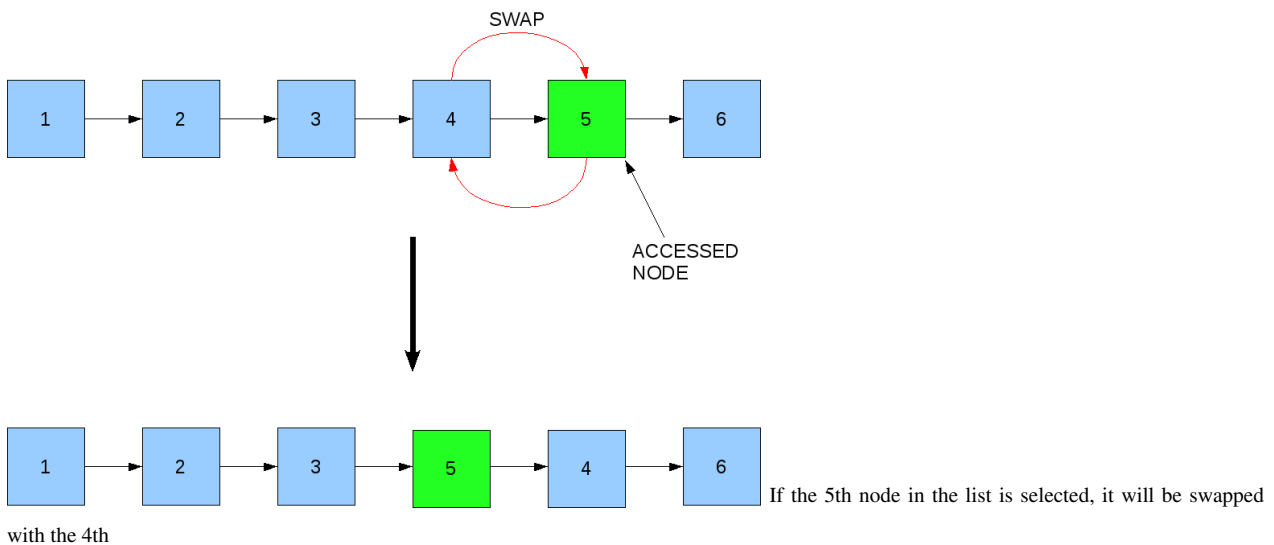
```

if item i is searched:
    count(i) = count(i) + 1
    rearrange items based on count

```

## Transpose Method

This technique involves swapping an accessed node with its predecessor. Therefore, if any node is accessed, it is swapped with the node in front unless it is the head node, thereby increasing its priority. This algorithm is again easy to implement and space efficient and is more likely to keep frequently accessed nodes at the front of the list. However, the transpose method is more cautious. i.e. it will take many accesses to move the element to the head of the list. This method also does not allow for rapid response to changes in the query distributions on the nodes in the list.



```

At the t-th item selection:
if item i is selected:
    if i is not the head of list:
        swap item i with item (i - 1)

```

## Other Methods

Research has been focused on fusing the above algorithms to achieve better efficiency.<sup>[1]</sup> Bitner's Algorithm uses MTF initially and then uses transpose method for finer rearrangements. Some algorithms are randomized and try to prevent the over-rewarding of infrequently accessed nodes that may occur in the MTF algorithm. Other techniques involve reorganizing the nodes based on the above algorithms after every  $n$  accesses on the list as a whole or after  $n$  accesses in a row on a particular node and so on. Some algorithms rearrange the nodes which are accessed based on their proximity to the head node, for example: Swap-With-Parent or Move-To-Parent algorithms. Another class of algorithms are used when the search pattern exhibits a property called locality of reference whereby in a given interval of time, only a smaller subset of the list is probabilistically most likely to be accessed. This is also referred to as dependent access where the probability of the access of a particular element depends on the probability of access of its neighboring elements. Such models are common in real world applications such as database or file systems and memory management and caching. A common framework for algorithms dealing with such dependent environments is to rearrange the list not only based on the record accessed but also on the records near it. This effectively involves reorganizing a sublist of the list to which the record belongs.

## Applications of Self Organizing Lists

Language translators like compilers and interpreters use self organizing lists to maintain symbol tables during compilation or interpretation of program source code. Currently research is underway to incorporate the self organizing list data structure in embedded systems to reduce bus transition activity which leads to power dissipation in those circuits. These lists are also used in artificial intelligence and neural networks as well as self adjusting programs. The algorithms used in self organizing lists are also used as caching algorithms as in the case of LFU algorithm.

## References

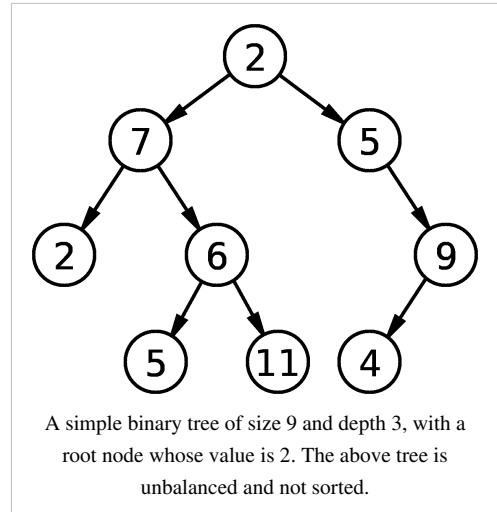
- [1] <http://www.springerlink.com/content/978-3-540-34597-8/#section=508698&page=1&locus=3> Lists on Lists: A Framework for Self Organizing-Lists in Environments with Locality of Reference
  - Vlajic, N (2003) (pdf), *Self-Organizing Lists* ([http://www.cse.yorku.ca/course\\_archive/2003-04/F/2011/2011A/DatStr\\_071\\_SOLists.pdf](http://www.cse.yorku.ca/course_archive/2003-04/F/2011/2011A/DatStr_071_SOLists.pdf))
  - (pdf) *Self Organization* (<http://courses.cs.vt.edu/~cs2604/spring04/Notes/C16.SelfOrganizingLists.pdf>), 2004
  - NIST DADS entry (<http://www.nist.gov/dads/HTML/selfOrganizingList.html>)
  - A Drozdek, Data Structures and Algorithms in Java Third edition
  - Amer, Abdelrehman; B. John Oommen (2006) (pdf), Lists on Lists: A Framework for Self Organizing-Lists in Environments with Locality of Reference *Lists on Lists: A Framework for Self-organizing Lists in Environments with Locality of Reference* (<http://www.springerlink.com/content/978-3-540-34597-8/#section=508698&page=1&locus=3>)
-

# Binary trees

## Binary tree

In computer science, a **binary tree** is a tree data structure in which each node has at most two child nodes, usually distinguished as "left" and "right". Nodes with children are parent nodes, and child nodes may contain references to their parents. Outside the tree, there is often a reference to the "root" node (the ancestor of all nodes), if it exists. Any node in the data structure can be reached by starting at root node and repeatedly following references to either the left or right child. A tree which does not have any node other than root node is called a null tree. In a binary tree a degree of every node is maximum two. A tree with 'n' number of nodes has exactly 'n-1' branches or degree.

Binary trees are used to implement binary search trees and binary heaps.



### Definitions for rooted trees

- A **directed edge** refers to the link from the parent to the child (the arrows in the picture of the tree).
- The root node of a tree is the node with no parents. There is at most one root node in a rooted tree.
- A leaf node has no children.
- The **depth** of a node  $n$  is the length of the path from the root to the node. The set of all nodes at a given depth is sometimes called a **level** of the tree. The root node is at depth zero.
- The **depth** (or *height*) of a tree is the length of the path from the root to the deepest node in the tree. A (rooted) tree with only one node (the root) has a depth of zero.
- **Siblings** are nodes that share the same parent node.
- A node  $p$  is an **ancestor** of a node  $q$  if it exists on the path from the root to node  $q$ . The node  $q$  is then termed as a **descendant** of  $p$ .
- The **size** of a node is the number of descendants it has including itself.
- **In-degree** of a node is the number of edges arriving at that node.
- **Out-degree** of a node is the number of edges leaving that node.
- The root is the only node in the tree with In-degree = 0.
- All the leaf nodes have Out-degree = 0.

## Types of binary trees

- A **rooted binary tree** is a tree with a root node in which every node has at most two children.
- A **full binary tree** (sometimes **proper binary tree** or **2-tree** or **strictly binary tree**) is a tree in which every node other than the leaves has two children. Sometimes a full tree is ambiguously defined as a perfect tree.
- A **perfect binary tree** is a *full binary tree* in which all *leaves* are at the same *depth* or same *level*, and in which every parent has two children.<sup>[1]</sup> (This is ambiguously also called a *complete binary tree*.)
- A **complete binary tree** is a binary tree in which every level, *except possibly the last*, is completely filled, and all nodes are as far left as possible.<sup>[2]</sup>
- An **infinite complete binary tree** is a tree with a countably infinite number of levels, in which every node has two children, so that there are  $2^d$  nodes at level  $d$ . The set of all nodes is countably infinite, but the set of all infinite paths from the root is uncountable: it has the cardinality of the continuum. These paths corresponding by an order preserving bijection to the points of the Cantor set, or (through the example of the Stern–Brocot tree) to the set of positive irrational numbers.
- A **balanced binary tree** is commonly defined as a binary tree in which the depth of the two subtrees of every node never differ by more than 1,<sup>[3]</sup> although in general it is a binary tree where no leaf is much farther away from the root than any other leaf. (Different balancing schemes allow different definitions of "much farther"<sup>[4]</sup>). Binary trees that are balanced according to this definition have a predictable depth (how many nodes are traversed from the root to a leaf, root counting as node 0 and subsequent as 1, 2, ..., depth). This depth is equal to the integer part of  $\log_2(n)$  where  $n$  is the number of nodes on the balanced tree. Example 1: balanced tree with 1 node,  $\log_2(1) = 0$  (depth = 0). Example 2: balanced tree with 3 nodes,  $\log_2(3) = 1.59$  (depth=1). Example 3: balanced tree with 5 nodes,  $\log_2(5) = 2.32$  (depth of tree is 2 nodes).
- A **rooted complete binary tree** can be identified with a free magma.
- A **degenerate tree** is a tree where for each parent node, there is only one associated child node. This means that in a performance measurement, the tree will behave like a linked list data structure.

Note that this terminology often varies in the literature, especially with respect to the meaning of "complete" and "full".

## Properties of binary trees

- The number  $n$  of nodes in a perfect binary tree can be found using this formula:  $n = 2^{h+1} - 1$  where  $h$  is the depth of the tree.
- The number  $n$  of nodes in a complete binary tree is at least  $n = 2^h$  and at most  $n = 2^{h+1} - 1$  where  $h$  is the depth of the tree.
- The number  $L$  of leaf nodes in a perfect binary tree can be found using this formula:  $L = 2^h$  where  $h$  is the depth of the tree.
- The number  $n$  of nodes in a perfect binary tree can also be found using this formula:  $n = 2L - 1$  where  $L$  is the number of leaf nodes in the tree.
- The number of null links (absent children of nodes) in a complete binary tree of  $n$  nodes is  $(n + 1)$ .
- The number  $n - L$  of internal nodes (non-leaf nodes) in a Complete Binary Tree of  $n$  nodes is  $\lfloor n/2 \rfloor$ .
- For any non-empty binary tree with  $n_0$  leaf nodes and  $n_2$  nodes of degree 2,  $n_0 = n_2 + 1$ .<sup>[5]</sup>

Proof:

Let  $n$  = the total number of nodes

$B$  = number of branches

$n_0, n_1, n_2$  represent the number of nodes with no children, a single child, and two children respectively.

$B = n - 1$  (since all nodes except the root node come from a single branch)

$$B = n_1 + 2 * n_2$$

$$n = n_1 + 2 * n_2 + 1$$

$$n = n_0 + n_1 + n_2$$

$$n_1 + 2 * n_2 + 1 = n_0 + n_1 + n_2 \implies n_0 = n_2 + 1$$

## Common operations

There are a variety of different operations that can be performed on binary trees. Some are mutator operations, while others simply return useful information about the tree.

### Insertion

Nodes can be inserted into binary trees in between two other nodes or added after an external node. In binary trees, a node that is inserted is specified as to which child it is.

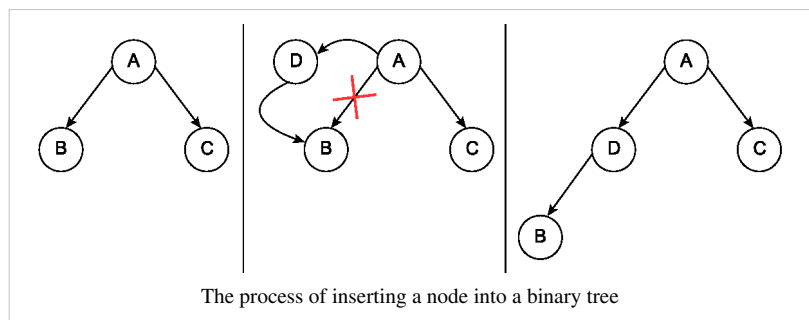
#### External nodes

Say that the external node being added on to is node A. To add a new node after node A, A assigns the new node as one of its children and the new node assigns node A as its parent.

#### Internal nodes

Insertion on internal nodes is slightly more complex than on external nodes. Say that the internal node is node A and that node B is the child of A. (If the insertion is to insert a right child, then B is the right child of A, and similarly with a left child insertion.) A assigns its child to the new node and the new node assigns its parent to A.

Then the new node assigns its child to B and B assigns its parent as the new node.

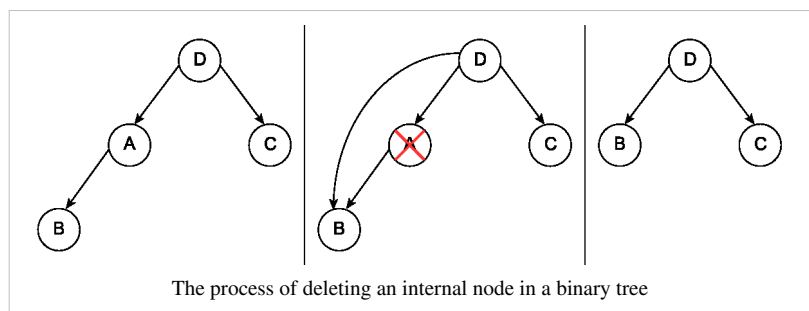


### Deletion

Deletion is the process whereby a node is removed from the tree. Only certain nodes in a binary tree can be removed unambiguously.<sup>[6]</sup>

#### Node with zero or one children

Say that the node to delete is node A. If a node has no children (external node), deletion is accomplished by setting the child of A's parent to null. If it has one child, set the parent of A's child to A's parent and set the child of A's parent to A's child.



### Node with two children

In a binary tree, a node with two children cannot be deleted unambiguously.<sup>[6]</sup> However, in certain binary trees these nodes *can* be deleted, including binary search trees.

### Iteration

Often, one wishes to visit each of the nodes in a tree and examine the value there, a process called iteration or enumeration. There are several common orders in which the nodes can be visited, and each has useful properties that are exploited in algorithms based on binary trees:

- In-Order: Left child, Root, Right child.
- Pre-Order: Root, Left child, Right child
- Post-Order: Left Child, Right child, Root

### Pre-order, in-order, and post-order traversal

Pre-order, in-order, and post-order traversal visit each node in a tree by recursively visiting each node in the left and right subtrees of the root.

### Depth-first order

In depth-first order, we always attempt to visit the node farthest from the root that we can, but with the caveat that it must be a child of a node we have already visited. Unlike a depth-first search on graphs, there is no need to remember all the nodes we have visited, because a tree cannot contain cycles. Pre-order is a special case of this. See depth-first search for more information.

### Breadth-first order

Contrasting with depth-first order is breadth-first order, which always attempts to visit the node closest to the root that it has not already visited. See breadth-first search for more information. Also called a *level-order traversal*.

## Type theory

In type theory, a binary tree with nodes of type  $A$  is defined inductively as  $T_A = \mu\alpha. 1 + A \times \alpha \times \alpha$ .

## Definition in graph theory

For each binary tree data structure, there is equivalent rooted binary tree in graph theory.

Graph theorists use the following definition: A binary tree is a connected acyclic graph such that the degree of each vertex is no more than three. It can be shown that in any binary tree of two or more nodes, there are exactly two more nodes of degree one than there are of degree three, but there can be any number of nodes of degree two. A **rooted binary tree** is such a graph that has one of its vertices of degree no more than two singled out as the root.

With the root thus chosen, each vertex will have a uniquely defined parent, and up to two children; however, so far there is insufficient information to distinguish a left or right child. If we drop the connectedness requirement, allowing multiple connected components in the graph, we call such a structure a forest.

Another way of defining binary trees is a recursive definition on directed graphs. A binary tree is either:

- A single vertex.
- A graph formed by taking two binary trees, adding a vertex, and adding an edge directed from the new vertex to the root of each binary tree.

This also does not establish the order of children, but does fix a specific root node.



## Combinatorics

In combinatorics one considers the problem of counting the number of full binary trees of a given size. Here the trees have no values attached to their nodes (this would just multiply the number of possible trees by an easily determined factor), and trees are distinguished only by their structure; however the left and right child of any node are distinguished (if they are different trees, then interchanging them will produce a tree distinct from the original one). The size of the tree is taken to be the number  $n$  of internal nodes (those with two children); the other nodes are leaf nodes and there are  $n + 1$  of them. The number of such binary trees of size  $n$  is equal to the number of ways of fully parenthesizing a string of  $n + 1$  symbols (representing leaves) separated by  $n$  binary operators (representing internal nodes), so as to determine the argument subexpressions of each operator. For instance for  $n = 3$  one has to parenthesize a string like  $X * X * X * X$ , which is possible in five ways:

$$((X*X)*X)*X, \quad (X*(X*X))*X, \quad (X*X)*(X*X), \quad X*((X*X)*X), \quad X*(X*(X*X)).$$

The correspondence to binary trees should be obvious, and the addition of redundant parentheses (around an already parenthesized expression or around the full expression) is disallowed (or at least not counted as producing a new possibility).

There is a unique binary tree of size 0 (consisting of a single leaf), and any other binary tree is characterized by the pair of its left and right children; if these have sizes  $i$  and  $j$  respectively, the full tree has size  $i + j + 1$ . Therefore the number  $C_n$  of binary trees of size  $n$  has the following recursive description  $C_0 = 1$ , and  $C_n = \sum_{i=0}^{n-1} C_i C_{n-1-i}$  for any positive integer  $n$ . It follows that  $C_n$  is the Catalan number of index  $n$ .

The above parenthesized strings should not be confused with the set of words of length  $2n$  in the Dyck language, which consist only of parentheses in such a way that they are properly balanced. The number of such strings satisfies the same recursive description (each Dyck word of length  $2n$  is determined by the Dyck subword enclosed by the initial '(' and its matching ')' together with the Dyck subword remaining after that closing parenthesis, whose lengths  $2i$  and  $2j$  satisfy  $i + j + 1 = n$ ); this number is therefore also the Catalan number  $C_n$ . So there are also five Dyck words of length 10:

$$()()(), \quad ()(()), \quad (())(), \quad (()), \quad ((())).$$

These Dyck words do not correspond in an obvious way to binary trees. A bijective correspondence can nevertheless be defined as follows: enclose the Dyck word in a extra pair of parentheses, so that the result can be interpreted as a Lisp list expression (with the empty list () as only occurring atom); then the dotted-pair expression for that proper list is a fully parenthesized expression (with NIL as symbol and '.' as operator) describing the corresponding binary tree (which is in fact the internal representation of the proper list).

The ability to represent binary trees as strings of symbols and parentheses implies that binary trees can represent the elements of a free magma on a singleton set.

## Methods for storing binary trees

Binary trees can be constructed from programming language primitives in several ways.

### Nodes and references

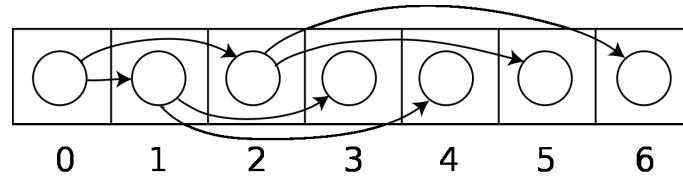
In a language with records and references, binary trees are typically constructed by having a tree node structure which contains some data and references to its left child and its right child. Sometimes it also contains a reference to its unique parent. If a node has fewer than two children, some of the child pointers may be set to a special null value, or to a special sentinel node.

In languages with tagged unions such as ML, a tree node is often a tagged union of two types of nodes, one of which is a 3-tuple of data, left child, and right child, and the other of which is a "leaf" node, which contains no data and functions much like the null value in a language with pointers.

## Arrays

Binary trees can also be stored in breadth-first order as an implicit data structure in arrays, and if the tree is a complete binary tree, this method wastes no space. In this compact arrangement, if a node has an index  $i$ , its children are found at indices  $2i + 1$  (for the left child) and  $2i + 2$  (for the right), while its parent (if any) is found at index  $\left\lfloor \frac{i-1}{2} \right\rfloor$  (assuming the root has index zero). This method benefits from more compact storage and better locality of reference, particularly during a preorder traversal. However, it is expensive to grow and wastes space proportional to  $2^h - n$  for a tree of depth  $h$  with  $n$  nodes.

This method of storage is often used for binary heaps. No space is wasted because nodes are added in breadth-first order.



## Encodings

### Succinct encodings

A succinct data structure is one which takes the absolute minimum possible space, as established by information theoretical lower bounds. The number of different binary trees on  $n$  nodes is  $C_n$ , the  $n$ th Catalan number (assuming we view trees with identical *structure* as identical). For large  $n$ , this is about  $4^n$ ; thus we need at least about  $\log_2 4^n = 2n$  bits to encode it. A succinct binary tree therefore would occupy only 2 bits per node.

One simple representation which meets this bound is to visit the nodes of the tree in preorder, outputting "1" for an internal node and "0" for a leaf. [7] If the tree contains data, we can simply simultaneously store it in a consecutive array in preorder. This function accomplishes this:

```
function EncodeSuccinct(node n, bitstring structure, array data) {
    if n = nil then
        append 0 to structure;
    else
        append 1 to structure;
        append n.data to data;
        EncodeSuccinct(n.left, structure, data);
        EncodeSuccinct(n.right, structure, data);
}
```

The string *structure* has only  $2n + 1$  bits in the end, where  $n$  is the number of (internal) nodes; we don't even have to store its length. To show that no information is lost, we can convert the output back to the original tree like this:

```
function DecodeSuccinct(bitstring structure, array data) {
    remove first bit of structure and put it in b
    if b = 1 then
        create a new node n
        remove first element of data and put it in n.data
        n.left = DecodeSuccinct(structure, data)
        n.right = DecodeSuccinct(structure, data)
    return n
```

```

else
    return nil
}

```

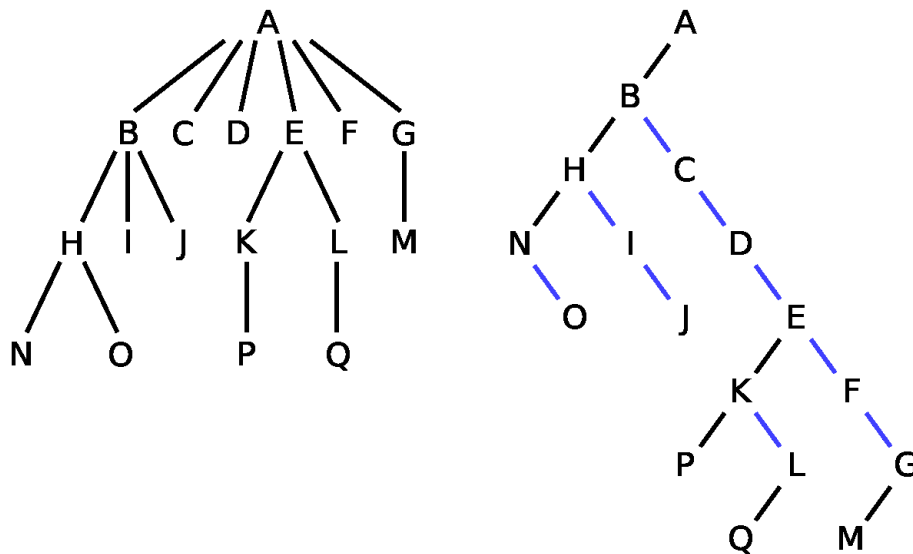
More sophisticated succinct representations allow not only compact storage of trees but even useful operations on those trees directly while they're still in their succinct form.

### Encoding general trees as binary trees

There is a one-to-one mapping between general ordered trees and binary trees, which in particular is used by Lisp to represent general ordered trees as binary trees. To convert a general ordered tree to binary tree, we only need to represent the general tree in left child-sibling way. The result of this representation will be automatically binary tree, if viewed from a different perspective. Each node  $N$  in the ordered tree corresponds to a node  $N'$  in the binary tree; the *left* child of  $N'$  is the node corresponding to the first child of  $N$ , and the *right* child of  $N'$  is the node corresponding to  $N$ 's next sibling --- that is, the next node in order among the children of the parent of  $N$ . This binary tree representation of a general order tree is sometimes also referred to as a left child-right sibling binary tree (LCRS tree), or a doubly chained tree, or a Filial-Heir chain.

One way of thinking about this is that each node's children are in a linked list, chained together with their *right* fields, and the node only has a pointer to the beginning or head of this list, through its *left* field.

For example, in the tree on the left, A has the 6 children {B,C,D,E,F,G}. It can be converted into the binary tree on the right.



The binary tree can be thought of as the original tree tilted sideways, with the black left edges representing *first child* and the blue right edges representing *next sibling*. The leaves of the tree on the left would be written in Lisp as:

```
((N O) I J) C D ((P) (Q)) F (M))
```

which would be implemented in memory as the binary tree on the right, without any letters on those nodes that have a left child.

## Notes

- [1] "perfect binary tree" (<http://www.nist.gov/dads/HTML/perfectBinaryTree.html>). NIST. .
- [2] "complete binary tree" (<http://www.nist.gov/dads/HTML/completeBinaryTree.html>). NIST. .
- [3] Aaron M. Tenenbaum, et. al Data Structures Using C, Prentice Hall, 1990 ISBN 0-13-199746-7
- [4] Paul E. Black (ed.), entry for *data structure* in *Dictionary of Algorithms and Data Structures*. U.S. National Institute of Standards and Technology. 15 December 2004. Online version (<http://xw2k.nist.gov/dads/HTML/balancedtree.html>) Accessed 2010-12-19.
- [5] Mehta, Dinesh; Sartaj Sahni (2004). *Handbook of Data Structures and Applications*. Chapman and Hall. ISBN 1-58488-435-5.
- [6] Dung X. Nguyen (2003). "Binary Tree Structure" (<http://www.clear.rice.edu/comp212/03-spring/lectures/22/>). rice.edu. . Retrieved December 28, 2010.
- [7] [http://theory.csail.mit.edu/classes/6.897/spring03/scribe\\_notes/L12/lecture12.pdf](http://theory.csail.mit.edu/classes/6.897/spring03/scribe_notes/L12/lecture12.pdf)

## References

- Donald Knuth. *The art of computer programming vol 1. Fundamental Algorithms*, Third Edition. Addison-Wesley, 1997. ISBN 0-201-89683-4. Section 2.3, especially subsections 2.3.1–2.3.2 (pp. 318–348).
- Kenneth A Berman, Jerome L Paul. *Algorithms: Parallel, Sequential and Distributed*. Course Technology, 2005. ISBN 0-534-42057-5. Chapter 4. (pp. 113–166).

## External links

- flash actionscript 3 opensource implementation of binary tree (<http://www.dpdn.nl/opensource>) — opensource library
- (<http://www.gamedev.net/reference/programming/features/trees2/>) — GameDev.net's article about binary trees
- (<http://www.brpreiss.com/books/opus4/html/page355.html>) — Binary Tree Proof by Induction
- Balanced binary search tree on array (<http://piergiu.wordpress.com/2010/02/21/balanced-binary-search-tree-on-array/>) How to create bottom-up an Ahnentafel list, or a balanced binary search tree on array
- A simple binary tree implementation in C++ (<http://www.technical-recipes.com/2011/simple-binary-tree-implementation/>)

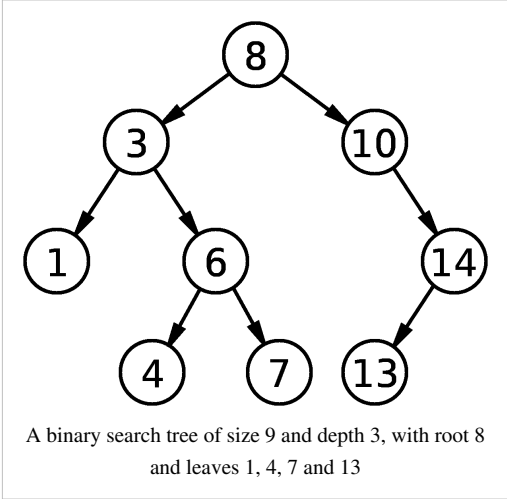
# Binary search tree

Binary search tree		
Type	Tree	
Time complexity in big O notation		
	Average	Worst case
Space	O(n)	O(n)
Search	O(log n)	O(n)
Insert	O(log n)	O(n)
Delete	O(log n)	O(n)

In computer science, a **binary search tree (BST)**, which may sometimes also be called an **ordered** or **sorted binary tree**, is a node-based binary tree data structure which has the following properties:<sup>[1]</sup>

- The left subtree of a node contains only nodes with keys less than the node's key.
- The right subtree of a node contains only nodes with keys greater than or equal to the node's key.
- Both the left and right subtrees must also be binary search trees.

Generally, the information represented by each node is a record rather than a single data element. However, for sequencing purposes, nodes are compared according to their keys rather than any part of their associated records.



The major advantage of binary search trees over other data structures is that the related sorting algorithms and search algorithms such as in-order traversal can be very efficient.

Binary search trees are a fundamental data structure used to construct more abstract data structures such as sets, multisets, and associative arrays.

## Operations

Operations on a binary search tree require comparisons between nodes. These comparisons are made with calls to a comparator, which is a subroutine that computes the total order (linear order) on any two values. This comparator can be explicitly or implicitly defined, depending on the language in which the BST is implemented.

## Searching

Searching a binary search tree for a specific value can be a recursive or iterative process. This explanation covers a recursive method.

We begin by examining the root node. If the tree is null, the value we are searching for does not exist in the tree. Otherwise, if the value equals the root, the search is successful. If the value is less than the root, search the left subtree. Similarly, if it is greater than the root, search the right subtree. This process is repeated until the value is found or the indicated subtree is null. If the searched value is not found before a null subtree is reached, then the item

must not be present in the tree.

Here is the search algorithm in the Python language:

```
# 'node' refers to the parent-node in this case
def search_binary_tree(node, key):
    if node is None:
        return None # key not found
    if key < node.key:
        return search_binary_tree(node.leftChild, key)
    if key > node.key:
        return search_binary_tree(node.rightChild, key)
    else: # key is equal to node key
        return node.value # found key
```

... or equivalent Haskell:

```
searchBinaryTree _ NullNode = Nothing
searchBinaryTree key (Node nodeKey nodeValue (leftChild, rightChild)) =
    case compare key nodeKey of
        LT -> searchBinaryTree key leftChild
        GT -> searchBinaryTree key rightChild
        EQ -> Just nodeValue
```

This operation requires  $O(\log n)$  time in the average case, but needs  $O(n)$  time in the worst case, when the unbalanced tree resembles a linked list (degenerate tree).

Assuming that `BinarySearchTree` is a class with a member function `search(int)` and a pointer to the root node, the algorithm is also easily implemented in terms of an iterative approach. The algorithm enters a loop, and decides whether to branch left or right depending on the value of the node at each parent node.

```
bool BinarySearchTree::search(int val)
{
    Node *next = this->root();

    while (next != NULL) {
        if (val == next->value()) {
            return true;
        } else if (val < next->value()) {
            next = next->left();
        } else {
            next = next->right();
        }
    }

    //not found
    return false;
}
```

## Insertion

Insertion begins as a search would begin; if the root is not equal to the value, we search the left or right subtrees as before. Eventually, we will reach an external node and add the value as its right or left child, depending on the node's value. In other words, we examine the root and recursively insert the new node to the left subtree if the new value is less than the root, or the right subtree if the new value is greater than or equal to the root.

Here's how a typical binary search tree insertion might be performed in C++:

```
/* Inserts the node pointed to by "newNode" into the subtree rooted at
"treeNode" */
void InsertNode(Node* &treeNode, Node *newNode)
{
    if (treeNode == NULL)
        treeNode = newNode;
    else if (newNode->key < treeNode->key)
        InsertNode(treeNode->left, newNode);
    else
        InsertNode(treeNode->right, newNode);
}
```

The above *destructive* procedural variant modifies the tree in place. It uses only constant space, but the prior version of the tree is lost. Alternatively, as in the following Python example, we can reconstruct all ancestors of the inserted node; any reference to the original tree root remains valid, making the tree a persistent data structure:

```
def binary_tree_insert(node, key, value):
    if node is None:
        return TreeNode(None, key, value, None)
    if key == node.key:
        return TreeNode(node.left, key, value, node.right)
    if key < node.key:
        return TreeNode(binary_tree_insert(node.left, key, value),
node.key, node.value, node.right)
    else:
        return TreeNode(node.left, node.key, node.value,
binary_tree_insert(node.right, key, value))
```

The part that is rebuilt uses  $\Theta(\log n)$  space in the average case and  $O(n)$  in the worst case (see big-O notation).

In either version, this operation requires time proportional to the height of the tree in the worst case, which is  $O(\log n)$  time in the average case over all trees, but  $O(n)$  time in the worst case.

Another way to explain insertion is that in order to insert a new node in the tree, its value is first compared with the value of the root. If its value is less than the root's, it is then compared with the value of the root's left child. If its value is greater, it is compared with the root's right child. This process continues, until the new node is compared with a leaf node, and then it is added as this node's right or left child, depending on its value.

There are other ways of inserting nodes into a binary tree, but this is the only way of inserting nodes at the leaves and at the same time preserving the BST structure.

Here is an iterative approach to inserting into a binary search tree in Java:

```
private Node m_root;

public void insert(int data) {
```

```

    if (m_root == null) {
        m_root = new TreeNode(data, null, null);
        return;
    }
    Node root = m_root;
    while (root != null) {
        // Not the same value twice
        if (data == root.getData()) {
            return;
        } else if (data < root.getData()) {
            // insert left
            if (root.getLeft() == null) {
                root.setLeft(new TreeNode(data, null, null));
                return;
            } else {
                root = root.getLeft();
            }
        } else {
            // insert right
            if (root.getRight() == null) {
                root.setRight(new TreeNode(data, null, null));
                return;
            } else {
                root = root.getRight();
            }
        }
    }
}

```

Below is a recursive approach to the insertion method.

```

private Node m_root;

public void insert(int data) {
    if (m_root == null) {
        m_root = new TreeNode(data, null, null);
    } else {
        internalInsert(m_root, data);
    }
}

private static void internalInsert(Node node, int data) {
    // Not the same value twice
    if (data == node.getValue()) {
        return;
    } else if (data < node.getValue()) {
        if (node.getLeft() == null) {
            node.setLeft(new TreeNode(data, null, null));
        }
    }
}

```



```

    }else{
        internalInsert (node.getLeft(), data);
    }
} else{
    if (node.getRight() == null) {
        node.setRight(new TreeNode(data, null, null));
    }else{
        internalInsert (node.getRight(), data);
    }
}
}
}

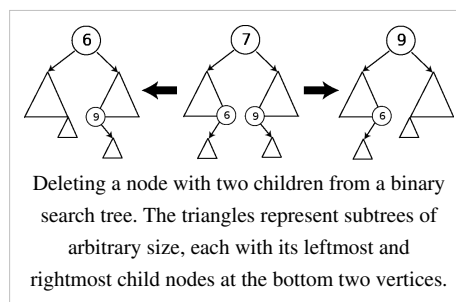
```

## Deletion

There are three possible cases to consider:

- **Deleting a leaf (node with no children):** Deleting a leaf is easy, as we can simply remove it from the tree.
- **Deleting a node with one child:** Remove the node and replace it with its child.
- **Deleting a node with two children:** Call the node to be deleted  $N$ . Do not delete  $N$ . Instead, choose either its in-order successor node or its in-order predecessor node,  $R$ . Replace the value of  $N$  with the value of  $R$ , then delete  $R$ .

As with all binary trees, a node's in-order successor is the left-most child of its right subtree, and a node's in-order predecessor is the right-most child of its left subtree. In either case, this node will have zero or one children. Delete it according to one of the two simpler cases above.



Consistently using the in-order successor or the in-order predecessor for every instance of the two-child case can lead to an unbalanced tree, so good implementations add inconsistency to this selection.

Running time analysis: Although this operation does not always traverse the tree down to a leaf, this is always a possibility; thus in the worst case it requires time proportional to the height of the tree. It does not require more even when the node has two children, since it still follows a single path and does not visit any node twice.

Here is the code in Python:

```

def findMin(self):
    """
    Finds the smallest element that is a child of *self*
    """
    current_node = self
    while current_node.left_child:
        current_node = current_node.left_child
    return current_node

```

```

def replace_node_in_parent(self, new_value=None):
    """
    Removes the reference to *self* from *self.parent* and replaces it
    with *new_value*.
    """
    if self.parent:
        if self == self.parent.left_child:
            self.parent.left_child = new_value
        else:
            self.parent.right_child = new_value
    if new_value:
        new_value.parent = self.parent

def binary_tree_delete(self, key):
    if key < self.key:
        self.left_child.binary_tree_delete(key)
    elif key > self.key:
        self.right_child.binary_tree_delete(key)
    else: # delete the key here
        if self.left_child and self.right_child: # if both children are
present
            # get the smallest node that's bigger than *self*
            successor = self.right_child.findMin()
            self.key = successor.key
            # if *successor* has a child, replace it with that
            # at this point, it can only have a *right_child*
            # if it has no children, *right_child* will be "None"
            successor.replace_node_in_parent(successor.right_child)
        elif self.left_child or self.right_child: # if the node has
only one child
            if self.left_child:
                self.replace_node_in_parent(self.left_child)
            else:
                self.replace_node_in_parent(self.right_child)
        else: # this node has no children
            self.replace_node_in_parent(None)

```

## Traversal

Once the binary search tree has been created, its elements can be retrieved in-order by recursively traversing the left subtree of the root node, accessing the node itself, then recursively traversing the right subtree of the node, continuing this pattern with each node in the tree as it's recursively accessed. As with all binary trees, one may conduct a pre-order traversal or a post-order traversal, but neither are likely to be useful for binary search trees.

The code for in-order traversal in Python is given below. It will call **callback** for every node in the tree.

```

def traverse_binary_tree(node, callback):
    if node is None:
        return

```

```

    traverse_binary_tree(node.leftChild, callback)
    callback(node.value)
    traverse_binary_tree(node.rightChild, callback)

```

Traversal requires  $O(n)$  time, since it must visit every node. This algorithm is also  $O(n)$ , so it is asymptotically optimal.

An in-order traversal algorithm for C is given below.

```

void in_order_traversal(struct Node *n, void (*cb)(void*))
{
    struct Node *cur, *pre;

    if(!n)
        return;

    cur = n;

    while(cur) {
        if(!cur->left) {
            cb(cur->val);
            cur = cur->right;
        } else {
            pre = cur->left;

            while(pre->right && pre->right != cur)
                pre = pre->right;

            if (!pre->right) {
                pre->right = cur;
                cur = cur->left;
            } else {
                pre->right = NULL;
                cb(cur->val);
                cur = cur->right;
            }
        }
    }
}

```

An alternate recursion-free algorithm for in-order traversal using a stack and `goto` statements is provided below. The stack contains nodes whose right subtrees have yet to be explored. If a node has an unexplored left subtree (a condition tested at the `try_left` label, then the node is pushed (marking its right subtree for future exploration) and the algorithm descends to the left subtree. The purpose of the `loop_top` label is to avoid moving to the left subtree when popping to a node (as popping to a node indicates that its left subtree has already been explored.)

```

void in_order_traversal(struct Node *n, void (*cb)(void*))
{
    struct Node *cur;
    struct Stack *stack;

```

```

    if (!n)
        return;

    stack = stack_create();
    cur = n;

try_left:
    /* check for the left subtree */
    if (cur->left) {
        stack_push(stack, cur);
        cur = cur->left;
        goto try_left;
    }

loop_top:
    /* call callback */
    cb(cur->val);

    /* check for the right subtree */
    if (cur->right) {
        cur = cur->right;
        goto try_left;
    }

    cur = stack_pop(stack);
    if (cur)
        goto loop_top;

    stack_destroy(stack);
}

```

## Sort

A binary search tree can be used to implement a simple but efficient sorting algorithm. Similar to heapsort, we insert all the values we wish to sort into a new ordered data structure—in this case a binary search tree—and then traverse it in order, building our result:

```

def build_binary_tree(values):
    tree = None
    for v in values:
        tree = binary_tree_insert(tree, v)
    return tree

def get_inorder_traversal(root):
    """
    Returns a list containing all the values in the tree, starting at
    *root*.
    Traverses the tree in-order(leftChild, root, rightChild).
    """

```

```
'''
result = []
traverse_binary_tree(root, lambda element: result.append(element))
return result
```

The worst-case time of `build_binary_tree` is  $O(n^2)$ —if you feed it a sorted list of values, it chains them into a linked list with no left subtrees. For example, `build_binary_tree([1, 2, 3, 4, 5])` yields the tree `(1 (2 (3 (4 (5))))))`.

There are several schemes for overcoming this flaw with simple binary trees; the most common is the self-balancing binary search tree. If this same procedure is done using such a tree, the overall worst-case time is  $O(n \log n)$ , which is asymptotically optimal for a comparison sort. In practice, the poor cache performance and added overhead in time and space for a tree-based sort (particularly for node allocation) make it inferior to other asymptotically optimal sorts such as heapsort for static list sorting. On the other hand, it is one of the most efficient methods of *incremental sorting*, adding items to a list over time while keeping the list sorted at all times.

## Types

There are many types of binary search trees. AVL trees and red-black trees are both forms of self-balancing binary search trees. A splay tree is a binary search tree that automatically moves frequently accessed elements nearer to the root. In a treap (*tree heap*), each node also holds a (randomly chosen) priority and the parent node has higher priority than its children. Tango trees are trees optimized for fast searches.

Two other titles describing binary search trees are that of a *complete* and *degenerate* tree.

A complete tree is a tree with  $n$  levels, where for each level  $d \leq n - 1$ , the number of existing nodes at level  $d$  is equal to  $2^d$ . This means all possible nodes exist at these levels. An additional requirement for a complete binary tree is that for the  $n$ th level, while every node does not have to exist, the nodes that do exist must fill from left to right.

A degenerate tree is a tree where for each parent node, there is only one associated child node. What this means is that in a performance measurement, the tree will essentially behave like a linked list data structure.

## Performance comparisons

D. A. Heger (2004)<sup>[2]</sup> presented a performance comparison of binary search trees. Treap was found to have the best average performance, while red-black tree was found to have the smallest amount of performance variations.

## Optimal binary search trees

If we don't plan on modifying a search tree, and we know exactly how often each item will be accessed, we can construct an *optimal binary search tree*, which is a search tree where the average cost of looking up an item (the *expected search cost*) is minimized.

Even if we only have estimates of the search costs, such a system can considerably speed up lookups on average. For example, if you have a BST of English words used in a spell checker, you might balance the tree based on word frequency in text corpora, placing words like *the* near the root and words like *agerasia* near the leaves. Such a tree might be compared with Huffman trees, which similarly seek to place frequently used items near the root in order to produce a dense information encoding; however, Huffman trees only store data elements in leaves and these elements need not be ordered.

If we do not know the sequence in which the elements in the tree will be accessed in advance, we can use splay trees which are asymptotically as good as any static search tree we can construct for any particular sequence of lookup operations.

*Alphabetic trees* are Huffman trees with the additional constraint on order, or, equivalently, search trees with the modification that all elements are stored in the leaves. Faster algorithms exist for *optimal alphabetic binary trees* (OABTs).

Example:

```
procedure Optimum Search Tree(f, f', c):
  for j = 0 to n do
    c[j, j] = 0, F[j, j] = f'j
  for d = 1 to n do
    for i = 0 to (n - d) do
      j = i + d
      F[i, j] = F[i, j - 1] + f' + f'j
      c[i, j] = MIN(i < k <= j) {c[i, k - 1] + c[k, j]} + F[i, j]
```

## References

- [1] Gilberg, R.; Forouzan, B. (2001), "8", *Data Structures: A Pseudocode Approach With C++*, Pacific Grove, CA: Brooks/Cole, p. 339, ISBN 0-534-95216-X
- [2] Heger, Dominique A. (2004), "A Disquisition on The Performance Behavior of Binary Search Tree Data Structures" (<http://www.cepis.org/upgrade/files/full-2004-V.pdf>), *European Journal for the Informatics Professional* **5** (5): 67–75,

## Further reading

- Paul E. Black, Binary Search Tree (<http://www.nist.gov/dads/HTML/binarySearchTree.html>) at the NIST Dictionary of Algorithms and Data Structures.
- Cormen, Thomas H.; Leiserson, Charles E.; Rivest, Ronald L.; Stein, Clifford (2001). "12: Binary search trees, 15.5: Optimal binary search trees". *Introduction to Algorithms* (2nd ed.). MIT Press & McGraw-Hill. pp. 253–272, 356–363. ISBN 0-262-03293-7.
- Jarc, Duane J. (3 December 2005). "Binary Tree Traversals" (<http://nova.umuc.edu/~jarc/idsv/lesson1.html>). *Interactive Data Structure Visualizations*. University of Maryland.
- Knuth, Donald (1997). "6.2.2: Binary Tree Searching". *The Art of Computer Programming. 3: "Sorting and Searching"* (3rd ed.). Addison-Wesley. pp. 426–458. ISBN 0-201-89685-0.
- Long, Sean. "Binary Search Tree" (<http://employees.oneonta.edu/zhangs/PowerPointPlatform/resources/samples/binarysearchtree.ppt>) (PPT). *Data Structures and Algorithms Visualization - A PowerPoint Slides Based Approach*. SUNY Oneonta.
- Parlante, Nick (2001). "Binary Trees" (<http://cslibrary.stanford.edu/110/BinaryTrees.html>). *CS Education Library*. Stanford University.

## External links

- Literate implementations of binary search trees in various languages ([http://en.literateprograms.org/Category:Binary\\_search\\_tree](http://en.literateprograms.org/Category:Binary_search_tree)) on LiteratePrograms
- Goleta, Maksim (27 November 2007). "Goletas.Collections" (<http://goletas.com/csharp-collections/>). *goletas.com*. Includes an iterative C# implementation of AVL trees.
- Jansens, Dana. "Persistent Binary Search Trees" (<http://cg.scs.carleton.ca/~dana/pbst>). Computational Geometry Lab, School of Computer Science, Carleton University. C implementation using GLib.
- Kovac, Kubo. "Binary Search Trees" (<http://people.ksp.sk/~kuko/bak/>) (Java applet). Korešpondenčný seminár z programovania.
- Madru, Justin (18 August 2009). "Binary Search Tree" ([http://jdserver.homelinux.org/wiki/Binary\\_Search\\_Tree](http://jdserver.homelinux.org/wiki/Binary_Search_Tree)). *JDServer*. C++ implementation.

- Tarreau, Willy (2011). "Elastic Binary Trees (ebtree)" (<http://1wt.eu/articles/ebtree/>). *1wt.eu*.
- Binary Search Tree Example in Python (<http://code.activestate.com/recipes/286239/>)
- "References to Pointers (C++)" ([http://msdn.microsoft.com/en-us/library/1sf8shae\(v=vs.80\).aspx](http://msdn.microsoft.com/en-us/library/1sf8shae(v=vs.80).aspx)). *MSDN*. Microsoft. 2005. Gives an example binary tree implementation.
- Igushev, Eduard. "Binary Search Tree C++ implementation" (<http://igushev.com/implementations/binary-search-tree-cpp/>).
- Stromberg, Daniel. "Python Search Tree Empirical Performance Comparison" (<http://stromberg.dnsalias.org/~strombrg/python-tree-and-heap-comparison/>).

## Self-balancing binary search tree

In computer science, a **self-balancing** (or **height-balanced**) **binary search tree** is any node based binary search tree that automatically keeps its height (number of levels below the root) small in the face of arbitrary item insertions and deletions.<sup>[1]</sup>

These structures provide efficient implementations for mutable ordered lists, and can be used for other abstract data structures such as associative arrays, priority queues and sets.

### Overview

Most operations on a binary search tree (BST) take time directly proportional to the height of the tree, so it is desirable to keep the height small. A binary tree with height  $h$  can contain at most  $2^0 + 2^1 + \dots + 2^h = 2^{h+1} - 1$  nodes. It follows that for a tree with  $n$  nodes and height  $h$ :

$$n \leq 2^{h+1} - 1$$

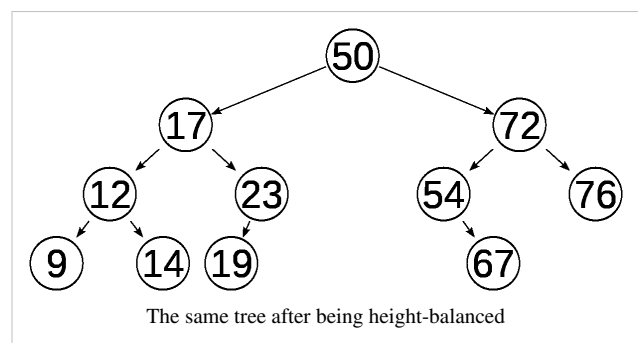
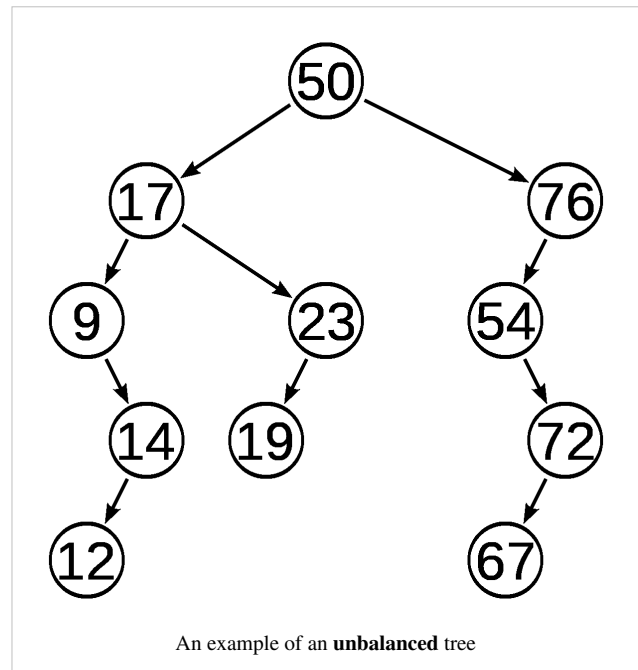
And that implies:

$$h \geq \lceil \log_2(n + 1) - 1 \rceil \geq \lfloor \log_2 n \rfloor.$$

In other words, the minimum height of a tree with  $n$  nodes is  $\log_2(n)$ , rounded down; that is,  $\lfloor \log_2 n \rfloor$ .<sup>[1]</sup>

However, the simplest algorithms for BST item insertion may yield a tree with height  $n$  in rather common situations. For example, when the items are inserted in sorted key order, the tree degenerates into a linked list with  $n$  nodes. The difference in performance between the two situations may be enormous: for  $n = 1,000,000$ , for example, the minimum height is  $\lfloor \log_2(n) \rfloor = 19$ .

If the data items are known ahead of time, the height can be kept small, in the average sense, by adding values in a random order, resulting in a random binary search tree. However, there are many situations (such as online algorithms) where this randomization is not viable.



Self-balancing binary trees solve this problem by performing transformations on the tree (such as tree rotations) at key times, in order to keep the height proportional to  $\log_2(n)$ . Although a certain overhead is involved, it may be justified in the long run by ensuring fast execution of later operations.

Maintaining the height always at its minimum value  $\lfloor \log_2(n) \rfloor$  is not always viable; it can be proven that any insertion algorithm which did so would have an excessive overhead. Therefore, most self-balanced BST algorithms keep the height within a constant factor of this lower bound.

In the asymptotic ("Big-O") sense, a self-balancing BST structure containing  $n$  items allows the lookup, insertion, and removal of an item in  $O(\log n)$  worst-case time, and ordered enumeration of all items in  $O(n)$  time. For some implementations these are per-operation time bounds, while for others they are amortized bounds over a sequence of operations. These times are asymptotically optimal among all data structures that manipulate the key only through comparisons.

## Implementations

Popular data structures implementing this type of tree include:

- AA tree
- AVL tree
- Red-black tree
- Scapegoat tree
- Splay tree
- Treap

## Applications

Self-balancing binary search trees can be used in a natural way to construct and maintain ordered lists, such as priority queues. They can also be used for associative arrays; key-value pairs are simply inserted with an ordering based on the key alone. In this capacity, self-balancing BSTs have a number of advantages and disadvantages over their main competitor, hash tables. One advantage of self-balancing BSTs is that they allow fast (indeed, asymptotically optimal) enumeration of the items *in key order*, which hash tables do not provide. One disadvantage is that their lookup algorithms get more complicated when there may be multiple items with the same key. Self-balancing BSTs have better worst-case lookup performance than hash tables ( $O(\log n)$  compared to  $O(n)$ ), but have worse average-case performance ( $O(\log n)$  compared to  $O(1)$ ).

Self-balancing BSTs can be used to implement any algorithm that requires mutable ordered lists, to achieve optimal worst-case asymptotic performance. For example, if binary tree sort is implemented with a self-balanced BST, we have a very simple-to-describe yet asymptotically optimal  $O(n \log n)$  sorting algorithm. Similarly, many algorithms in computational geometry exploit variations on self-balancing BSTs to solve problems such as the line segment intersection problem and the point location problem efficiently. (For average-case performance, however, self-balanced BSTs may be less efficient than other solutions. Binary tree sort, in particular, is likely to be slower than merge sort, quicksort, or heapsort, because of the tree-balancing overhead as well as cache access patterns.)

Self-balancing BSTs are flexible data structures, in that it's easy to extend them to efficiently record additional information or perform new operations. For example, one can record the number of nodes in each subtree having a certain property, allowing one to count the number of nodes in a certain key range with that property in  $O(\log n)$  time. These extensions can be used, for example, to optimize database queries or other list-processing algorithms.



## References

- [1] Donald Knuth. *The Art of Computer Programming*, Volume 3: *Sorting and Searching*, Second Edition. Addison-Wesley, 1998. ISBN 0-201-89685-0. Section 6.2.3: Balanced Trees, pp.458–481.

## External links

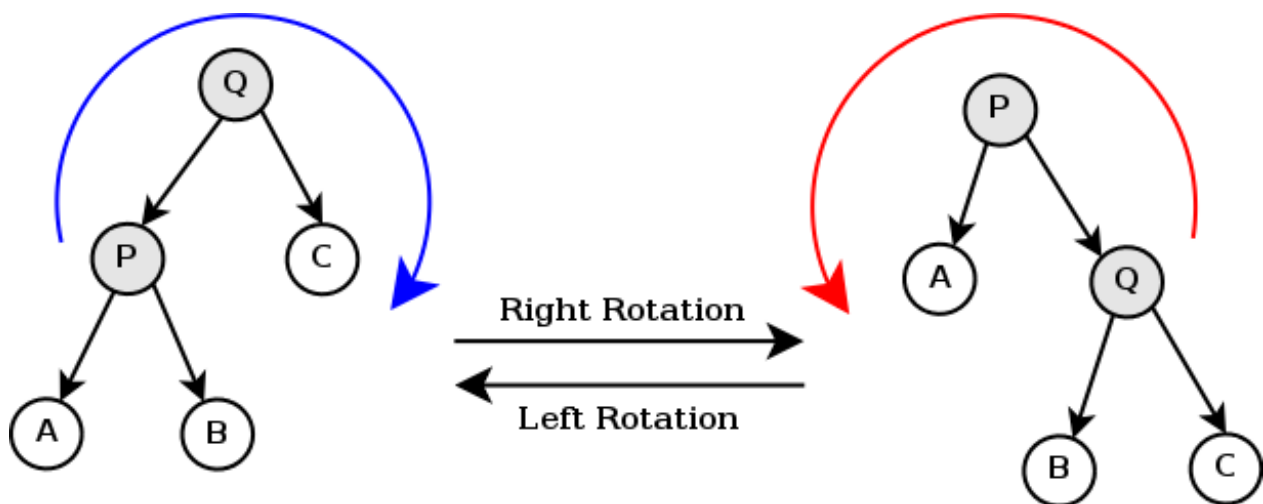
- Dictionary of Algorithms and Data Structures: Height-balanced binary search tree (<http://www.nist.gov/dads/HTML/heightBalancedTree.html>)
- GNU libavl (<http://adtnfo.org/>), a LGPL-licensed library of binary tree implementations in C, with documentation

# Tree rotation

In discrete mathematics, **tree rotation** is an operation on a binary tree that changes the structure without interfering with the order of the elements. A tree rotation moves one node up in the tree and one node down. It is used to change the shape of the tree, and in particular to decrease its height by moving smaller subtrees down and larger subtrees up, resulting in improved performance of many tree operations.

There exists an inconsistency in different descriptions as to the definition of the **direction of rotations**. Some say that the direction of a rotation depends on the side which the tree nodes are shifted upon whilst others say that it depends on which child takes the root's place (opposite of the former). This article takes the approach of the side where the nodes get shifted to.

## Illustration



The right rotation operation as shown in the image above is performed with  $Q$  as the root and hence is a right rotation on, or rooted at,  $Q$ . This operation results in a rotation of the tree in the clockwise direction. The inverse operation is the left rotation, which results in a movement in a counter-clockwise direction (the left rotation shown above is rooted at  $P$ ). The key to understanding how a rotation functions is to understand its constraints. In particular the order of the leaves of the tree (when read left to right for example) cannot change (another way to think of it is that the order that the leaves would be visited in a depth first search must be the same after the operation as before). Another constraint is the main property of a binary search tree, namely that the right child is greater than the parent and the left child is lesser than the parent. Notice that the right child of a left child of the root of a sub-tree (for example node  $B$  in the diagram for the tree rooted at  $Q$ ) can become the left child of the root, that itself becomes the right child of the "new" root in the rotated sub-tree, without violating either of those constraints. As you can see in

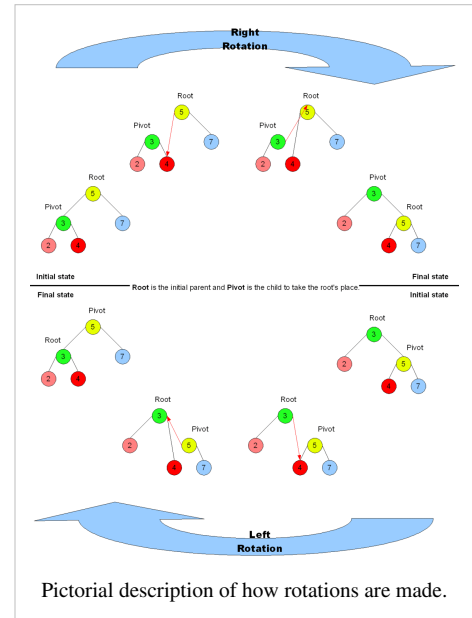
the diagram, the order of the leaves doesn't change. The opposite operation also preserves the order and is the second kind of rotation.

Assuming this is a binary search tree, as stated above, the elements must be interpreted as variables that can be compared to each other. The alphabetic characters above are used as placeholders for these variables.

## Detailed illustration

When a subtree is rotated, the subtree side upon which it is rotated decreases its height by one node while the other subtree increases its height. This makes tree rotations useful for rebalancing a tree.

Using the terminology of **Root** for the parent node of the subtrees to rotate, **Pivot** for the node which will become the new parent node, **RS** for rotation side upon to rotate and **OS** for opposite side of rotation. In the above diagram for the root Q, the **RS** is C and the **OS** is P. The pseudo code for the rotation is:



```
Pivot = Root.OS
Root.OS = Pivot.RS
Pivot.RS = Root
Root = Pivot
```

This is a constant time operation.

The programmer must also make sure that the root's parent points to the pivot after the rotation. Also, the programmer should note that this operation may result in a new root for the entire tree and take care to update pointers accordingly.

## Inorder Invariance

The tree rotation renders the inorder traversal of the binary tree invariant. This implies the order of the elements are not affected when a rotation is performed in any part of the tree. Here are the inorder traversals of the trees shown above:

```
Left tree: ((A, P, B), Q, C)           Right tree: (A, P, (B, Q, C))
```

Computing one from the other is very simple. The following is example Python code that performs that computation:

```
def right_rotation(treenode):
    left, Q, C = treenode
    A, P, B = left
    return (A, P, (B, Q, C))
```

Another way of looking at it is:

**Right Rotation of node Q:**

```

Let P be Q's left child.
Set P to be the new root.
Set Q's left child to be P's right child.
Set P's right child to be Q.

```

**Left Rotation of node P:**

```

Let Q be P's right child.
Set Q to be the new root.
Set P's right child to be Q's left child.
Set Q's left child to be P.

```

All other connections are left as-is.

There are also *double rotations*, which are combinations of left and right rotations. A *double left* rotation at X can be defined to be a right rotation at the right child of X followed by a left rotation at X; similarly, a *double right* rotation at X can be defined to be a left rotation at the left child of X followed by a right rotation at X.

Tree rotations are used in a number of tree data structures such as AVL trees, red-black trees, splay trees, and treaps. They require only constant time because they are *local* transformations: they only operate on 5 nodes, and need not examine the rest of the tree.

**Rotations for rebalancing**

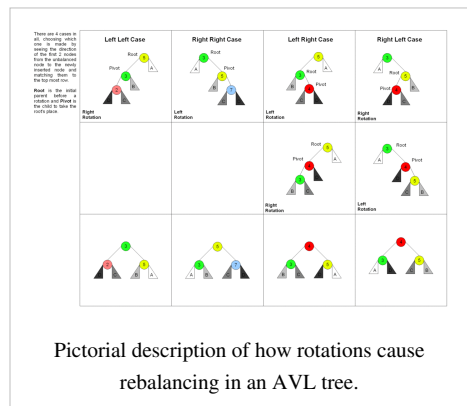
A tree can be rebalanced using rotations. After a rotation, the side of the rotation increases its height by 1 whilst the side opposite the rotation decreases its height similarly. Therefore, one can strategically apply rotations to nodes whose left child and right child differ in height by more than 1. Self-balancing binary search trees apply this operation automatically. A type of tree which uses this rebalancing technique is the AVL tree.

**Rotation distance**

The **rotation distance** between any two binary trees with the same number of nodes is the minimum number of rotations needed to transform one into the other. With this distance, the set of  $n$ -node binary trees becomes a metric space: the distance is symmetric, positive when given two different trees, and satisfies the triangle inequality.

It is an open problem whether there exists a polynomial time algorithm for calculating rotation distance.

Daniel Sleator, Robert Tarjan and William Thurston showed that the rotation distance between any two  $n$ -node trees (for  $n \geq 11$ ) is at most  $2n - 6$ , and that infinitely many pairs of trees are this far apart.<sup>[1]</sup>



## References

- [1] Sleator, Daniel D.; Tarjan, Robert E.; Thurston, William P. (1988), "Rotation distance, triangulations, and hyperbolic geometry", *Journal of the American Mathematical Society* (American Mathematical Society) **1** (3): 647–681, doi:10.2307/1990951, JSTOR 1990951, MR928904.

## External links

- Java applets demonstrating tree rotations (<http://www.cs.queensu.ca/home/jstewart/applets/bst/bst-rotation.html>)
- The AVL Tree Rotations Tutorial (<http://fortheloot.com/public/AVLTreeTutorial.rtf>) (RTF) by John Hargrove

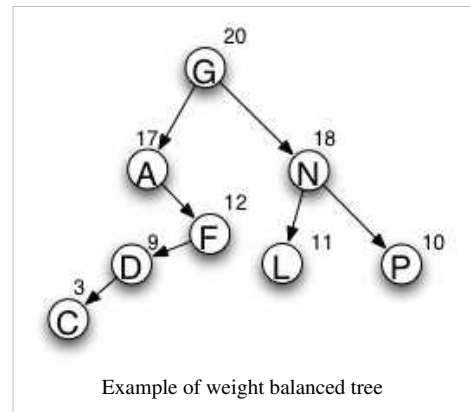
# Weight-balanced tree

A **weight-balanced binary tree** is a binary tree which is balanced based on knowledge of the probabilities of searching for each individual node. Within each subtree, the node with the highest weight appears at the root. This can result in more efficient searching performance.

Construction of such a tree is similar to that of a Treap, but node weights are chosen randomly in the latter.

## The diagram

In the diagram to the right, the letters represent node *values* and the numbers represent node *weights*. Values are used to order the tree, as in a general binary search tree. The weight may be thought of as a probability or activity count associated with the node. In the diagram, the root is G because its weight is the greatest in the tree. The left subtree begins with A because, out of all nodes with values that come before G, A has the highest weight. Similarly, N is the highest-weighted node that comes after G.



## Timing analysis

A weight balanced tree gives close to optimal values for the expected length of successful search calculations. From the above example we get

$$\text{ELOSS} = \text{depth}(\text{node A}) * \text{probability}(\text{node A}) + \text{depth}(\text{node C}) * \text{probability}(\text{node C}) + \dots$$

$$\text{ELOSS} = 2 * 0.17 + 5 * 0.03 + 4 * 0.09 + 3 * 0.12 + 1 * 0.20 + 3 * 0.11 + 3 * 0.10 + 2 * 0.18$$

$$\text{ELOSS} = 2.4$$

This is the expected number of nodes that will be examined before finding the desired node.

## References

- Jean-Paul Tremblay and Grant A. Cheston. *Data Structures and Software Development in an object-oriented domain*, Eiffel Edition. Prentice Hall, 2001. ISBN 0-13-787946-6.

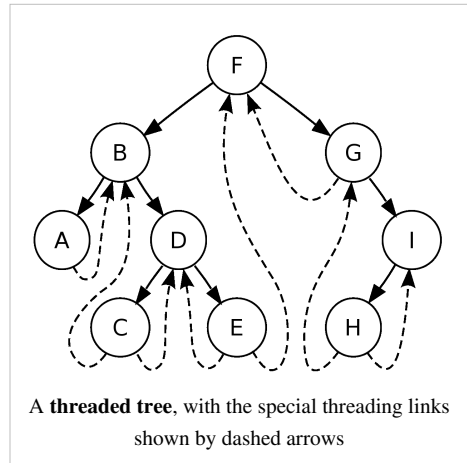
# Threaded binary tree

A **threaded binary tree** defined as follows:

"A binary tree is *threaded* by making all right child pointers that would normally be null point to the inorder successor of the node, and all left child pointers that would normally be null point to the inorder predecessor of the node."<sup>[1]</sup>

A threaded binary tree makes it possible to traverse the values in the binary tree via a linear traversal that is more rapid than a recursive in-order traversal. It is also possible to discover the parent of a node from a threaded binary tree, without explicit use of parent pointers or a stack, albeit slowly. This can be useful where stack space is limited, or where a stack of parent pointers is unavailable (for finding the parent pointer via DFS).

To see how this is possible, consider a node  $k$  that has a right child  $r$ . Then the left pointer of  $r$  must be either a child or a thread back to  $k$ . In the case that  $r$  has a left child, that left child must in turn have either a left child of its own or a thread back to  $k$ , and so on for all successive left children. So by following the chain of left pointers from  $r$ , we will eventually find a thread pointing back to  $k$ . The situation is symmetrically similar when  $q$  is the left child of  $p$ —we can follow  $q$ 's right children to a thread pointing ahead to  $p$ .



## Types of binary trees

1. Single Threaded: each node is threaded towards **either** the inorder predecessor **or** successor.
2. Double threaded: each node is threaded towards **both** the inorder predecessor **and** successor.

In Python:

```
def parent (node) :
    if node is node.tree.root:
        return None
    else:
        x = node
        y = node
        while True:
            if is_thread(y.right):
                p = y.right
                if p is None or p.left is not node:
                    p = x
                    while not is_thread(p.left):
                        p = p.left
                    p = p.left
                return p
            elif is_thread(x.left):
                p = x.left
                if p is None or p.right is not node:
                    p = y
```

```

while not is_thread(p.right):
    p = p.right
    p = p.right
return p
x = x.left
y = y.right

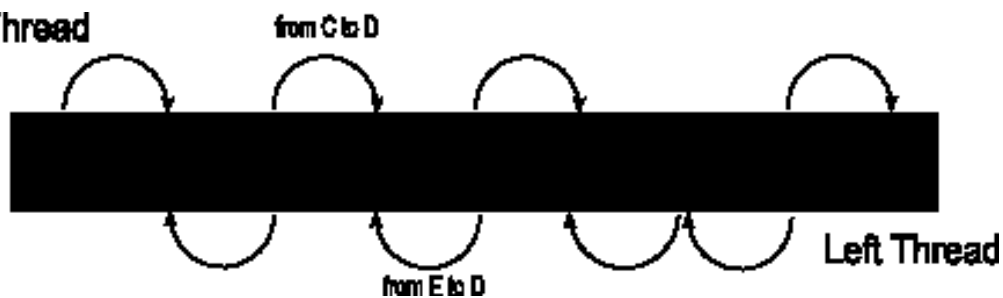
```

## The array of Inorder traversal

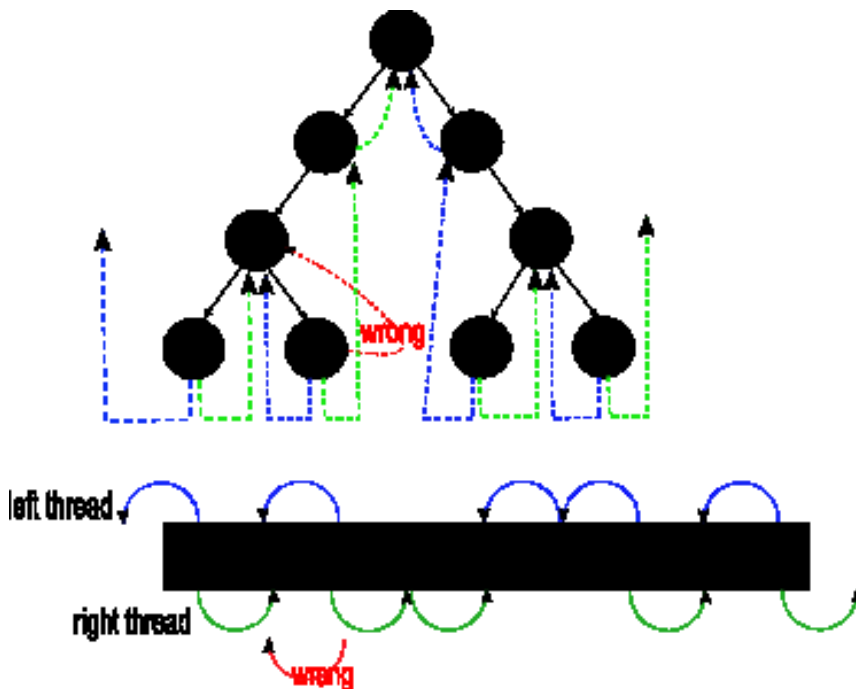
Threads are reference to the predecessors and successors of the node according to an inorder traversal.

Inorder of the threaded tree is ABCDEFGHI, the predecessor of E is D, the successor of E is F.

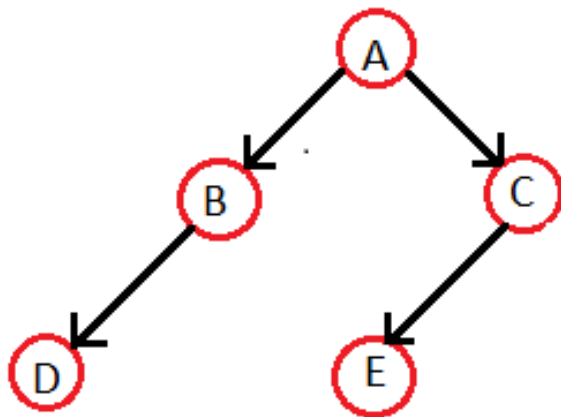
### Right Thread



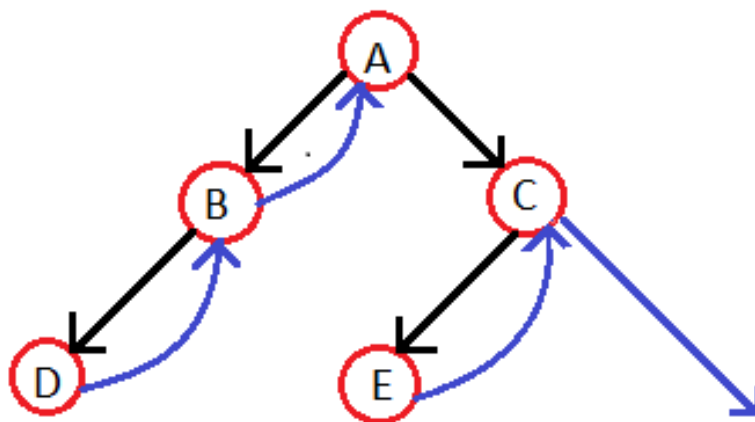
### Example



Let's make the Threaded Binary tree out of a normal binary tree...



The INORDER traversal for the above tree is—D B A E C. So, the respective Threaded Binary tree will be --

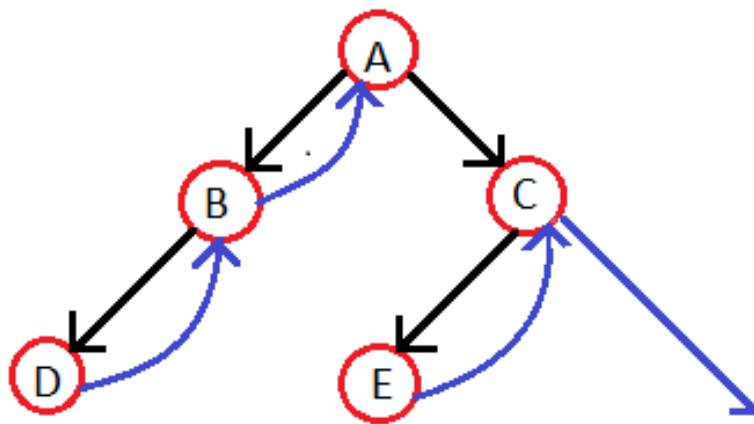


## Null link

An m-way threaded binary tree, there are  $n*m - (n-1)$  links are void in a tree with n nodes.

## Non recursive Inorder traversal for a Threaded Binary Tree

As this is a non-recursive method for traversal, it has to be an iterative procedure; meaning, all the steps for the traversal of a node have to be under a loop so that the same can be applied to all the nodes in the tree. I'll consider the INORDER traversal again. Here, for every node, we'll visit the left sub-tree (if it exists) first (if and only if we haven't visited it earlier); then we visit (i.e. print its value, in our case) the node itself and then the right sub-tree (if it exists). If the right sub-tree is not there, we check for the threaded link and make the threaded node the current node in consideration. Please, follow the example given below.



## Algorithm

Step-1: For the current node check whether it has a left child which is not there in the visited list. If it has then go to step-2 or else step-3.

Step-2: Put that left child in the list of visited nodes and make it your current node in consideration. Go to step-6.

Step-3: For the current node check whether it has a right child. If it has then go to step-4 else go to step-5

Step-4: Make that right child as your current node in consideration. Go to step-6.

Step-5: Check for the threaded node and if its there make it your current node.

Step-6: Go to step-1 if all the nodes are not over otherwise quit



		<b>List of visited nodes</b>	<b>Inorder</b>
step-1	'A' has a left child i.e. B, which has not been visited. So, we put B in our "list of visited nodes" and B becomes our current node in consideration.	B	
step-2	'B' also has a left child, 'D', which is not there in our list of visited nodes. So, we put 'D' in that list and make it our current node in consideration.	B D	
step-3	'D' has no left child, so we print 'D'. Then we check for its right child. 'D' has no right child and thus we check for its thread-link. It has a thread going till node 'B'. So, we make 'B' as our current node in consideration.	B D	D
step-4	'B' certainly has a left child but its already in our list of visited nodes. So, we print 'B'. Then we check for its right child but it doesn't exist. So, we make its threaded node (i.e. 'A') as our current node in consideration.	B D	D B
step-5	'A' has a left child, 'B', but its already there in the list of visited nodes. So, we print 'A'. Then we check for its right child. 'A' has a right child, 'C' and it's not there in our list of visited nodes. So, we add it to that list and we make it our current node in consideration.	B D C	D B A
step-6	'C' has 'E' as the left child and it's not there in our list of visited nodes even. So, we add it to that list and make it our current node in consideration.	B D C E	D B A
step-7		and finally.....	D B A E C

## References

[1] Van Wyk, Christopher J. Data Structures and C Programs, Addison-Wesley, 1988, p. 175. ISBN 978-0-201-16116-8.

## External links

- Tutorial on threaded binary trees ([http://www.etrnallyconfuzzled.com/tuts/datastructures/jsw\\_tut\\_bst1.aspx#thread](http://www.etrnallyconfuzzled.com/tuts/datastructures/jsw_tut_bst1.aspx#thread))
- GNU libavl 2.0.2, Section on threaded binary search trees (<http://www.stanford.edu/~blp/avl/libavl.html/Threaded-Binary-Search-Trees.html>)

# AVL tree

---

AVL tree		
Type	Tree	
Invented	1962	
Invented by	G. M. Adelson-Velskii and E. M. Landis	
Time complexity in big O notation		
	Average	Worst case
Space	O(n)	O(n)
Search	O(log n)	O(log n)
Insert	O(log n)	O(log n)
Delete	O(log n)	O(log n)

In computer science, an **AVL tree** is a self-balancing binary search tree, and it was the first such data structure to be invented.<sup>[1]</sup> In an AVL tree, the heights of the two child subtrees of any node differ by at most one. Lookup, insertion, and deletion all take  $O(\log n)$  time in both the average and worst cases, where  $n$  is the number of nodes in the tree prior to the operation. Insertions and deletions may require the tree to be rebalanced by one or more tree rotations. In other words AVL tree is a binary search tree where the height of the left subtree differs from the height of the right subtree by at most 1 level, if it exceeds 1 level then rebalancing occurs.

The AVL tree is named after its two Soviet inventors, G. M. Adelson-Velskii and E. M. Landis, who published it in their 1962 paper "An algorithm for the organization of information."<sup>[2]</sup>

AVL trees are often compared with red-black trees because they support the same set of operations and because red-black trees also take  $O(\log n)$  time for the basic operations. Because AVL trees are more rigidly balanced, they are faster than red-black trees for lookup intensive applications.<sup>[3]</sup> Similar to red-black trees, AVL trees are in general not weight-balanced<sup>[4]</sup>, that is sibling nodes can have hugely differing numbers of descendants.

## Operations

Basic operations of an AVL tree involve carrying out the same actions as would be carried out on an unbalanced binary search tree, but modifications are preceded or followed by one or more operations called tree rotations, which help to restore the height balance of the subtrees.

## Searching

Lookup in an AVL tree is performed exactly like in any unbalanced binary search tree. Because of the height-balancing of the tree, a lookup takes  $O(\log n)$  time. No special actions need to be taken, and the tree's structure is not modified by lookups. (This is in contrast to splay tree lookups, which do modify their tree's structure.)

If each node additionally records the size of its subtree (including itself and its descendants), then the nodes can be retrieved by index in  $O(\log n)$  time as well.

Once a node has been found in a balanced tree, the *next* or *previous* nodes can be explored in amortized constant time. Some instances of exploring these "nearby" nodes require traversing up to  $2 \times \log(n)$  links (particularly when moving from the rightmost leaf of the root's left subtree to the leftmost leaf of the root's right subtree). However, exploring all  $n$  nodes of the tree in this manner would use each link exactly twice: one traversal to enter the subtree

---

rooted at that node, and another to leave that node's subtree after having explored it. And since there are  $n-1$  links in any tree, the amortized cost is found to be  $2 \times (n-1)/n$ , or approximately 2.

## Insertion

After inserting a node, it is necessary to check each of the node's ancestors for consistency with the rules of AVL. The balance factor is calculated as follows:  $\text{balanceFactor} = \text{height}(\text{left-subtree}) - \text{height}(\text{right-subtree})$ . For each node checked, if the balance factor remains  $-1$ ,  $0$ , or  $+1$  then no rotations are necessary. However, if balance factor becomes less than  $-1$  or greater than  $+1$ , the subtree rooted at this node is unbalanced. If insertions are performed serially, after each insertion, at most one of the following cases needs to be resolved to restore the entire tree to the rules of AVL.

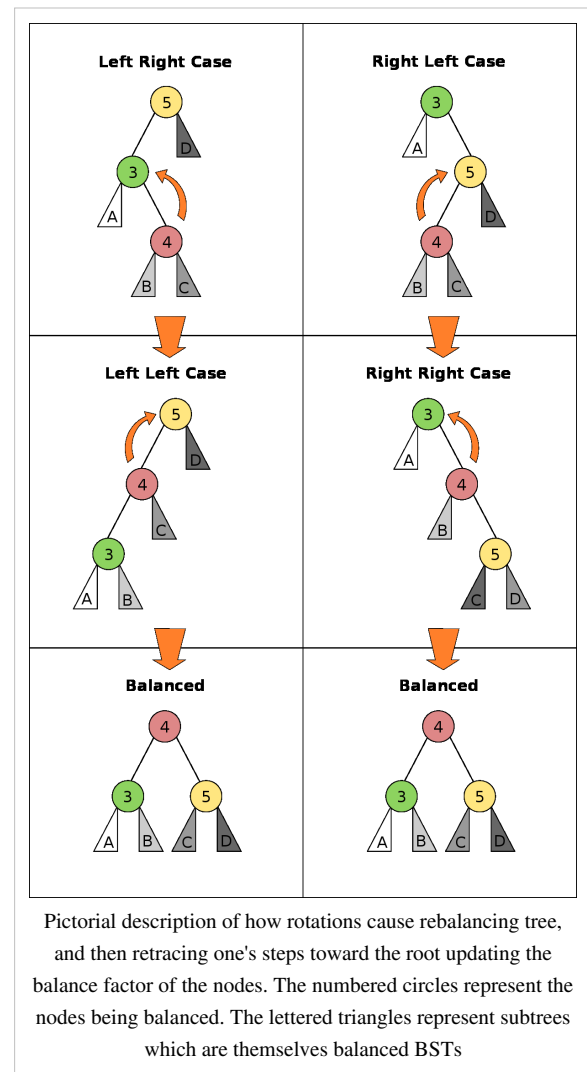
There are four cases which need to be considered, of which two are symmetric to the other two. Let  $P$  be the root of the unbalanced subtree, with  $R$  and  $L$  denoting the right and left children of  $P$  respectively.

### Right-Right case and Right-Left case:

- If the balance factor of  $P$  is  $-2$  then the right subtree outweighs the left subtree of the given node, and the balance factor of the right child ( $R$ ) must be checked. The left rotation with  $P$  as the root is necessary.
- If the balance factor of  $R$  is  $-1$ , a **single left rotation** (with  $P$  as the root) is needed (Right-Right case).
- If the balance factor of  $R$  is  $+1$ , two different rotations are needed. The first rotation is a **right rotation** with  $R$  as the root. The second is a **left rotation** with  $P$  as the root (Right-Left case).

### Left-Left case and Left-Right case:

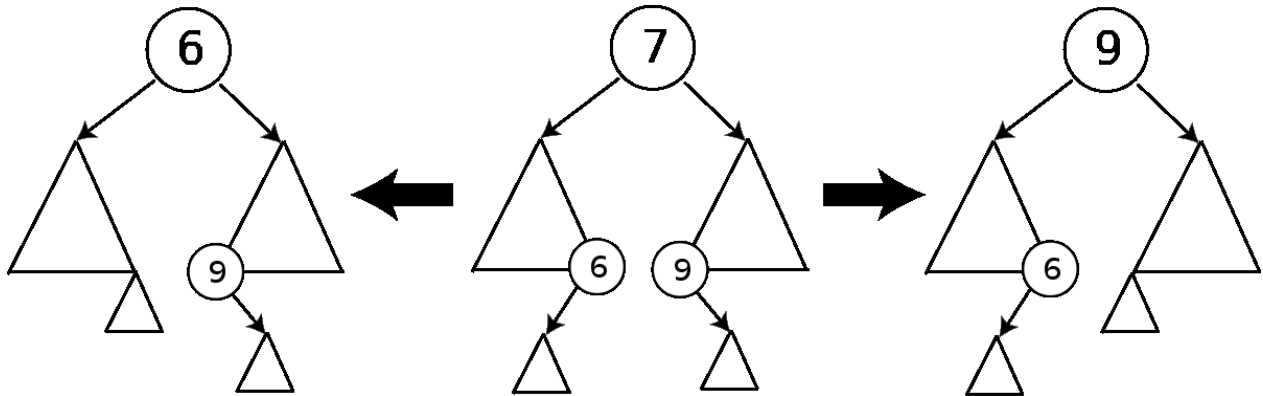
- If the balance factor of  $P$  is  $2$ , then the left subtree outweighs the right subtree of the given node, and the balance factor of the left child ( $L$ ) must be checked. The right rotation with  $P$  as the root is necessary.
- If the balance factor of  $L$  is  $+1$ , a **single right rotation** (with  $P$  as the root) is needed (Left-Left case).
- If the balance factor of  $L$  is  $-1$ , two different rotations are needed. The first rotation is a **left rotation** with  $L$  as the root. The second is a **right rotation** with  $P$  as the root (Left-Right case).



## Deletion

If the node is a leaf or has only one child, remove it. Otherwise, replace it with either the largest in its left sub tree (in order predecessor) or the smallest in its right sub tree (in order successor), and remove that node. The node that was found as a replacement has at most one sub tree. After deletion, retrace the path back up the tree (parent of the replacement) to the root, adjusting the balance factors as needed.

As with all binary trees, a node's in-order successor is the left-most child of its right subtree, and a node's in-order predecessor is the right-most child of its left subtree. In either case, this node will have zero or one children. Delete it according to one of the two simpler cases above.



In addition to the balancing described above for insertions, if the balance factor for the tree is 2 and that of the left subtree is 0, a right rotation must be performed on P. The mirror of this case is also necessary.

The retracing can stop if the balance factor becomes  $-1$  or  $+1$  indicating that the height of that subtree has remained unchanged. If the balance factor becomes 0 then the height of the subtree has decreased by one and the retracing needs to continue. If the balance factor becomes  $-2$  or  $+2$  then the subtree is unbalanced and needs to be rotated to fix it. If the rotation leaves the subtree's balance factor at 0 then the retracing towards the root must continue since the height of this subtree has decreased by one. This is in contrast to an insertion where a rotation resulting in a balance factor of 0 indicated that the subtree's height has remained unchanged.

The time required is  $O(\log n)$  for lookup, plus a maximum of  $O(\log n)$  rotations on the way back to the root, so the operation can be completed in  $O(\log n)$  time.

## Comparison to other structures

Both AVL trees and red-black trees are self-balancing binary search trees, so they are very similar mathematically. The operations to balance the trees are different, but both occur in  $O(\log n)$  time. The real difference between the two is the limiting height. For a tree of size  $n$ :

- An AVL tree's height is strictly less than:<sup>[5][6]</sup>

$$\log_{\varphi}(\sqrt{5}(n+2))-2 = \frac{\log_2(\sqrt{5}(n+2))}{\log_2(\varphi)}-2 = \log_{\varphi}(2) \cdot \log_2(\sqrt{5}(n+2))-2 \approx 1.44 \log_2(n+2)-0.328$$

where  $\varphi$  is the golden ratio.

- A red-black tree's height is at most  $2 \log_2(n+1)$ <sup>[7]</sup>

AVL trees are more rigidly balanced than red-black trees, leading to slower insertion and removal but faster retrieval.

## References

- [1] Robert Sedgewick, *Algorithms*, Addison-Wesley, 1983, ISBN 0-201-06672-6, page 199, chapter 15: Balanced Trees.
- [2] Adelson-Velskii, G.; E. M. Landis (1962). "An algorithm for the organization of information". *Proceedings of the USSR Academy of Sciences* **146**: 263–266. **(Russian)** English translation by Myron J. Ricci in *Soviet Math. Doklady*, 3:1259–1263, 1962.
- [3] Pfaff, Ben (June 2004). "Performance Analysis of BSTs in System Software" (<http://www.stanford.edu/~blp/papers/libavl.pdf>) (PDF). Stanford University. .
- [4] AVL trees are not weight-balanced? (<http://cs.stackexchange.com/questions/421/avl-trees-are-not-weight-balanced>)
- [5] Burkhard, Walt (Spring 2012). "AVL Dictionary Data Type Implementation" (<http://ieng6.ucsd.edu/~cs100s/public/Notes/CS100s12.pdf>). *Advanced Data Structures*. La Jolla: A.S. Soft Reserves (<http://softreserves.ucsd.edu/>), UC San Diego. p. 103. .
- [6] Knuth, Donald E. (2000). *Sorting and searching* (2. ed., 6. printing, newly updated and rev. ed.). Boston [u.a.]: Addison-Wesley. pp. 460. ISBN 0-201-89685-0.
- [7] Proof of asymptotic bounds

## Further reading

- Donald Knuth. *The Art of Computer Programming*, Volume 3: *Sorting and Searching*, Third Edition. Addison-Wesley, 1997. ISBN 0-201-89685-0. Pages 458–475 of section 6.2.3: Balanced Trees.

## External links

- xdg library (<https://github.com/vilkov/libxdg/wiki>) by Dmitriy Vilkov: Serializable straight C-implementation could easily be taken from this library under GNU-LGPL and AFL v2.0 licenses.
- Description from the Dictionary of Algorithms and Data Structures (<http://www.nist.gov/dads/HTML/avltree.html>)
- Python Implementation (<http://github.com/pgrafv/python-avl-tree/>)
- Single C header file by Ian Piumarta (<http://piumarta.com/software/tree/>)
- AVL Tree Demonstration ([http://www.strille.net/works/media\\_technology\\_projects/avl-tree\\_2001/](http://www.strille.net/works/media_technology_projects/avl-tree_2001/))
- AVL tree applet – all the operations (<http://webdiis.unizar.es/asignaturas/EDA/AVLTree/avltree.html>)
- Fast and efficient implementation of AVL Trees (<http://github.com/fbuihuu/libtree>)
- PHP Implementation (<https://github.com/mondrake/Rbppavl>)
- C++ implementation which can be used as an array (<http://www.codeproject.com/Articles/12347/AVL-Binary-Tree-for-C>)
- Self balancing AVL tree with Concat and Split operations (<http://code.google.com/p/self-balancing-avl-tree/>)

# Red-black tree

---

Red–black tree		
Type	Tree	
Invented	1972	
Invented by	Rudolf Bayer	
Time complexity in big O notation		
	Average	Worst case
Space	O(n)	O(n)
Search	O(log n)	O(log n)
Insert	O(log n)	O(log n)
Delete	O(log n)	O(log n)

A **red–black tree** is a type of self-balancing binary search tree, a data structure used in computer science, typically to implement associative arrays.

Since it is a balanced tree, it guarantees insertion, search and delete to be  $O(\log n)$  in time, where  $n$  is the total number of elements in the tree.<sup>[1]</sup>

## History

The original structure was invented in 1972 by Rudolf Bayer<sup>[2]</sup> and named "symmetric binary B-tree," but acquired its modern name in a paper in 1978 by Leonidas J. Guibas and Robert Sedgwick.<sup>[3]</sup>

## Terminology

A red–black tree is a special type of binary tree, used in computer science to organize pieces of comparable data, such as text fragments or numbers.

The leaf nodes of red–black trees do not contain data. These leaves need not be explicit in computer memory—a null child pointer can encode the fact that this child is a leaf—but it simplifies some algorithms for operating on red–black trees if the leaves really are explicit nodes. To save memory, sometimes a single sentinel node performs the role of all leaf nodes; all references from internal nodes to leaf nodes then point to the sentinel node.

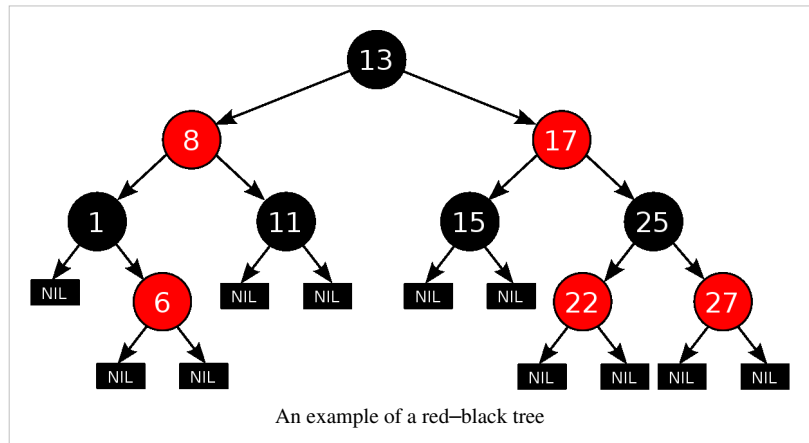
Red–black trees, like all binary search trees, allow efficient in-order traversal (that is: in the order Left–Root–Right) of their elements. The search-time results from the traversal from root to leaf, and therefore a balanced tree, having the least possible tree height, results in  $O(\log n)$  search time.

---

## Properties

In addition to the requirements imposed on a binary search trees, with red–black trees:

1. A node is either red or black.
2. The root is black. (This rule is sometimes omitted. Since the root can always be changed from red to black, but not necessarily vice-versa, this rule has little effect on analysis.)
3. All leaves (NIL) are black. (All leaves are same color as the root.)
4. Both children of every red node are black.
5. Every simple path from a given node to any of its descendant leaves contains the same number of black nodes.



These constraints enforce a critical property of red–black trees: that the path from the root to the furthest leaf is no more than twice as long as the path from the root to the nearest leaf. The result is that the tree is roughly height-balanced. Since operations such as inserting, deleting, and finding values require worst-case time proportional to the height of the tree, this theoretical upper bound on the height allows red–black trees to be efficient in the worst case, unlike ordinary binary search trees. Red-black trees are in general not weight-balanced<sup>[4]</sup>, that is sibling nodes can have hugely differing numbers of descendants.

To see why this is guaranteed, it suffices to consider the effect of properties 4 and 5 together. For a red–black tree  $T$ , let  $B$  be the number of black nodes in property 5. Therefore the shortest possible path from the root of  $T$  to any leaf consists of  $B$  black nodes. Longer possible paths may be constructed by inserting red nodes. However, property 4 makes it impossible to insert more than one consecutive red node. Therefore the longest possible path consists of  $2B$  nodes, alternating black and red.

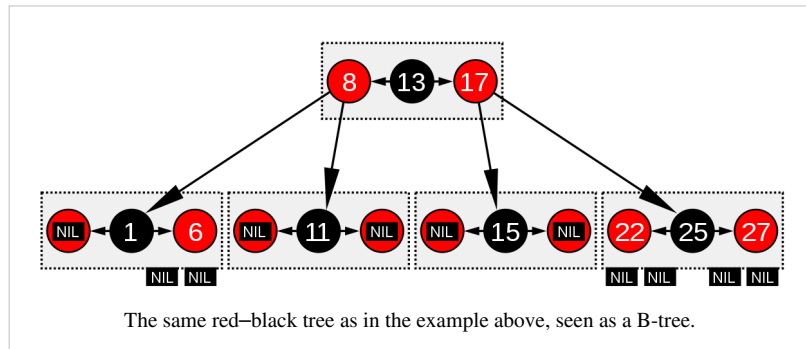
The shortest possible path has all black nodes, and the longest possible path alternates between red and black nodes. Since all maximal paths have the same number of black nodes, by property 5, this shows that no path is more than twice as long as any other path.

In many of the presentations of tree data structures, it is possible for a node to have only one child, and leaf nodes contain data. It is possible to present red–black trees in this paradigm, but it changes several of the properties and complicates the algorithms. For this reason, this article uses "null leaves", which contain no data and merely serve to indicate where the tree ends, as shown above. These nodes are often omitted in drawings, resulting in a tree that seems to contradict the above principles, but in fact does not. A consequence of this is that all internal (non-leaf) nodes have two children, although one or both of those children may be null leaves. Property 5 ensures that a red node must have either two black null leaves or two black non-leaves as children. For a black node with one null leaf child and one non-null-leaf child, properties 3, 4 and 5 ensure that the non-null-leaf child must be a red node with two black null leaves as children.

Some explain a red–black tree as a binary search tree whose edges, instead of nodes, are colored in red or black, but this does not make any difference. The color of a node in this article's terminology corresponds to the color of the edge connecting the node to its parent, except that the root node is always black (property 2) whereas the corresponding edge does not exist.

## Analogy to B-trees of order 4

A red-black tree is similar in structure to a B-tree of order 4, where each node can contain between 1 to 3 values and (accordingly) between 2 to 4 child pointers. In such B-tree, each node will contain only one value matching the value in a black node of the red-black tree, with an optional value before and/or after it in the same node, both matching an equivalent red node of the red-black tree.



One way to see this equivalence is to "move up" the red nodes in a graphical representation of the red-black tree, so that they align horizontally with their parent black node, by creating together a horizontal cluster. In the B-tree, or in the modified graphical representation of the red-black tree, all leaf nodes are at the same depth.

The red-black tree is then structurally equivalent to a B-tree of order 4, with a minimum fill factor of 33% of values per cluster with a maximum capacity of 3 values.

This B-tree type is still more general than a red-black tree though, as it allows ambiguity in a red-black tree conversion—multiple red-black trees can be produced from an equivalent B-tree of order 4. If a B-tree cluster contains only 1 value, it is the minimum, black, and has two child pointers. If a cluster contains 3 values, then the central value will be black and each value stored on its sides will be red. If the cluster contains two values, however, either one can become the black node in the red-black tree (and the other one will be red).

So the order-4 B-tree does not maintain which of the values contained in each cluster is the root black tree for the whole cluster and the parent of the other values in the same cluster. Despite this, the operations on red-black trees are more economical in time because you don't have to maintain the vector of values. It may be costly if values are stored directly in each node rather than being stored by reference. B-tree nodes, however, are more economical in space because you don't need to store the color attribute for each node. Instead, you have to know which slot in the cluster vector is used. If values are stored by reference, e.g. objects, null references can be used and so the cluster can be represented by a vector containing 3 slots for value pointers plus 4 slots for child references in the tree. In that case, the B-tree can be more compact in memory, improving data locality.

The same analogy can be made with B-trees with larger orders that can be structurally equivalent to a colored binary tree: you just need more colors. Suppose that you add blue, then the blue-red-black tree defined like red-black trees but with the additional constraint that no two successive nodes in the hierarchy will be blue and all blue nodes will be children of a red node, then it becomes equivalent to a B-tree whose clusters will have at most 7 values in the following colors: blue, red, blue, black, blue, red, blue (For each cluster, there will be at most 1 black node, 2 red nodes, and 4 blue nodes).

For moderate volumes of values, insertions and deletions in a colored binary tree are faster compared to B-trees because colored trees don't attempt to maximize the fill factor of each horizontal cluster of nodes (only the minimum fill factor is guaranteed in colored binary trees, limiting the number of splits or junctions of clusters). B-trees will be faster for performing rotations (because rotations will frequently occur within the same cluster rather than with multiple separate nodes in a colored binary tree). However for storing large volumes, B-trees will be much faster as they will be more compact by grouping several children in the same cluster where they can be accessed locally.

All optimizations possible in B-trees to increase the average fill factors of clusters are possible in the equivalent multicolored binary tree. Notably, maximizing the average fill factor in a structurally equivalent B-tree is the same as reducing the total height of the multicolored tree, by increasing the number of non-black nodes. The worst case



occurs when all nodes in a colored binary tree are black, the best case occurs when only a third of them are black (and the other two thirds are red nodes).

## Applications and related data structures

Red-black trees offer worst-case guarantees for insertion time, deletion time, and search time. Not only does this make them valuable in time-sensitive applications such as real-time applications, but it makes them valuable building blocks in other data structures which provide worst-case guarantees; for example, many data structures used in computational geometry can be based on red-black trees, and the Completely Fair Scheduler used in current Linux kernels uses red-black trees.

The AVL tree is another structure supporting  $O(\log n)$  search, insertion, and removal. It is more rigidly balanced than red-black trees, leading to slower insertion and removal but faster retrieval. This makes it attractive for data structures that may be built once and loaded without reconstruction, such as language dictionaries (or program dictionaries, such as the opcodes of an assembler or interpreter).

Red-black trees are also particularly valuable in functional programming, where they are one of the most common persistent data structures, used to construct associative arrays and sets which can retain previous versions after mutations. The persistent version of red-black trees requires  $O(\log n)$  space for each insertion or deletion, in addition to time.

For every 2-4 tree, there are corresponding red-black trees with data elements in the same order. The insertion and deletion operations on 2-4 trees are also equivalent to color-flipping and rotations in red-black trees. This makes 2-4 trees an important tool for understanding the logic behind red-black trees, and this is why many introductory algorithm texts introduce 2-4 trees just before red-black trees, even though 2-4 trees are not often used in practice.

In 2008, Sedgwick introduced a simpler version of the red-black tree called the left-leaning red-black tree<sup>[5]</sup> by eliminating a previously unspecified degree of freedom in the implementation. The LLRB maintains an additional invariant that all red links must lean left except during inserts and deletes. Red-black trees can be made isometric to either 2-3 trees,<sup>[6]</sup> or 2-4 trees,<sup>[5]</sup> for any sequence of operations. The 2-4 tree isometry was described in 1978 by Sedgwick. With 2-4 trees, the isometry is resolved by a "color flip," corresponding to a split, in which the red color of two children nodes leaves the children and moves to the parent node. The tango tree, a type of tree optimized for fast searches, usually uses red-black trees as part of its data structure.

## Operations

Read-only operations on a red-black tree require no modification from those used for binary search trees, because every red-black tree is a special case of a simple binary search tree. However, the immediate result of an insertion or removal may violate the properties of a red-black tree. Restoring the red-black properties requires a small number ( $O(\log n)$  or amortized  $O(1)$ ) of color changes (which are very quick in practice) and no more than three tree rotations (two for insertion). Although insert and delete operations are complicated, their times remain  $O(\log n)$ .

### Insertion

Insertion begins by adding the node as any binary search tree insertion does and by coloring it red. Whereas in the binary search tree, we always add a leaf, in the red-black tree leaves contain no information, so instead we add a red interior node, with two black leaves, in place of an existing black leaf.

What happens next depends on the color of other nearby nodes. The term *uncle node* will be used to refer to the sibling of a node's parent, as in human family trees. Note that:

- property 3 (all leaves are black) always holds.
  - property 4 (both children of every red node are black) is threatened only by adding a red node, repainting a black node red, or a rotation.
-

- property 5 (all paths from any given node to its leaf nodes contain the same number of black nodes) is threatened only by adding a black node, repainting a red node black (or vice versa), or a rotation.

*Note:* The label **N** will be used to denote the current node (colored red). At the beginning, this is the new node being inserted, but the entire procedure may also be applied recursively to other nodes (see case 3). **P** will denote **N**'s parent node, **G** will denote **N**'s grandparent, and **U** will denote **N**'s uncle. Note that in between some cases, the roles and labels of the nodes are exchanged, but in each case, every label continues to represent the same node it represented at the beginning of the case. Any color shown in the diagram is either assumed in its case or implied by those assumptions.

Each case will be demonstrated with example C code. The uncle and grandparent nodes can be found by these functions:

```
struct node *grandparent(struct node *n)
{
    if ((n != NULL) && (n->parent != NULL))
        return n->parent->parent;
    else
        return NULL;
}

struct node *uncle(struct node *n)
{
    struct node *g = grandparent(n);
    if (g == NULL)
        return NULL; // No grandparent means no uncle
    if (n->parent == g->left)
        return g->right;
    else
        return g->left;
}
```

**Case 1:** The current node **N** is at the root of the tree. In this case, it is repainted black to satisfy property 2 (the root is black). Since this adds one black node to every path at once, property 5 (all paths from any given node to its leaf nodes contain the same number of black nodes) is not violated.

```
void insert_case1(struct node *n)
{
    if (n->parent == NULL)
        n->color = BLACK;
    else
        insert_case2(n);
}
```

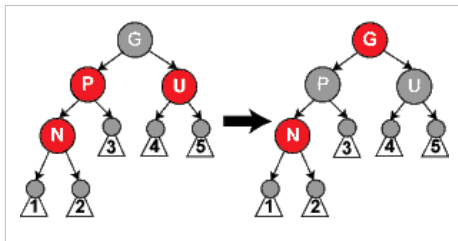
**Case 2:** The current node's parent **P** is black, so property 4 (both children of every red node are black) is not invalidated. In this case, the tree is still valid. property 5 (all paths from any given node to its leaf nodes contain the same number of black nodes) is not threatened, because the current node **N** has two black leaf children, but because **N** is red, the paths through each of its children have the same number of black nodes as the path through the leaf it replaced, which was black, and so this property remains satisfied.

```

void insert_case2(struct node *n)
{
    if (n->parent->color == BLACK)
        return; /* Tree is still valid */
    else
        insert_case3(n);
}

```

*Note:* In the following cases it can be assumed that **N** has a grandparent node **G**, because its parent **P** is red, and if it were the root, it would be black. Thus, **N** also has an uncle node **U**, although it may be a leaf in cases 4 and 5.



**Case 3:** If both the parent **P** and the uncle **U** are red, then both of them can be repainted black and the grandparent **G** becomes red (to maintain property 5 (all paths from any given node to its leaf nodes contain the same number of black nodes)). Now, the current red node **N** has a black parent. Since any path through the parent or uncle must pass through the grandparent, the number of black nodes on these paths has not changed. However, the grandparent **G** may now violate properties 2 (The root is black) or 4 (Both children of every red node are black) (property 4 possibly being violated since **G** may have a red parent). To fix this, the entire procedure is recursively performed on **G** from case 1. Note that this is a tail-recursive call, so it could be rewritten as a loop; since this is the only loop, and any rotations occur after this loop, this proves that a constant number of rotations occur.

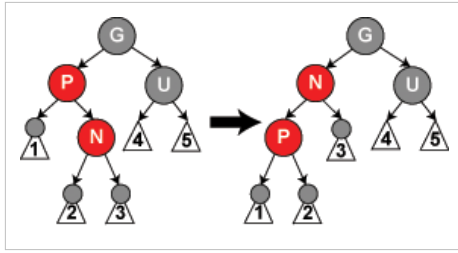
```

void insert_case3(struct node *n)
{
    struct node *u = uncle(n), *g;

    if ((u != NULL) && (u->color == RED)) {
        n->parent->color = BLACK;
        u->color = BLACK;
        g = grandparent(n);
        g->color = RED;
        insert_case1(g);
    } else {
        insert_case4(n);
    }
}

```

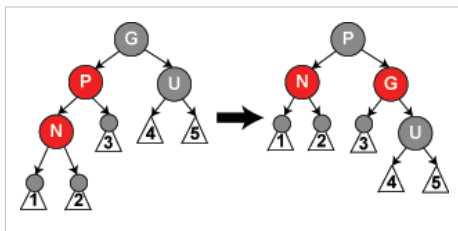
*Note:* In the remaining cases, it is assumed that the parent node **P** is the left child of its parent. If it is the right child, *left* and *right* should be reversed throughout cases 4 and 5. The code samples take care of this.



**Case 4:** The parent **P** is red but the uncle **U** is black; also, the current node **N** is the right child of **P**, and **P** in turn is the left child of its parent **G**. In this case, a left rotation that switches the roles of the current node **N** and its parent **P** can be performed; then, the former parent node **P** is dealt with using case 5 (relabeling **N** and **P**) because property 4 (both children of every red node are black) is still violated. The rotation causes some paths (those in the sub-tree labelled "1") to pass through the node **N** where they did not before. It also causes some paths (those in the sub-tree labelled "3") not to pass through the node **P** where they did before. However, both of these nodes are red, so property 5 (all paths from any given node to its leaf nodes contain the same number of black nodes) is not violated by the rotation. After this case has been completed, property 4 (both children of every red node are black) is still violated, but now we can resolve this by continuing to case 5.

```
void insert_case4(struct node *n)
{
    struct node *g = grandparent(n);

    if ((n == n->parent->right) && (n->parent == g->left)) {
        rotate_left(n->parent);
        n = n->left;
    } else if ((n == n->parent->left) && (n->parent == g->right)) {
        rotate_right(n->parent);
        n = n->right;
    }
    insert_case5(n);
}
```



**Case 5:** The parent **P** is red but the uncle **U** is black, the current node **N** is the left child of **P**, and **P** is the left child of its parent **G**. In this case, a right rotation on **G** is performed; the result is a tree where the former parent **P** is now the parent of both the current node **N** and the former grandparent **G**. **G** is known to be black, since its former child **P** could not have been red otherwise (without violating property 4). Then, the colors of **P** and **G** are switched, and the resulting tree satisfies property 4 (both children of every red node are black). Property 5 (all paths from any given node to its leaf nodes contain the same number of black nodes) also remains satisfied, since all paths that went through any of these three nodes went through **G** before, and now they all go through **P**. In each case, this is the only black node of the three.

```
void insert_case5(struct node *n)
{
    struct node *g = grandparent(n);

    n->parent->color = BLACK;
    g->color = RED;
}
```

```

    if (n == n->parent->left)
        rotate_right(g);
    else
        rotate_left(g);
}

```

Note that inserting is actually in-place, since all the calls above use tail recursion.

## Removal

In a regular binary search tree when deleting a node with two non-leaf children, we find either the maximum element in its left subtree (which is the in-order predecessor) or the minimum element in its right subtree (which is the in-order successor) and move its value into the node being deleted (as shown here). We then delete the node we copied the value from, which must have fewer than two non-leaf children. (Non-leaf children, rather than all children, are specified here because unlike normal binary search trees, red-black trees have leaf nodes anywhere they can have them, so that all nodes are either internal nodes with two children or leaf nodes with, by definition, zero children. In effect, internal nodes having two leaf children in a red-black tree are like the leaf nodes in a regular binary search tree.) Because merely copying a value does not violate any red-black properties, this reduces to the problem of deleting a node with at most one non-leaf child. Once we have solved that problem, the solution applies equally to the case where the node we originally want to delete has at most one non-leaf child as to the case just considered where it has two non-leaf children.

Therefore, for the remainder of this discussion we address the deletion of a node with at most one non-leaf child. We use the label **M** to denote the node to be deleted; **C** will denote a selected child of **M**, which we will also call "its child". If **M** does have a non-leaf child, call that its child, **C**; otherwise, choose either leaf as its child, **C**.

If **M** is a red node, we simply replace it with its child **C**, which must be black by property 4. (This can only occur when **M** has two leaf children, because if the red node **M** had a black non-leaf child on one side but just a leaf child on the other side, then the count of black nodes on both sides would be different, thus the tree would violate property 5.) All paths through the deleted node will simply pass through one less red node, and both the deleted node's parent and child must be black, so property 3 (all leaves are black) and property 4 (both children of every red node are black) still hold.

Another simple case is when **M** is black and **C** is red. Simply removing a black node could break Properties 4 ("Both children of every red node are black") and 5 ("All paths from any given node to its leaf nodes contain the same number of black nodes"), but if we repaint **C** black, both of these properties are preserved.

The complex case is when both **M** and **C** are black. (This can only occur when deleting a black node which has two leaf children, because if the black node **M** had a black non-leaf child on one side but just a leaf child on the other side, then the count of black nodes on both sides would be different, thus the tree would have been an invalid red-black tree by violation of property 5.) We begin by replacing **M** with its child **C**. We will call (or *label*—that is, *relabel*) this child (in its new position) **N**, and its sibling (its new parent's other child) **S**. (**S** was previously the sibling of **M**.) In the diagrams below, we will also use **P** for **N**'s new parent (**M**'s old parent), **S<sub>L</sub>** for **S**'s left child, and **S<sub>R</sub>** for **S**'s right child (**S** cannot be a leaf because if **N** is black, which we presumed, then **P**'s one subtree which includes **N** counts two black-height and thus **P**'s other subtree which includes **S** must also count two black-height, which cannot be the case if **S** is a leaf node).

*Note:* In between some cases, we exchange the roles and labels of the nodes, but in each case, every label continues to represent the same node it represented at the beginning of the case. Any color shown in the diagram is either assumed in its case or implied by those assumptions. White represents an unknown color (either red or black).

We will find the sibling using this function:

```

struct node *sibling(struct node *n)
{
    if (n == n->parent->left)
        return n->parent->right;
    else
        return n->parent->left;
}

```

*Note:* In order that the tree remains well-defined, we need that every null leaf remains a leaf after all transformations (that it will not have any children). If the node we are deleting has a non-leaf (non-null) child **N**, it is easy to see that the property is satisfied. If, on the other hand, **N** would be a null leaf, it can be verified from the diagrams (or code) for all the cases that the property is satisfied as well.

We can perform the steps outlined above with the following code, where the function `replace_node` substitutes `child` into `n`'s place in the tree. For convenience, code in this section will assume that null leaves are represented by actual node objects rather than `NULL` (the code in the *Insertion* section works with either representation).

```

void delete_one_child(struct node *n)
{
    /*
     * Precondition: n has at most one non-null child.
     */
    struct node *child = is_leaf(n->right) ? n->left : n->right;

    replace_node(n, child);
    if (n->color == BLACK) {
        if (child->color == RED)
            child->color = BLACK;
        else
            delete_case1(child);
    }
    free(n);
}

```

*Note:* If **N** is a null leaf and we do not want to represent null leaves as actual node objects, we can modify the algorithm by first calling `delete_case1()` on its parent (the node that we delete, `n` in the code above) and deleting it afterwards. We can do this because the parent is black, so it behaves in the same way as a null leaf (and is sometimes called a 'phantom' leaf). And we can safely delete it at the end as `n` will remain a leaf after all operations, as shown above.

If both **N** and its original parent are black, then deleting this original parent causes paths which proceed through **N** to have one fewer black node than paths that do not. As this violates property 5 (all paths from any given node to its leaf nodes contain the same number of black nodes), the tree must be rebalanced. There are several cases to consider:

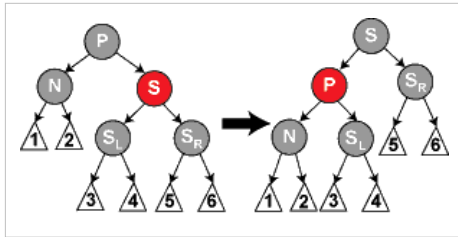
**Case 1:** **N** is the new root. In this case, we are done. We removed one black node from every path, and the new root is black, so the properties are preserved.

```

void delete_case1(struct node *n)
{
    if (n->parent != NULL)
        delete_case2(n);
}

```

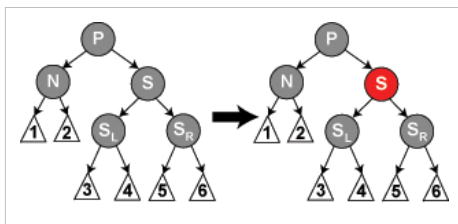
*Note:* In cases 2, 5, and 6, we assume **N** is the left child of its parent **P**. If it is the right child, *left* and *right* should be reversed throughout these three cases. Again, the code examples take both cases into account.



**Case 2:** **S** is red. In this case we reverse the colors of **P** and **S**, and then rotate left at **P**, turning **S** into **N**'s grandparent. Note that **P** has to be black as it had a red child. Although all paths still have the same number of black nodes, now **N** has a black sibling and a red parent, so we can proceed to step 4, 5, or 6. (Its new sibling is black because it was once the child of the red **S**.) In later cases, we will relabel **N**'s new sibling as **S**.

```
void delete_case2(struct node *n)
{
    struct node *s = sibling(n);

    if (s->color == RED) {
        n->parent->color = RED;
        s->color = BLACK;
        if (n == n->parent->left)
            rotate_left(n->parent);
        else
            rotate_right(n->parent);
    }
    delete_case3(n);
}
```



**Case 3:** **P**, **S**, and **S**'s children are black. In this case, we simply repaint **S** red. The result is that all paths passing through **S**, which are precisely those paths *not* passing through **N**, have one less black node. Because deleting **N**'s original parent made all paths passing through **N** have one less black node, this evens things up. However, all paths through **P** now have one fewer black node than paths that do not pass through **P**, so property 5 (all paths from any given node to its leaf nodes contain the same number of black nodes) is still violated. To correct this, we perform the rebalancing procedure on **P**, starting at case 1.

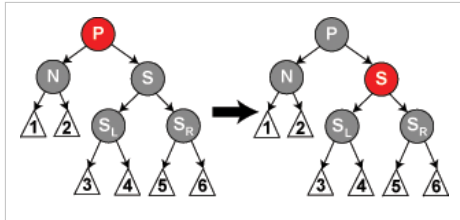
```
void delete_case3(struct node *n)
{
    struct node *s = sibling(n);

    if ((n->parent->color == BLACK) &&
        (s->color == BLACK) &&
        (s->left->color == BLACK) &&
        (s->right->color == BLACK)) {
```

```

        s->color = RED;
        delete_case1(n->parent);
    } else
        delete_case4(n);
}

```



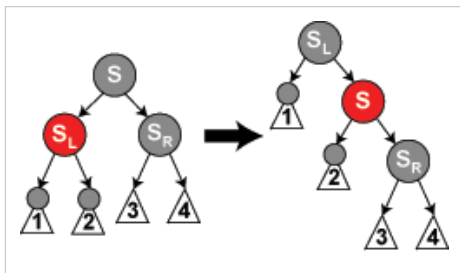
**Case 4:** **S** and **S**'s children are black, but **P** is red. In this case, we simply exchange the colors of **S** and **P**. This does not affect the number of black nodes on paths going through **S**, but it does add one to the number of black nodes on paths going through **N**, making up for the deleted black node on those paths.

```

void delete_case4(struct node *n)
{
    struct node *s = sibling(n);

    if ((n->parent->color == RED) &&
        (s->color == BLACK) &&
        (s->left->color == BLACK) &&
        (s->right->color == BLACK)) {
        s->color = RED;
        n->parent->color = BLACK;
    } else
        delete_case5(n);
}

```



**Case 5:** **S** is black, **S**'s left child is red, **S**'s right child is black, and **N** is the left child of its parent. In this case we rotate right at **S**, so that **S**'s left child becomes **S**'s parent and **N**'s new sibling. We then exchange the colors of **S** and its new parent. All paths still have the same number of black nodes, but now **N** has a black sibling whose right child is red, so we fall into case 6. Neither **N** nor its parent are affected by this transformation. (Again, for case 6, we relabel **N**'s new sibling as **S**.)

```

void delete_case5(struct node *n)
{
    struct node *s = sibling(n);

    if (s->color == BLACK) { /* this if statement is trivial,

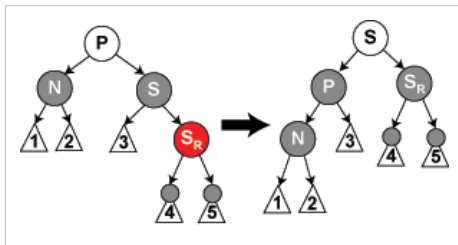
```



```

due to case 2 (even though case 2 changed the sibling to a sibling's
child,
the sibling's child can't be red, since no red parent can have a red
child). */
/* the following statements just force the red to be on the left of the
left of the parent,
or right of the right, so case six will rotate correctly. */
    if ((n == n->parent->left) &&
        (s->right->color == BLACK) &&
        (s->left->color == RED)) { /* this last test is trivial
too due to cases 2-4. */
        s->color = RED;
        s->left->color = BLACK;
        rotate_right(s);
    } else if ((n == n->parent->right) &&
        (s->left->color == BLACK) &&
        (s->right->color == RED)) { /* this last test is
trivial too due to cases 2-4. */
        s->color = RED;
        s->right->color = BLACK;
        rotate_left(s);
    }
}
delete_case6(n);
}

```



**Case 6:** **S** is black, **S**'s right child is red, and **N** is the left child of its parent **P**. In this case we rotate left at **P**, so that **S** becomes the parent of **P** and **S**'s right child. We then exchange the colors of **P** and **S**, and make **S**'s right child black. The subtree still has the same color at its root, so Properties 4 (Both children of every red node are black) and 5 (All paths from any given node to its leaf nodes contain the same number of black nodes) are not violated. However, **N** now has one additional black ancestor: either **P** has become black, or it was black and **S** was added as a black grandparent. Thus, the paths passing through **N** pass through one additional black node.

Meanwhile, if a path does not go through **N**, then there are two possibilities:

- It goes through **N**'s new sibling. Then, it must go through **S** and **P**, both formerly and currently, as they have only exchanged colors and places. Thus the path contains the same number of black nodes.
- It goes through **N**'s new uncle, **S**'s right child. Then, it formerly went through **S**, **S**'s parent, and **S**'s right child (which was red), but now only goes through **S**, which has assumed the color of its former parent, and **S**'s right child, which has changed from red to black (assuming **S**'s color: black). The net effect is that this path goes through the same number of black nodes.

Either way, the number of black nodes on these paths does not change. Thus, we have restored Properties 4 (Both children of every red node are black) and 5 (All paths from any given node to its leaf nodes contain the same number of black nodes). The white node in the diagram can be either red or black, but must refer to the same color both before and after the transformation.

```

void delete_case6(struct node *n)
{
    struct node *s = sibling(n);

    s->color = n->parent->color;
    n->parent->color = BLACK;

    if (n == n->parent->left) {
        s->right->color = BLACK;
        rotate_left(n->parent);
    } else {
        s->left->color = BLACK;
        rotate_right(n->parent);
    }
}

```

Again, the function calls all use tail recursion, so the algorithm is in-place. In the algorithm above, all cases are chained in order, except in delete case 3 where it can recurse to case 1 back to the parent node: this is the only case where an in-place implementation will effectively loop (after only one rotation in case 3).

Additionally, no tail recursion ever occurs on a child node, so the tail recursion loop can only move from a child back to its successive ancestors. No more than  $O(\log n)$  loops back to case 1 will occur (where  $n$  is the total number of nodes in the tree before deletion). If a rotation occurs in case 2 (which is the only possibility of rotation within the loop of cases 1–3), then the parent of the node  $N$  becomes red after the rotation and we will exit the loop. Therefore at most one rotation will occur within this loop. Since no more than two additional rotations will occur after exiting the loop, at most three rotations occur in total.

## Proof of asymptotic bounds

A red black tree which contains  $n$  internal nodes has a height of  $O(\log(n))$ .

Definitions:

- $h(v)$  = height of subtree rooted at node  $v$
- $bh(v)$  = the number of black nodes (not counting  $v$  if it is black) from  $v$  to any leaf in the subtree (called the black-height).

**Lemma:** A subtree rooted at node  $v$  has at least  $2^{bh(v)} - 1$  internal nodes.

Proof of Lemma (by induction height):

Basis:  $h(v) = 0$

If  $v$  has a height of zero then it must be *null*, therefore  $bh(v) = 0$ . So:

$$2^{bh(v)} - 1 = 2^0 - 1 = 1 - 1 = 0$$

Inductive Step:  $v$  such that  $h(v) = k$ , has at least  $2^{bh(v)} - 1$  internal nodes implies that  $v'$  such that  $h(v') = k+1$  has at least  $2^{bh(v')} - 1$  internal nodes.

Since  $v'$  has  $h(v') > 0$  it is an internal node. As such it has two children each of which have a black-height of either  $bh(v')$  or  $bh(v')-1$  (depending on whether the child is red or black, respectively). By the inductive hypothesis each child has at least  $2^{bh(v')-1} - 1$  internal nodes, so  $v'$  has at least:

$$2^{bh(v')-1} - 1 + 2^{bh(v')-1} - 1 + 1 = 2^{bh(v')} - 1$$

internal nodes.

Using this lemma we can now show that the height of the tree is logarithmic. Since at least half of the nodes on any path from the root to a leaf are black (property 4 of a red–black tree), the black-height of the root is at least  $h(\text{root})/2$ . By the lemma we get:

$$n \geq 2^{\frac{h(\text{root})}{2}} - 1 \leftrightarrow \log_2(n+1) \geq \frac{h(\text{root})}{2} \leftrightarrow h(\text{root}) \leq 2 \log_2(n+1).$$

Therefore the height of the root is  $O(\log(n))$ .

### Insertion complexity

In the tree code there is only one loop where the node of the root of the red–black property that we wish to restore,  $x$ , can be moved up the tree by one level at each iteration.

Since the original height of the tree is  $O(\log n)$ , there are  $O(\log n)$  iterations. So overall the insert routine has  $O(\log n)$  complexity.

### Parallel algorithms

Parallel algorithms for constructing red–black trees from sorted lists of items can run in constant time or  $O(\log \log n)$  time, depending on the computer model, if the number of processors available is proportional to the number of items. Fast search, insertion, and deletion parallel algorithms are also known.<sup>[7]</sup>

### Notes

- [1] John Morris. "Red-Black Trees" ([http://www.cs.auckland.ac.nz/~jmor159/PLDS210/red\\_black.html](http://www.cs.auckland.ac.nz/~jmor159/PLDS210/red_black.html)). .>
- [2] Rudolf Bayer (1972). "Symmetric binary B-Trees: Data structure and maintenance algorithms" ([http://www.springerlink.com/content/](http://www.springerlink.com/content/qh51m2014673513j/)[qh51m2014673513j/](http://www.springerlink.com/content/qh51m2014673513j/)). *Acta Informatica* **1** (4): 290–306. doi:10.1007/BF00289509. .
- [3] Leonidas J. Guibas and Robert Sedgwick (1978). "A Dichromatic Framework for Balanced Trees" (<http://doi.ieeecomputersociety.org/10.1109/SFCS.1978.3>). *Proceedings of the 19th Annual Symposium on Foundations of Computer Science*. pp. 8–21. doi:10.1109/SFCS.1978.3. .
- [4] Not all Red-Black trees are balanced? (<http://cs.stackexchange.com/questions/342/not-all-red-black-trees-are-balanced>)
- [5] <http://www.cs.princeton.edu/~rs/talks/LLRB/RedBlack.pdf>
- [6] <http://www.cs.princeton.edu/courses/archive/fall08/cos226/lectures/10BalancedTrees-2x2.pdf>
- [7] H. Park and K. Park (2001). "Parallel algorithms for red–black trees" (<http://www.sciencedirect.com/science/article/pii/S0304397500002875>). *Theoretical computer science* (Elsevier) **262** (1–2): 415–435. doi:10.1016/S0304-3975(00)00287-5. .

### References

- Mathworld: Red–Black Tree (<http://mathworld.wolfram.com/Red-BlackTree.html>)
- San Diego State University: CS 660: Red–Black tree notes (<http://www.eli.sdsu.edu/courses/fall95/cs660/notes/RedBlackTree/RedBlack.html#RTFTToC2>), by Roger Whitney
- Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7 . Chapter 13: Red–Black Trees, pp. 273–301.
- Pfaff, Ben (June 2004). "Performance Analysis of BSTs in System Software" (<http://www.stanford.edu/~blp/papers/libavl.pdf>) (PDF). Stanford University.
- Okasaki, Chris. "Red–Black Trees in a Functional Setting" (<http://www.eecs.usma.edu/webs/people/okasaki/jfp99.ps>) (PS).

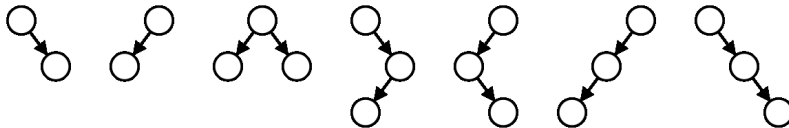
## External links

- In the C++ Standard Template Library, the containers `std::set<Value>` and `std::map<Key, Value>` are typically based on red–black trees
- Tutorial and code for top-down Red–Black Trees ([http://eternallyconfuzzled.com/tuts/datastructures/jsw\\_tut\\_rbtrees.aspx](http://eternallyconfuzzled.com/tuts/datastructures/jsw_tut_rbtrees.aspx))
- C code for Red–Black Trees (<http://github.com/fbuihoo/libtree>)
- Red–Black Tree in GNU libavl C library by Ben Pfaff ([http://www.stanford.edu/~blp/avl/libavl.html/Red\\_002dBlack-Trees.html](http://www.stanford.edu/~blp/avl/libavl.html/Red_002dBlack-Trees.html))
- Red–Black Tree C Code ([http://www.mit.edu/~emin/source\\_code/red\\_black\\_tree/index.html](http://www.mit.edu/~emin/source_code/red_black_tree/index.html))
- Lightweight Java implementation of Persistent Red–Black Trees (<http://wiki.edinburghhacklab.com/PersistentRedBlackTreeSet>)
- VBScript implementation of stack, queue, deque, and Red–Black Tree (<http://www.ludvikjerabek.com/downloads.html>)
- Red–Black Tree Demonstration (<http://www.ece.uc.edu/~franco/C321/html/RedBlack/redblack.html>)
- Red–Black Tree PHP5 Code (<http://code.google.com/p/redblacktreephp/source/browse/#svn/trunk>)
- In Java a freely available red black tree implementation is that of apache commons (<http://commons.apache.org/collections/api-release/org/apache/commons/collections/bidimap/TreeBidiMap.html>)
- Java's TreeSet class internally stores its elements in a red black tree: <http://java.sun.com/docs/books/tutorial/collections/interfaces/set.html>
- Left Leaning Red Black Trees (<http://www.cs.princeton.edu/~rs/talks/LLRB/LLRB.pdf>)
- Left Leaning Red Black Trees Slides (<http://www.cs.princeton.edu/~rs/talks/LLRB/RedBlack.pdf>)
- Left-Leaning Red–Black Tree in ANS-Forth by Hugh Aguilar (<http://www.forth.org/novice.html>) See ASSOCIATION.4TH for the LLRB tree.
- An implementation of left-leaning red-black trees in C# (<http://blogs.msdn.com/b/delay/archive/2009/06/02/maintaining-balance-a-versatile-red-black-tree-implementation-for-net-via-silverlight-wpf-charting.aspx>)
- PPT slides demonstration of manipulating red black trees to facilitate teaching (<http://employees.oneonta.edu/zhangs/PowerPointplatform/>)
- OCW MIT Lecture by Prof. Erik Demaine on Red Black Trees (<http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-046j-introduction-to-algorithms-sma-5503-fall-2005/video-lectures/lecture-10-red-black-trees-rotations-insertions-deletions/>) -
- 1 (<http://www.boyet.com/Articles/RedBlack1.html>) 2 (<http://www.boyet.com/Articles/RedBlack2.html>) 3 (<http://www.boyet.com/Articles/RedBlack3.html>) 4 (<http://www.boyet.com/Articles/RedBlack4.html>) 5 (<http://www.boyet.com/Articles/RedBlack5.html>), a C# Article series by Julian M. Bucknall.
- Open Data Structures - Chapter 9 - Red-Black Trees ([http://opendatastructures.org/versions/edition-0.1e/ods-java/9\\_Red\\_Black\\_Trees.html](http://opendatastructures.org/versions/edition-0.1e/ods-java/9_Red_Black_Trees.html))
- Binary Search Tree Insertion Visualization ([https://www.youtube.com/watch?v=\\_VbTnLV8plU](https://www.youtube.com/watch?v=_VbTnLV8plU)) on YouTube – Visualization of random and pre-sorted data insertions, in elementary binary search trees, and left-leaning red–black trees
- Red Black Tree API in the Linux kernel (<http://lwn.net/Articles/184495/>)

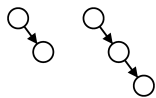
# AA tree

An **AA tree** in computer science is a form of balanced tree used for storing and retrieving ordered data efficiently. AA trees are named for Arne Andersson, their inventor.

AA trees are a variation of the red-black tree, which in turn is an enhancement to the binary search tree. Unlike red-black trees, red nodes on an AA tree can only be added as a right subchild. In other words, no red node can be a left sub-child. This results in the simulation of a 2-3 tree instead of a 2-3-4 tree, which greatly simplifies the maintenance operations. The maintenance algorithms for a red-black tree need to consider seven different shapes to properly balance the tree:



An AA tree on the other hand only needs to consider two shapes due to the strict requirement that only right links can be red:



## Balancing rotations

Whereas red-black trees require one bit of balancing metadata per node (the color), AA trees require  $O(\log(N))$  bits of metadata per node, in the form of an integer "level". The following invariants hold for AA trees:

1. The level of every leaf node is one.
2. The level of every left child is exactly one less than that of its parent.
3. The level of every right child is equal to or one less than that of its parent.
4. The level of every right grandchild is strictly less than that of its grandparent.
5. Every node of level greater than one has two children.

A link where the child's level is equal to that of its parent is called a *horizontal* link, and is analogous to a red link in the red-black tree. Individual right horizontal links are allowed, but consecutive ones are forbidden; all left horizontal links are forbidden. These are more restrictive constraints than the analogous ones on red-black trees, with the result that re-balancing an AA tree is procedurally much simpler than re-balancing a red-black tree.

Insertions and deletions may transiently cause an AA tree to become unbalanced (that is, to violate the AA tree invariants). Only two distinct operations are needed for restoring balance: "skew" and "split". Skew is a right rotation to replace a subtree containing a left horizontal link with one containing a right horizontal link instead. Split is a left rotation and level increase to replace a subtree containing two or more consecutive right horizontal links with one containing two fewer consecutive right horizontal links. Implementation of balance-preserving insertion and deletion is simplified by relying on the skew and split operations to modify the tree only if needed, instead of making their callers decide whether to skew or split.

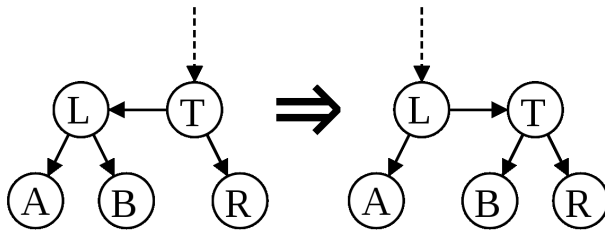
```
function skew is
  input: T, a node representing an AA tree that needs to be rebalanced.
  output: Another node representing the rebalanced AA tree.

  if nil(T) then
    return Nil
```

```

else if nil(left(T)) then
    return T
else if level(left(T)) == level(T) then
    Swap the pointers of horizontal left links.
    L = left(T)
    left(T) := right(L)
    right(L) := T
    return L
else
    return T
end if
end function

```



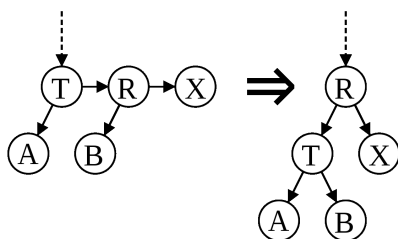
Skew:

```

function split is
    input: T, a node representing an AA tree that needs to be rebalanced.
    output: Another node representing the rebalanced AA tree.

    if nil(T) then
        return Nil
    else if nil(right(T)) or nil(right(right(T))) then
        return T
    else if level(T) == level(right(right(T))) then
        We have two horizontal right links. Take the middle node, elevate it, and return it.
        R = right(T)
        right(T) := left(R)
        left(R) := T
        level(R) := level(R) + 1
        return R
    else
        return T
    end if
end function

```



Split:

## Insertion

Insertion begins with the normal binary tree search and insertion procedure. Then, as the call stack unwinds (assuming a recursive implementation of the search), it's easy to check the validity of the tree and perform any rotations as necessary. If a horizontal left link arises, a skew will be performed, and if two horizontal right links arise, a split will be performed, possibly incrementing the level of the new root node of the current subtree. Note, in the code as given above, the increment of `level(T)`. This makes it necessary to continue checking the validity of the tree as the modifications bubble up from the leaves.

```
function insert is
  input: X, the value to be inserted, and T, the root of the tree to insert it into.
  output: A balanced version T including X.

  Do the normal binary tree insertion procedure. Set the result of the
recursive call to the correct child in case a new node was created or the
root of the subtree changes.
  if nil(T) then
    Create a new leaf node with X.
    return node(X, 1, Nil, Nil)
  else if X < value(T) then
    left(T) := insert(X, left(T))
  else if X > value(T) then
    right(T) := insert(X, right(T))
  end if

  Note that the case of X == value(T) is unspecified. As given, an insert
will have no effect. The implementor may desire different behavior.

  Perform skew and then split. The conditionals that determine whether or
not a rotation will occur or not are inside of the procedures, as given
above.
  T := skew(T)
  T := split(T)

  return T
end function
```

## Deletion

As in most balanced binary trees, the deletion of an internal node can be turned into the deletion of a leaf node by swapping the internal node with either its closest predecessor or successor, depending on which are in the tree or on the implementor's whims. Retrieving a predecessor is simply a matter of following one left link and then all of the remaining right links. Similarly, the successor can be found by going right once and left until a null pointer is found. Because of the AA property of all nodes of level greater than one having two children, the successor or predecessor node will be in level 1, making their removal trivial.

To re-balance a tree, there are a few approaches. The one described by Andersson in his original paper <sup>[1]</sup> is the simplest, and it is described here, although actual implementations may opt for a more optimized approach. After a removal, the first step to maintaining tree validity is to lower the level of any nodes whose children are two levels below them, or who are missing children. Then, the entire level must be skewed and split. This approach was favored, because when laid down conceptually, it has three easily understood separate steps:

1. Decrease the level, if appropriate.
2. Skew the level.
3. Split the level.

However, we have to skew and split the entire level this time instead of just a node, complicating our code.

```
function delete is
  input: X, the value to delete, and T, the root of the tree from which it should be deleted.
  output: T, balanced, without the value X.

  if nil(T) then
    return T
  else if X > value(T) then
    right(T) := delete(X, right(T))
  else if X < value(T) then
    left(T) := delete(X, left(T))
  else
    If we're a leaf, easy, otherwise reduce to leaf case.
    if leaf(T) then
      return Nil
    else if nil(left(T)) then
      L := successor(T)
      right(T) := delete(L, right(T))
      value(T) := L
    else
      L := predecessor(T)
      left(T) := delete(L, left(T))
      value(T) := L
    end if
  end if

  Rebalance the tree. Decrease the level of all nodes in this level if
  necessary, and then skew and split all nodes in the new level.
  T := decrease_level(T)
  T := skew(T)
  right(T) := skew(right(T))
  right(right(T)) := skew(right(right(T)))
  T := split(T)
  right(T) := split(right(T))
  return T
end function
```

```
function decrease_level is
  input: T, a tree for which we want to remove links that skip levels.
  output: T with its level decreased.

  should_be = min(level(left(T)), level(right(T))) + 1
  if should_be < level(T) then
    level(T) := should_be
```



```
    if should_be < level(right(T)) then
        level(right(T)) := should_be
    end if
end if
return T
end function
```

A good example of deletion by this algorithm is present in the Andersson paper <sup>[1]</sup>.

## Performance

The performance of an AA tree is equivalent to the performance of a red-black tree. While an AA tree makes more rotations than a red-black tree, the simpler algorithms tend to be faster, and all of this balances out to result in similar performance. A red-black tree is more consistent in its performance than an AA tree, but an AA tree tends to be flatter, which results in slightly faster search times.<sup>[2]</sup>

## References

[1] <http://user.it.uu.se/~arnea/abs/simp.html>

[2] "A Disquisition on The Performance Behavior of Binary Search Tree Data Structures (pages 67-75)" (<http://www.cepis.org/upgrade/files/full-2004-V.pdf>). .

## External links

- A. Andersson. Balanced search trees made simple (<http://user.it.uu.se/~arnea/abs/simp.html>)
- A. Andersson. A note on searching in a binary search tree (<http://user.it.uu.se/~arnea/abs/searchproc.html>)
- AA-Tree Applet (<http://people.ksp.sk/~kuko/bak/index.html>) by Kubo Kovac
- BSTlib (<http://bitbucket.org/trijezdci/bstlib/src/>) - Open source AA tree library for C by trijezdci
- AA Visual 2007 1.5 - OpenSource Delphi program for educating AA tree structures (<http://www.softpedia.com/get/Others/Home-Education/AA-Visual-2007.shtml>)
- Thorough tutorial ([http://www.eternallyconfuzzled.com/tuts/datastructures/jsw\\_tut\\_andersson.aspx](http://www.eternallyconfuzzled.com/tuts/datastructures/jsw_tut_andersson.aspx)) Julianne Walker with lots of code, including a practical implementation
- Object Oriented implementation with tests ([http://www.cs.fiu.edu/~weiss/dsaa\\_c++3/code/](http://www.cs.fiu.edu/~weiss/dsaa_c++3/code/))
- A Disquisition on The Performance Behavior of Binary Search Tree Data Structures (pages 67-75) (<http://www.cepis.org/upgrade/files/full-2004-V.pdf>) - Comparison of AA trees, red-black trees, treaps, skip lists, and radix trees
- An example C implementation (<http://www.rational.co.za/aatree.c>)
- An Objective-C implementation (<http://code.google.com/p/objc-aatree/>)

# Scapegoat tree

In computer science, a **scapegoat tree** is a self-balancing binary search tree, discovered by Arne Anderson<sup>[1]</sup> and again by Igal Galperin and Ronald L. Rivest.<sup>[2]</sup> It provides worst-case  $O(\log n)$  lookup time, and  $O(\log n)$  amortized insertion and deletion time.

Unlike most other self-balancing binary search trees that provide worst case  $O(\log n)$  lookup time, scapegoat trees have no additional per-node memory overhead compared to a regular binary search tree: a node stores only a key and two pointers to the child nodes. This makes scapegoat trees easier to implement and, due to data structure alignment, can reduce node overhead by up to one-third.

## Theory

A binary search tree is said to be weight balanced if half the nodes are on the left of the root, and half on the right. An  $\alpha$ -weight-balanced is therefore defined as meeting the following conditions:

```
size(left)  <=  $\alpha$ *size(node)
size(right) <=  $\alpha$ *size(node)
```

Where size can be defined recursively as:

```
function size(node)
  if node = nil
    return 0
  else
    return size(node->left) + size(node->right) + 1
end
```

An  $\alpha$  of 1 therefore would describe a linked list as balanced, whereas an  $\alpha$  of 0.5 would only match almost complete binary trees.

A binary search tree that is  $\alpha$ -weight-balanced must also be  **$\alpha$ -height-balanced**, that is

```
height(tree) <=  $\log_{1/\alpha}$ (NodeCount)
```

Scapegoat trees are not guaranteed to keep  $\alpha$ -weight-balance at all times, but are always loosely  $\alpha$ -height-balance in that

```
height(scapegoat tree) <=  $\log_{1/\alpha}$ (NodeCount) + 1
```

This makes scapegoat trees similar to red-black trees in that they both have restrictions on their height. They differ greatly though in their implementations of determining where the rotations (or in the case of scapegoat trees, rebalances) take place. Whereas red-black trees store additional 'color' information in each node to determine the location, scapegoat trees find a **scapegoat** which isn't  $\alpha$ -weight-balanced to perform the rebalance operation on. This is loosely similar to AVL trees, in that the actual rotations depend on 'balances' of nodes, but the means of determining the balance differs greatly. Since AVL trees check the balance value on every insertion/deletion, it is typically stored in each node; scapegoat trees are able to calculate it only as needed, which is only when a scapegoat needs to be found.

Unlike most other self-balancing search trees, scapegoat trees are entirely flexible as to their balancing. They support any  $\alpha$  such that  $0.5 < \alpha < 1$ . A high  $\alpha$  value results in fewer balances, making insertion quicker but lookups and deletions slower, and vice versa for a low  $\alpha$ . Therefore in practical applications, an  $\alpha$  can be chosen depending on how frequently these actions should be performed.

## Operations

### Insertion

Insertion is implemented with the same basic ideas as an unbalanced binary search tree, however with a few significant changes.

When finding the insertion point, the depth of the new node must also be recorded. This is implemented via a simple counter that gets incremented during each iteration of the lookup, effectively counting the number of edges between the root and the inserted node. If this node violates the  $\alpha$ -height-balance property (defined above), a rebalance is required.

To rebalance, an entire subtree rooted at a **scapegoat** undergoes a balancing operation. The scapegoat is defined as being an ancestor of the inserted node which isn't  $\alpha$ -weight-balanced. There will always be at least one such ancestor. Rebalancing any of them will restore the  $\alpha$ -height-balanced property.

One way of finding a scapegoat, is to climb from the new node back up to the root and select the first node that isn't  $\alpha$ -weight-balanced.

Climbing back up to the root requires  $O(\log n)$  storage space, usually allocated on the stack, or parent pointers. This can actually be avoided by pointing each child at its parent as you go down, and repairing on the walk back up.

To determine whether a potential node is a viable scapegoat, we need to check its  $\alpha$ -weight-balanced property. To do this we can go back to the definition:

```
size(left) <=  $\alpha$ *size(node)
size(right) <=  $\alpha$ *size(node)
```

However a large optimisation can be made by realising that we already know two of the three sizes, leaving only the third having to be calculated.

Consider the following example to demonstrate this. Assuming that we're climbing back up to the root:

```
size(parent) = size(node) + size(sibling) + 1
```

But as:

```
size(inserted node) = 1.
```

The case is trivialized down to:

```
size[x+1] = size[x] + size(sibling) + 1
```

Where  $x$  = this node,  $x + 1$  = parent and `size(sibling)` is the only function call actually required.

Once the scapegoat is found, the subtree rooted at the scapegoat is completely rebuilt to be perfectly balanced.<sup>[2]</sup> This can be done in  $O(n)$  time by traversing the nodes of the subtree to find their values in sorted order and recursively choosing the median as the root of the subtree.

As rebalance operations take  $O(n)$  time (dependent on the number of nodes of the subtree), insertion has a worst case performance of  $O(n)$  time. However, because these worst-case scenarios are spread out, insertion takes  $O(\log n)$  amortized time.

### Sketch of proof for cost of insertion

Define the Imbalance of a node  $v$  to be the absolute value of the difference in size between its left node and right node minus 1, or 0, whichever is greater. In other words:

$$I(v) = \max(|\text{left}(v) - \text{right}(v)| - 1, 0)$$

Immediately after rebuilding a subtree rooted at  $v$ ,  $I(v) = 0$ .

**Lemma:** Immediately before rebuilding the subtree rooted at  $v$ ,

$$I(v) = \Omega(|v|)$$

( $\Omega$  is Big O Notation.)

Proof of lemma:

Let  $v_0$  be the root of a subtree immediately after rebuilding.  $h(v_0) = \log(|v_0| + 1)$ . If there are  $\Omega(|v_0|)$  degenerate insertions (that is, where each inserted node increases the height by 1), then

$$I(v) = \Omega(|v_0|),$$

$$h(v) = h(v_0) + \Omega(|v_0|) \text{ and}$$

$$\log(|v|) < \log(|v_0| + 1) + 1.$$

Since  $I(v) = \Omega(|v|)$  before rebuilding, there were  $\Omega(|v|)$  insertions into the subtree rooted at  $v$  that did not result in rebuilding. Each of these insertions can be performed in  $O(\log n)$  time. The final insertion that causes rebuilding costs  $O(|v|)$ . Using aggregate analysis it becomes clear that the amortized cost of an insertion is

$$\frac{O(\log n) \cdot \Omega(|v|) + O(|v|)}{\Omega(|v|)} = O(\log n)$$

### Deletion

Scapegoat trees are unusual in that deletion is easier than insertion. To enable deletion, scapegoat trees need to store an additional value with the tree data structure. This property, which we will call `MaxNodeCount` simply represents the highest achieved `NodeCount`. It is set to `NodeCount` whenever the entire tree is rebalanced, and after insertion is set to  $\max(\text{MaxNodeCount}, \text{NodeCount})$ .

To perform a deletion, we simply remove the node as you would in a simple binary search tree, but if

```
NodeCount <= MaxNodeCount / 2
```

then we rebalance the entire tree about the root, remembering to set `MaxNodeCount` to `NodeCount`.

This gives deletion its worst case performance of  $O(n)$  time, however it is amortized to  $O(\log n)$  average time.

### Sketch of proof for cost of deletion

Suppose the scapegoat tree has  $n$  elements and has just been rebuilt (in other words, it is a complete binary tree). At most  $n/2 - 1$  deletions can be performed before the tree must be rebuilt. Each of these deletions take  $O(\log n)$  time (the amount of time to search for the element and flag it as deleted). The  $n/2$  deletion causes the tree to be rebuilt and takes  $O(\log n) + O(n)$  (or just  $O(n)$ ) time. Using aggregate analysis it becomes clear that the amortized cost of a deletion is  $O(\log n)$ :

$$\frac{\sum_{i=1}^{n/2} O(\log n) + O(n)}{n/2} = \frac{n/2 O(\log n) + O(n)}{n/2} = O(\log n)$$

## Lookup

Lookup is not modified from a standard binary search tree, and has a worst-case time of  $O(\log n)$ . This is in contrast to splay trees which have a worst-case time of  $O(n)$ . The reduced node memory overhead compared to other self-balancing binary search trees can further improve locality of reference and caching.

## References

- [1] Andersson, Arne (1989). "Improving partial rebuilding by using simple balance criteria". *Proc. Workshop on Algorithms and Data Structures*. Springer-Verlag. pp. 393–402. doi:10.1007/3-540-51542-9\_33.
- [2] Galperin, Igal; Rivest, Ronald L. (1993), "Scapegoat trees" (<http://portal.acm.org/citation.cfm?id=313676>), *Proceedings of the fourth annual ACM-SIAM Symposium on Discrete algorithms*: 165–174,

## External links

- Scapegoat Tree Applet (<http://people.ksp.sk/~kuko/bak/index.html>) by Kubo Kovac
- Scapegoat Trees: Galperin and Rivest's paper describing scapegoat trees (<http://cg.scs.carleton.ca/~morin/teaching/5408/refs/gr93.pdf>)
- On Consulting a Set of Experts and Searching (full version paper) (<http://publications.csail.mit.edu/lcs/pubs/pdf/MIT-LCS-TR-700.pdf>)
- Open Data Structures - Chapter 8 - Scapegoat Trees ([http://opendatastructures.org/versions/edition-0.1e/ods-java/8\\_Scapegoat\\_Trees.html](http://opendatastructures.org/versions/edition-0.1e/ods-java/8_Scapegoat_Trees.html))

# Splay tree

Splay tree		
Type	Tree	
Invented	1985	
Invented by	Daniel Dominic Sleator and Robert Endre Tarjan	
Time complexity in big O notation		
	Average	Worst case
Space	O(n)	O(n)
Search	O(log n)	amortized O(n)
Insert	O(log n)	amortized O(log n)
Delete	O(log n)	amortized O(log n)

A **splay tree** is a self-adjusting binary search tree with the additional property that recently accessed elements are quick to access again. It performs basic operations such as insertion, look-up and removal in  $O(\log n)$  amortized time. For many sequences of nonrandom operations, splay trees perform better than other search trees, even when the specific pattern of the sequence is unknown. The splay tree was invented by Daniel Dominic Sleator and Robert Endre Tarjan in 1985.<sup>[1]</sup>

All normal operations on a binary search tree are combined with one basic operation, called *splaying*. Splaying the tree for a certain element rearranges the tree so that the element is placed at the root of the tree. One way to do this is to first perform a standard binary tree search for the element in question, and then use tree rotations in a specific fashion to bring the element to the top. Alternatively, a top-down algorithm can combine the search and the tree reorganization into a single phase.

## Advantages

Good performance for a splay tree depends on the fact that it is self-optimizing, in that frequently accessed nodes will move nearer to the root where they can be accessed more quickly. The worst-case height—though unlikely—is  $O(n)$ , with the average being  $O(\log n)$ . Having frequently used nodes near the root is an advantage for nearly all practical applications (also see Locality of reference), and is particularly useful for implementing caches and garbage collection algorithms.

Advantages include:

- Simple implementation—simpler than other self-balancing binary search trees, such as red-black trees or AVL trees.
- Comparable performance—average-case performance is as efficient as other trees.
- Small memory footprint—splay trees do not need to store any bookkeeping data.
- Possibility of creating a "persistent data structure" version of splay trees—which allows access to both the previous and new versions after an update. This can be useful in functional programming, and requires amortized  $O(\log n)$  space per update.
- Working well with nodes containing identical keys—contrary to other types of self-balancing trees. Even with identical keys, performance remains amortized  $O(\log n)$ . All tree operations preserve the order of the identical nodes within the tree, which is a property similar to stable sorting algorithms. A carefully designed find operation can return the leftmost or rightmost node of a given key.

## Disadvantages

Perhaps the most significant disadvantage of splay trees is that the height of a splay tree can be linear. For example, this will be the case after accessing all  $n$  elements in non-decreasing order. Since the height of a tree corresponds to the worst-case access time, this means that the actual cost of an operation can be slow. However the amortized access cost of this worst case is logarithmic,  $O(\log n)$ . Also, the expected access cost can be reduced to  $O(\log n)$  by using a randomized variant<sup>[2]</sup>.

A splay tree can be worse than a static tree by at most a constant factor.

Splay trees can change even when they are accessed in a 'read-only' manner (i.e. by *find* operations). This complicates the use of such splay trees in a multi-threaded environment. Specifically, extra management is needed if multiple threads are allowed to perform *find* operations concurrently.

## Operations

### Splaying

When a node  $x$  is accessed, a splay operation is performed on  $x$  to move it to the root. To perform a splay operation we carry out a sequence of *splay steps*, each of which moves  $x$  closer to the root. By performing a splay operation on the node of interest after every access, the recently accessed nodes are kept near the root and the tree remains roughly balanced, so that we achieve the desired amortized time bounds.

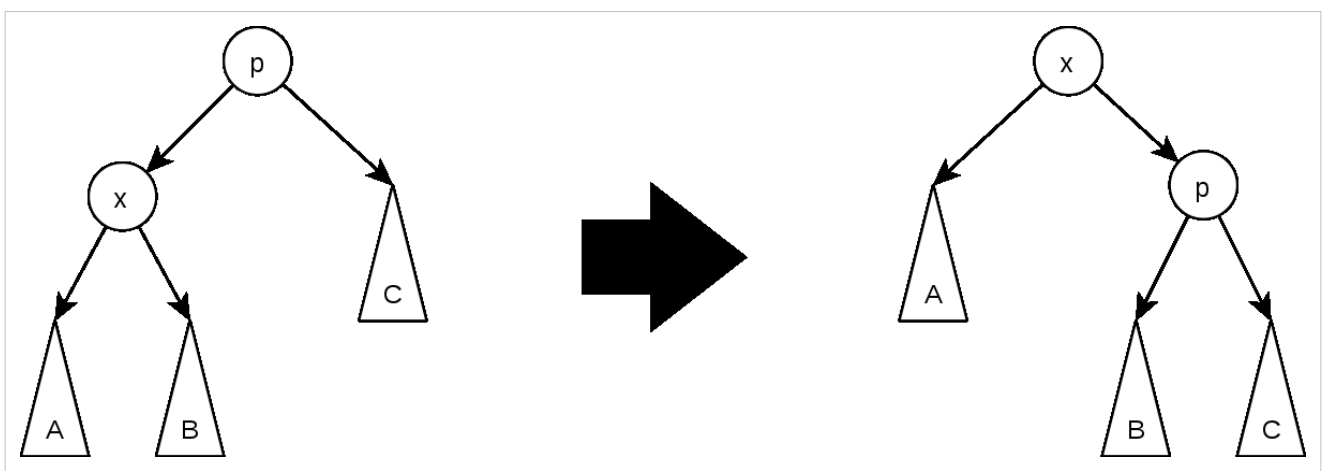
Each particular step depends on three factors:

- Whether  $x$  is the left or right child of its parent node,  $p$ ,
- whether  $p$  is the root or not, and if not
- whether  $p$  is the left or right child of *its* parent,  $g$  (the *grandparent* of  $x$ ).

It is important to remember to set  $gg$  (the *great-grandparent* of  $x$ ) to now point to  $x$  after any splay operation. If  $gg$  is null, then  $x$  obviously is now the root and must be updated as such.

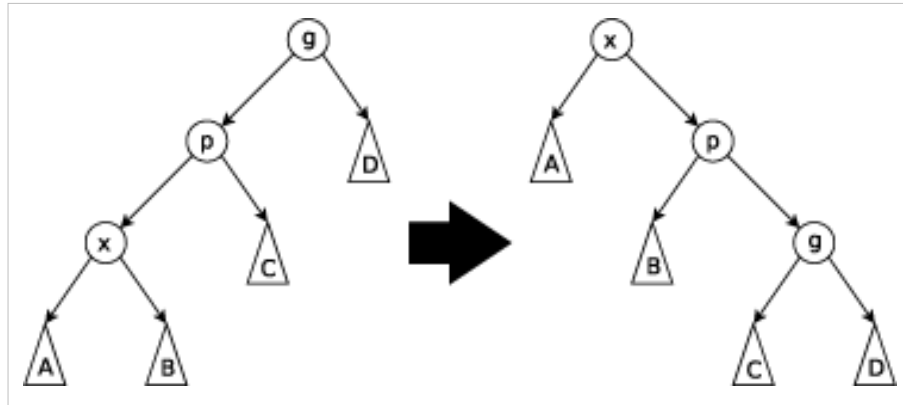
The three types of splay steps are:

**Zig Step:** This step is done when  $p$  is the root. The tree is rotated on the edge between  $x$  and  $p$ . Zig steps exist to deal with the parity issue and will be done only as the last step in a splay operation and only when  $x$  has odd depth at the beginning of the operation.

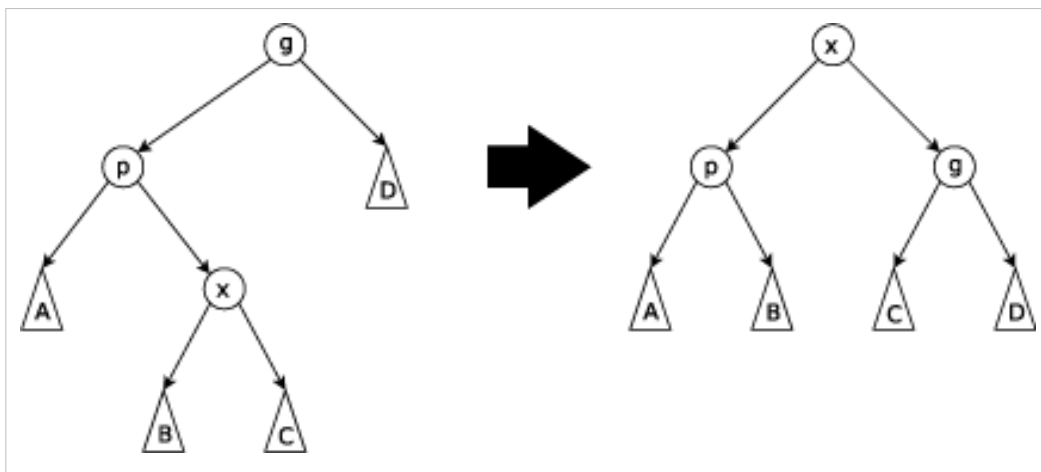


**Zig-zig Step:** This step is done when  $p$  is not the root and  $x$  and  $p$  are either both right children or are both left children. The picture below shows the case where  $x$  and  $p$  are both left children. The tree is rotated on the edge joining  $p$  with *its* parent  $g$ , then rotated on the edge joining  $x$  with  $p$ . Note that zig-zig steps are the only thing that differentiate splay trees from the *rotate to root* method introduced by Allen and Munro<sup>[3]</sup> prior to the introduction of

splay trees.



**Zig-zag Step:** This step is done when  $p$  is not the root and  $x$  is a right child and  $p$  is a left child or vice versa. The tree is rotated on the edge between  $x$  and  $p$ , then rotated on the edge between  $x$  and its new parent  $g$ .



## Insertion

To insert a node  $x$  into a splay tree:

1. First insert the node as with a normal binary search tree.
2. Then splay the newly inserted node  $x$  to the top of the tree.

## Deletion

To delete a node  $x$ , we use the same method as with a binary search tree: if  $x$  has two children, we swap its value with that of either the rightmost node of its left sub tree (its in-order predecessor) or the leftmost node of its right subtree (its in-order successor). Then we remove that node instead. In this way, deletion is reduced to the problem of removing a node with 0 or 1 children.

Unlike a binary search tree, in a splay tree after deletion, we splay the parent of the removed node to the top of the tree. **OR** The node to be deleted is first splayed, i.e. brought to the root of the tree and then deleted. This leaves the tree with two sub trees. The maximum element of the left sub tree (: **METHOD 1**), or minimum of the right sub tree (: **METHOD 2**) is then splayed to the root. The right sub tree is made the right child of the resultant left sub tree (for **METHOD 1**). The root of left sub tree is the root of melded tree.



## Code in C language

### Splay operation in BST

Here x is the node on which the splay operation is performed and root is the root node of the tree.

```
#include<stdio.h>
#include<malloc.h>
#include<stdlib.h>
struct node
{
    int data;
    struct node *parent;
    struct node *left;
    struct node *right;
};
int data_print(struct node *x);
struct node *rightrotation(struct node *p, struct node *root);
struct node *leftrotation(struct node *p, struct node *root);
void splay (struct node *x, struct node *root);
struct node *insert(struct node *p, int value);
struct node *inorder(struct node *p);
struct node *delete(struct node *p, int value);
struct node *successor(struct node *x);
struct node *lookup(struct node *p, int value);

void splay (struct node *x, struct node *root)
{
    struct node *p, *g;
    /*check if node x is the root node*/
    if(x==root)
        return;
    /*Performs Zig step*/
    else if(x->parent==root)
    {
        if(x==x->parent->left)
            root=rightrotation(root, root);
        else
            root=leftrotation(root, root);
    }
    else
    {
        p=x->parent; /*now points to parent of x*/
        g=p->parent; /*now points to parent of x's parent*/
        /*Performs the Zig-zig step when x is left and x's parent
is left*/
        if(x==p->left&&p==g->left)
        {
            root=rightrotation(g, root);
        }
    }
}
```

```

        root=rightrotation(p, root);
    }
    /*Performs the Zig-zig step when x is right and x's parent
is right*/
    else if(x==p->right&&g==p->right)
    {
        root=leftrotation(g, root);
        root=leftrotation(p, root);
    }
    /*Performs the Zig-zag step when x's is right and x's
parent is left*/
    else if(x==p->right&&g==p->left)
    {
        root=leftrotation(p, root);
        root=rightrotation(g, root);
    }
    /*Performs the Zig-zag step when x's is left and x's parent
is right*/
    else if(x==p->left&&g==p->right)
    {
        root=rightrotation(p, root);
        root=leftrotation(g, root);
    }
    splay(x, root);
}
}

struct node *rightrotation(struct node *p, struct node *root)
{
    struct node *x;
    x = p->left;
    p->left = x->right;
    if (x->right!=NULL) x->right->parent = p;
    x->right = p;
    if (p->parent!=NULL)
        if(p==p->parent->right) p->parent->right=x;
        else
            p->parent->left=x;
    x->parent = p->parent;
    p->parent = x;
    if (p==root)
        return x;
    else
        return root;
}

struct node *leftrotation(struct node *p, struct node *root)
{
    struct node *x;

```

```

    x = p->right;
    p->right = x->left;
    if (x->left!=NULL) x->left->parent = p;
    x->left = p;
    if (p->parent!=NULL)
        if (p==p->parent->left) p->parent->left=x;
        else
            p->parent->right=x;
    x->parent = p->parent;
    p->parent = x;
    if(p==root)
        return x;
    else
        return root;
}

struct node *insert(struct node *p,int value)
{
    struct node *temp1,*temp2,*par,*x;
    if(p == NULL)
    {
        p=(struct node *)malloc(sizeof(struct node));
        if(p != NULL)
        {
            p->data = value;
            p->parent = NULL;
            p->left = NULL;
            p->right = NULL;
        }
        else
        {
            printf("No memory is allocated\n");
            exit(0);
        }
        return(p);
    } //the case 2 says that we must splay newly inserted node to
root
    else
    {
        temp2 = p;
        while(temp2 != NULL)
        {
            temp1 = temp2;
            if(temp2->data > value)
                temp2 = temp2->left;
            else if(temp2->data < value)
                temp2 = temp2->right;
            else

```

```

        if(temp2->data == value)
            return temp2;
    }
    if(temp1->data > value)
    {
        par = temp1; //temp1 having the parent
address, so that's it
        temp1->left = (struct node
*) malloc(sizeof(struct node));
        temp1 = temp1->left;
        if(temp1 != NULL)
        {
            temp1->data = value;
            temp1->parent = par; //store the parent
address.
            temp1->left = NULL;
            temp1->right = NULL;
        }
        else
        {
            printf("No memory is allocated\n");
            exit(0);
        }
    }
    else
    {
        par = temp1; //temp1 having the parent node
address.
        temp1->right = (struct node
*) malloc(sizeof(struct node));
        temp1 = temp1->right;
        if(temp1 != NULL)
        {
            temp1->data = value;
            temp1->parent = par; //store the parent
address
            temp1->left = NULL;
            temp1->right = NULL;
        }
        else
        {
            printf("No memory is allocated\n");
            exit(0);
        }
    }
}
splay(temp1, p); //temp1 will be new root after splaying

```

```

    return (templ);
}
struct node *inorder(struct node *p)
{
    if(p != NULL)
    {
        inorder(p->left);
        printf("CURRENT %d\t",p->data);
        printf("LEFT %d\t",data_print(p->left));
        printf("PARENT %d\t",data_print(p->parent));
        printf("RIGHT %d\t\n",data_print(p->right));
        inorder(p->right);
    }
}
struct node *delete(struct node *p,int value)
{
    struct node *x,*y,*pl;
    struct node *root;
    struct node *s;
    root = p;
    x = lookup(p,value);
    if(x->data == value)
    {
        //if the deleted element is leaf
        if((x->left == NULL) && (x->right == NULL))
        {
            y = x->parent;
            if(x == (x->parent->right))
                y->right = NULL;
            else
                y->left = NULL;
            free(x);
        }
        //if deleted element having left child only
        else if((x->left != NULL) && (x->right == NULL))
        {
            if(x == (x->parent->left))
            {
                y = x->parent;
                x->left->parent = y;
                y->left = x->left;
                free(x);
            }
            else
            {
                y = x->parent;
                x->left->parent = y;
                y->right = x->left;
            }
        }
    }
}

```

```
        free(x);
    }
}
//if deleted element having right child only
else if((x->left == NULL) && (x->right != NULL))
{
    if(x == (x->parent->left))
    {
        y = x->parent;
        x->right->parent = y;
        y->left = x->right;
        free(x);
    }
    else
    {
        y = x->parent;
        x->right->parent = y;
        y->right = x->right;
        free(x);
    }
}
//if the deleted element having two children
else if((x->left != NULL) && (x->right != NULL))
{
    if(x == (x->parent->left))
    {
        s = successor(x);
        if(s != x->right)
        {
            y = s->parent;
            if(s->right != NULL)
            {
                s->right->parent = y;
                y->left = s->right;
            }
            else y->left = NULL;
            s->parent = x->parent;
            x->right->parent = s;
            x->left->parent = s;
            s->right = x->right;
            s->left = x->left;
            x->parent->left = s;
        }
        else
        {
            y = s;
            s->parent = x->parent;
```

```

        x->left->parent = s;
        s->left = x->left;
        x->parent->left = s;
    }
    free(x);
}
else if(x == (x->parent->right))
{
    s = successor(x);
    if(s != x->right)
    {
        y = s->parent;
        if(s->right != NULL)
        {
            s->right->parent = y;
            y->left = s->right;
        }
        else y->left = NULL;
        s->parent = x->parent;
        x->right->parent = s;
        x->left->parent = s;
        s->right = x->right;
        s->left = x->left;
        x->parent->right = s;
    }
    else
    {
        y = s;
        s->parent = x->parent;
        x->left->parent = s;
        s->left = x->left;
        x->parent->right = s;
    }
    free(x);
}

}
splay(y, root);
}
else
{
    splay(x, root);
}
}
struct node *successor(struct node *x)
{
    struct node *temp, *temp2;

```

```
temp=temp2=x->right;
while(temp != NULL)
{
    temp2 = temp;
    temp = temp->left;
}
return temp2;
}
//p is a root element of the tree
struct node *lookup(struct node *p,int value)
{
    struct node *temp1,*temp2;
    if(p != NULL)
    {
        temp1 = p;
        while(temp1 != NULL)
        {
            temp2 = temp1;
            if(temp1->data > value)
                temp1 = temp1->left;
            else if(temp1->data < value)
                temp1 = temp1->right;
            else
                return temp1;
        }
        return temp2;
    }
    else
    {
        printf("NO element in the tree\n");
        exit(0);
    }
}
struct node *search(struct node *p,int value)
{
    struct node *x,*root;
    root = p;
    x = lookup(p,value);
    if(x->data == value)
    {
        printf("Inside search if\n");
        splay(x,root);
    }
    else
    {
        printf("Inside search else\n");
        splay(x,root);
    }
}
```



```

    }
}
main()
{
    struct node *root; //the root element
    struct node *x; //x is which element will come to root.
    int i;
    root = NULL;
    int choice = 0;
    int ele;
    while(1)
    {
        printf("\n\n 1.Insert");
        printf("\n\n 2.Delete");
        printf("\n\n 3.Search");
        printf("\n\n 4.Display\n");
        printf("\n\n Enter your choice:");
        scanf("%d",&choice);
        if(choice==5)
            exit(0);
        switch(choice)
        {
            case 1:
                printf("\n\n Enter the element to be
inserted:");

                scanf("%d",&ele);
                x = insert(root,ele);
                if(root != NULL)
                {
                    splay(x,root);
                }
                root = x;
                break;
            case 2:
                if(root == NULL)
                {
                    printf("\n Empty tree...");
                    continue;
                }
                printf("\n\n Enter the element to be delete:");
                scanf("%d",&ele);
                root = delete(root,ele);
                break;
            case 3:
                printf("Enter the element to be search\n");
                scanf("%d",&ele);
                x = lookup(root,ele);

```

```

        splay(x, root);
        root = x;
        break;
    case 4:
        printf("The elements are\n");
        inorder(root);
        break;
    default:
        printf("Wrong choice\n");
        break;
    }
}

int data_print(struct node *x)
{
    if ( x==NULL )
        return 0;
    else
        return x->data;
}

/*some suggestion this code is not fully functional for example
if you have inserted some elements then try to delete root then it may
not work
because we are calling right and left child of a null value(parent of
root)
which is not allowed and will give segmentation fault

Also for inserting second element because of splaying twice(once in
insert and one in main)
will give error So I have made those changes but mainly in my cpp( c
plus plus file) file,
but I guess wiki will itself look into this and made these changes */

```

## Analysis

A simple amortized analysis of static splay trees can be carried out using the potential method. Suppose that  $\text{size}(r)$  is the number of nodes in the subtree rooted at  $r$  (including  $r$ ) and  $\text{rank}(r) = \log_2(\text{size}(r))$ . Then the potential function  $P(t)$  for a splay tree  $t$  is the sum of the ranks of all the nodes in the tree. This will tend to be high for poorly balanced trees, and low for well-balanced trees. We can bound the amortized cost of any zig-zig or zig-zag operation by:

$$\text{amortized cost} = \text{cost} + P(t_f) - P(t_i) \leq 3(\text{rank}_f(x) - \text{rank}_i(x)),$$

where  $x$  is the node being moved towards the root, and the subscripts "f" and "i" indicate after and before the operation, respectively. When summed over the entire splay operation, this telescopes to  $3(\text{rank}(\text{root}))$  which is  $O(\log n)$ . Since there's at most one zig operation, this only adds a constant.

## Performance theorems

There are several theorems and conjectures regarding the worst-case runtime for performing a sequence  $S$  of  $m$  accesses in a splay tree containing  $n$  elements.

Balance Theorem<sup>[1]</sup>

The cost of performing the sequence  $S$  is  $O[m(1 + \log n) + n \log n]$ . In other words, splay trees perform as well as static balanced binary search trees on sequences of at least  $n$  accesses.

Static Optimality Theorem<sup>[1]</sup>

Let  $q_i$  be the number of times element  $i$  is accessed in  $S$ . The cost of performing  $S$  is  $O\left[m + \sum_{i=1}^n q_i \log \frac{m}{q_i}\right]$ . In other words, splay trees perform as well as optimum static binary search trees on sequences of at least  $n$  accesses.

Static Finger Theorem<sup>[1]</sup>

Let  $i_j$  be the element accessed in the  $j^{\text{th}}$  access of  $S$  and let  $f$  be any fixed element (the finger). The cost of performing  $S$  is  $O\left[m + n \log n + \sum_{j=1}^m \log(|i_j - f| + 1)\right]$ .

Working Set Theorem<sup>[1]</sup>

Let  $t(j)$  be the number of distinct elements accessed between access  $j$  and the previous time element  $i_j$  was accessed. The cost of performing  $S$  is  $O\left[m + n \log n + \sum_{j=1}^m \log(t(j) + 1)\right]$ .

Dynamic Finger Theorem<sup>[4][5]</sup>

The cost of performing  $S$  is  $O\left[m + n + \sum_{j=1}^m \log(|i_{j+1} - i_j| + 1)\right]$ .

Scanning Theorem<sup>[6]</sup>

Also known as the **Sequential Access Theorem**. Accessing the  $n$  elements of a splay tree in symmetric order takes  $O(n)$  time, regardless of the initial structure of the splay tree. The tightest upper bound proven so far is  $4.5n$ .<sup>[7]</sup>

## Dynamic optimality conjecture

In addition to the proven performance guarantees for splay trees there is an unproven conjecture of great interest from the original Sleator and Tarjan paper. This conjecture is known as the *dynamic optimality conjecture* and it basically claims that splay trees perform as well as any other binary search tree algorithm up to a constant factor.

**Dynamic Optimality Conjecture:**<sup>[1]</sup> Let  $A$  be any binary search tree algorithm that accesses an element  $x$  by traversing the path from the root to  $x$  at a cost of  $d(x) + 1$ , and that between accesses can make any rotations in the tree at a cost of 1 per rotation. Let  $A(S)$  be the cost for  $A$  to perform the sequence  $S$  of accesses. Then the cost for a splay tree to perform the same accesses is  $O[n + A(S)]$ .

There are several corollaries of the dynamic optimality conjecture that remain unproven:

**Traversal Conjecture:**<sup>[1]</sup> Let  $T_1$  and  $T_2$  be two splay trees containing the same elements. Let  $S$  be the sequence obtained by visiting the elements in  $T_2$  in preorder (i.e., depth first search order). The total cost of performing the sequence  $S$  of accesses on  $T_1$  is  $O(n)$ .

**Deque Conjecture:**<sup>[8][6][9]</sup> Let  $S$  be a sequence of  $m$  double-ended queue operations (push, pop, inject, eject). Then the cost of performing  $S$  on a splay tree is  $O(m + n)$ .

**Split Conjecture:**<sup>[10]</sup> Let  $\mathcal{S}$  be any permutation of the elements of the splay tree. Then the cost of deleting the elements in the order  $\mathcal{S}$  is  $O(n)$ .

## References

- [1] Sleator, Daniel D.; Tarjan, Robert E. (1985), "Self-Adjusting Binary Search Trees" (<http://www.cs.cmu.edu/~sleator/papers/self-adjusting.pdf>), *Journal of the ACM (Association for Computing Machinery)* **32** (3): 652–686, doi:10.1145/3828.3835,
- [3] Allen, Brian; and Munro, Ian (1978), "Self-organizing search trees", *Journal of the ACM* **25** (4): 526–535, doi:10.1145/322092.322094
- [4] Cole, Richard; Mishra, Bud; Schmidt, Jeanette; and Siegel, Alan (2000), "On the Dynamic Finger Conjecture for Splay Trees. Part I: Splay Sorting log n-Block Sequences", *SIAM (Society for Industrial and Applied Mathematics) Journal on Computing* **30**: 1–43
- [5] Cole, Richard (2000), "On the Dynamic Finger Conjecture for Splay Trees. Part II: The Proof", *SIAM Journal on Computing* **30**: 44–85, doi:10.1137/S009753979732699X
- [6] Tarjan, Robert E. (1985), "Sequential access in splay trees takes linear time", *Combinatorica* **5** (4): 367–378, doi:10.1007/BF02579253
- [7] Elmasry, Amr (2004), "On the sequential access theorem and Deque conjecture for splay trees", *Theoretical Computer Science* **314** (3): 459–466, doi:10.1016/j.tcs.2004.01.019
- [8] Pettie, Seth (2008), "Splay Trees, Davenport-Schinzel Sequences, and the Deque Conjecture", *Proceedings of the 19th ACM-SIAM Symposium on Discrete Algorithms*: 1115–1124
- [9] Sundar, Rajamani (1992), "On the Deque conjecture for the splay algorithm", *Combinatorica* **12** (1): 95–124, doi:10.1007/BF01191208
- [10] Lucas, Joan M. (1991), "On the Competitiveness of Splay Trees: Relations to the Union-Find Problem", *Online Algorithms, Center for Discrete Mathematics and Theoretical Computer Science (DIMACS) Series in Discrete Mathematics and Theoretical Computer Science Vol. 7*: 95–124

## External links

- NIST's Dictionary of Algorithms and Data Structures: Splay Tree (<http://www.nist.gov/dads/HTML/splaytree.html>)
- Implementations in C and Java (by Daniel Sleator) (<http://www.link.cs.cmu.edu/link/ftp-site/splaying/>)
- Pointers to splay tree visualizations (<http://wiki.algoviz.org/AlgovizWiki/SplayTrees>)
- Fast and efficient implentation of Splay trees (<http://github.com/fbuihuu/libtree>)
- Top-Down Splay Tree Java implementation (<http://github.com/cpdomina/SplayTree>)
- Zipper Trees (<http://arxiv.org/abs/1003.0139>)

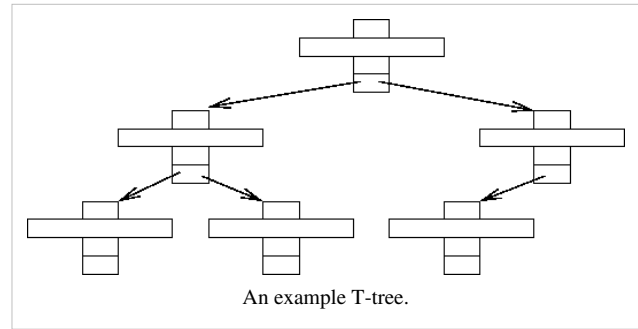
# T-tree

In computer science a **T-tree** is a type of binary tree data structure that is used by main-memory databases, such as Datablitz, *eXtremeDB*, MySQL Cluster, Oracle TimesTen and MobileLite.

A T-tree is a balanced index tree data structure optimized for cases where both the index and the actual data are fully kept in memory, just as a B-tree is an index structure optimized for storage on block oriented secondary storage devices like hard disks. T-trees seek to gain the performance benefits of in-memory tree structures such as AVL trees while avoiding the large storage space overhead which is common to them.

T-trees do not keep copies of the indexed data fields within the index tree nodes themselves. Instead, they take advantage of the fact that the actual data is always in main memory together with the index so that they just contain pointers to the actual data fields.

The "T" in T-tree refers to the shape of the node data structures in the original paper that first described this type of index.<sup>[1]</sup>



## Performance

Although T-trees seem to be widely used for main-memory databases, recent research indicates that they actually do not perform better than B-trees on modern hardware:

Rao, Jun; Kenneth A. Ross (1999). "Cache conscious indexing for decision-support in main memory" <sup>[2]</sup>. *Proceedings of the 25th International Conference on Very Large Databases (VLDB 1999)*. Morgan Kaufmann. pp. 78–89.

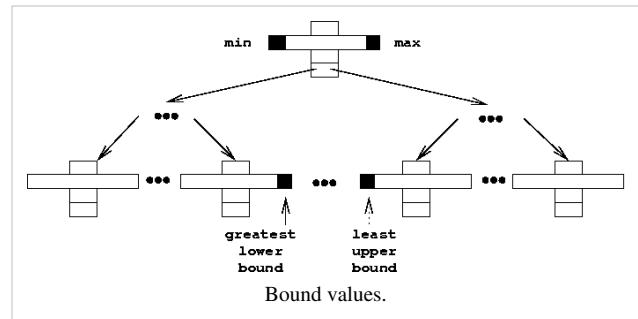
Kim, Kyungwha; Junho Shim, and Ig-hoon Lee (2007). "Cache conscious trees: How do they perform on contemporary commodity microprocessors?". *Proceedings of the 5th International Conference on Computational Science and Its Applications (ICCSA 2007)*. Springer. pp. 189–200. doi:10.1007/978-3-540-74472-6\_15.

The main reason seems to be that the traditional assumption of memory references having uniform cost is no longer valid given the current speed gap between cache access and main memory access.

## Node structures

A T-tree node usually consists of pointers to the parent node, the left and right child node, an ordered array of data pointers and some extra control data. Nodes with two subtrees are called *internal nodes*, nodes without subtrees are called *leaf nodes* and nodes with only one subtree are named *half-leaf nodes*. A node is called the *bounding node* for a value if the value is between the node's current minimum and maximum value, inclusively.

For each internal node, leaf or half leaf nodes exist that contain the predecessor of its smallest data value (called the *greatest lower bound*) and one that contains the successor of its largest data value (called the *least upper bound*). Leaf and half-leaf nodes can contain any number of data elements from one to the maximum size of the data array. Internal nodes keep their occupancy between predefined minimum and maximum numbers of elements



## Algorithms

### Search

- Search starts at the root node
- If the current node is the bounding node for the search value then search its data array. Search fails if the value is not found in the data array.
- If the search value is less than the minimum value of the current node then continue search in its left subtree. Search fails if there is no left subtree.
- If the search value is greater than the maximum value of the current node then continue search in its right subtree. Search fails if there is no right subtree.

### Insertion

- Search for a bounding node for the new value. If such a node exist then
  - check whether there is still space in its data array, if so then insert the new value and finish
  - if no space is available then remove the minimum value from the node's data array and insert the new value. Now proceed to the node holding the greatest lower bound for the node that the new value was inserted to. If the removed minimum value still fits in there then add it as the new maximum value of the node, else create a new right subnode for this node.
- If no bounding node was found then insert the value into the last node searched if it still fits into it. In this case the new value will either become the new minimum or maximum value. If the value doesn't fit anymore then create a new left or right subtree.

If a new node was added then the tree might need to be rebalanced, as described below.

## Deletion

- Search for bounding node of the value to be deleted. If no bounding node is found then finish.
- If the bounding node does not contain the value then finish.
- delete the value from the node's data array

Now we have to distinguish by node type:

- Internal node:

If the node's data array now has less than the minimum number of elements then move the greatest lower bound value of this node to its data value. Proceed with one of the following two steps for the half leaf or leaf node the value was removed from.

- Leaf node:

If this was the only element in the data array then delete the node. Rebalance the tree if needed.

- Half leaf node:

If the node's data array can be merged with its leaf's data array without overflow then do so and remove the leaf node. Rebalance the tree if needed.

## Rotation and balancing

A T-tree is implemented on top of an underlying self-balancing binary search tree. Specifically, Lehman and Carey's article describes a T-tree balanced like an AVL tree: it becomes out of balance when a node's child trees differ in height by at least two levels. This can happen after an insertion or deletion of a node. After an insertion or deletion, the tree is scanned from the leaf to the root. If an imbalance is found, one tree rotation or pair of rotations is performed, which is guaranteed to balance the whole tree.

When the rotation results in an internal node having fewer than the minimum number of items, items from the node's new child(ren) are moved into the internal node.

## References

- [1] Tobin J. Lehman and Michael J. Carey, A Study of Index Structures for Main Memory Database Management Systems. VLDB 1986 (<http://www.vldb.org/conf/1986/P294.PDF>)
- [2] <http://www.vldb.org/dblp/db/conf/vldb/RaoR99.html>

## External links

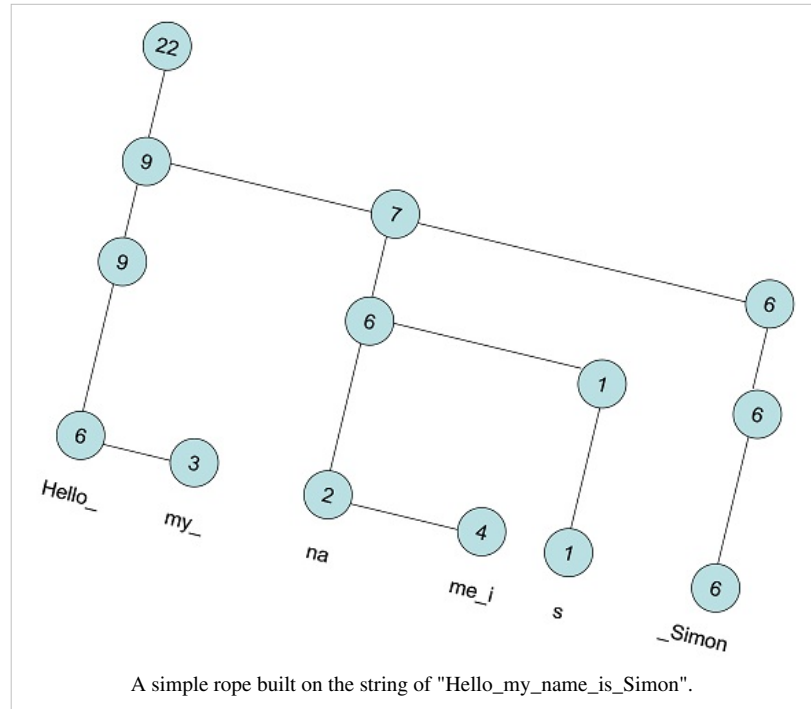
- Oracle TimesTen FAQ entry on index mentioning T-Trees ([http://www.oracle.com/technology/products/timesten/htdocs/faq/technical\\_faq.html##6](http://www.oracle.com/technology/products/timesten/htdocs/faq/technical_faq.html##6))
- Oracle Whitepaper: Oracle TimesTen Products and Technologies ([http://www.oracle.com/technology/products/timesten/pdf/wp/timesten\\_tech\\_wp\\_dec\\_2005.pdf](http://www.oracle.com/technology/products/timesten/pdf/wp/timesten_tech_wp_dec_2005.pdf))
- DataBlitz presentation mentioning T-Trees (<http://www.dependability.org/wg10.4/timedepend/08-Rasto.pdf>)
- An Open-source T\*-tree Library (<http://code.google.com/p/ttree/>)

# Rope

In computer programming a **rope**, or **cord**, is a data structure for efficiently storing and manipulating a very long string. For example, a text editing program may use a rope to represent the text being edited, so that operations such as insertion, deletion, and random access can be done efficiently.<sup>[1]</sup>

## Description

A rope is a binary tree. Leaf nodes (as well as some single-child internal nodes) contain a short string. Each node has a "weight" equal to the length of its string plus the sum of all the weights in its left subtree. Thus a node with two children divides the whole string into two parts: the left subtree stores the first part of the string. The right subtree stores the second part and its weight is the sum of the two parts.



The binary tree can be seen as several levels of nodes. The bottom level contains all the nodes that contain a string. Higher levels have fewer and fewer nodes. The top level consists of a single "root" node. The rope is built by putting the nodes with short strings in the bottom level, then attaching a random half of the nodes to parent nodes in the next level. Nodes with no parent (for example, the two nodes storing the strings "my\_" and "me\_i" in the diagram above) become the right child of the node located immediately to their left, thus forming a chain.

## Operations

### Index

Definition: Index(i): return the character at position i

Time complexity:  $O(\log N)$  where N is the length of the rope

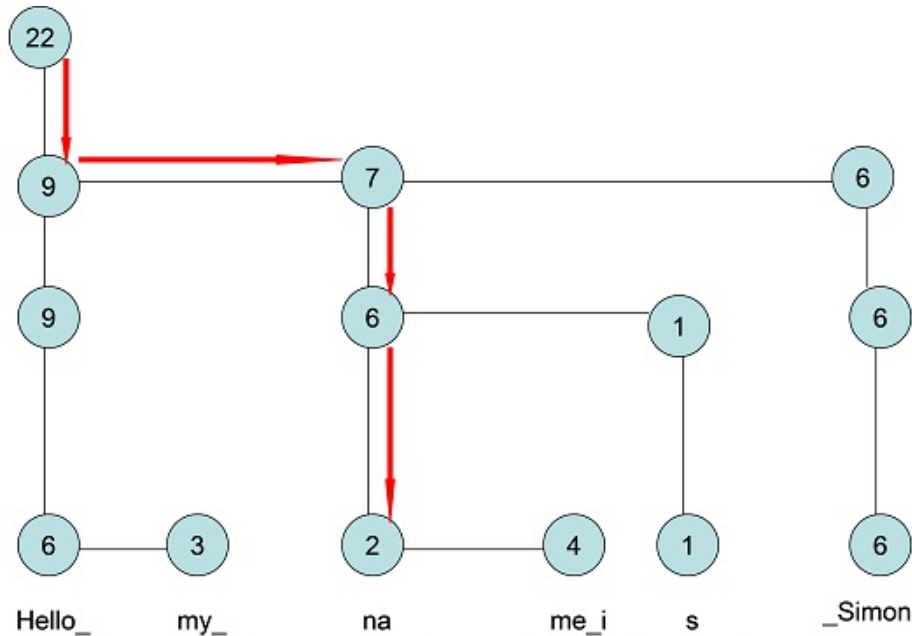
To retrieve the i-th character, we begin a recursive search from the root node:

```
// Note: Assumes 1-based indexing.
function index(RopeNode node, integer i)
  if node.weight < i then
    return index(node.right, i - node.weight)
  else
    if exists(node.left) then
      return index(node.left, i)
    else
      return node.string[i]
    endif
  endif
endif
```



**end**

For example, to find the character at  $i=10$  in the following rope, start at the root node, find that 22 is greater than 10 and there is a left child, so go to the left child. 9 is less than 10, so subtract 9 from 10 (leaving  $i=1$ ) and go to the right child. Then because 7 is greater than 1 and there's a left child, go to the left child. 6 is greater than 1 and there's a left child, so go to the left child again. Finally 2 is greater than 1 but there is no left child, so the character at index 1 of the short string "na", is the answer.

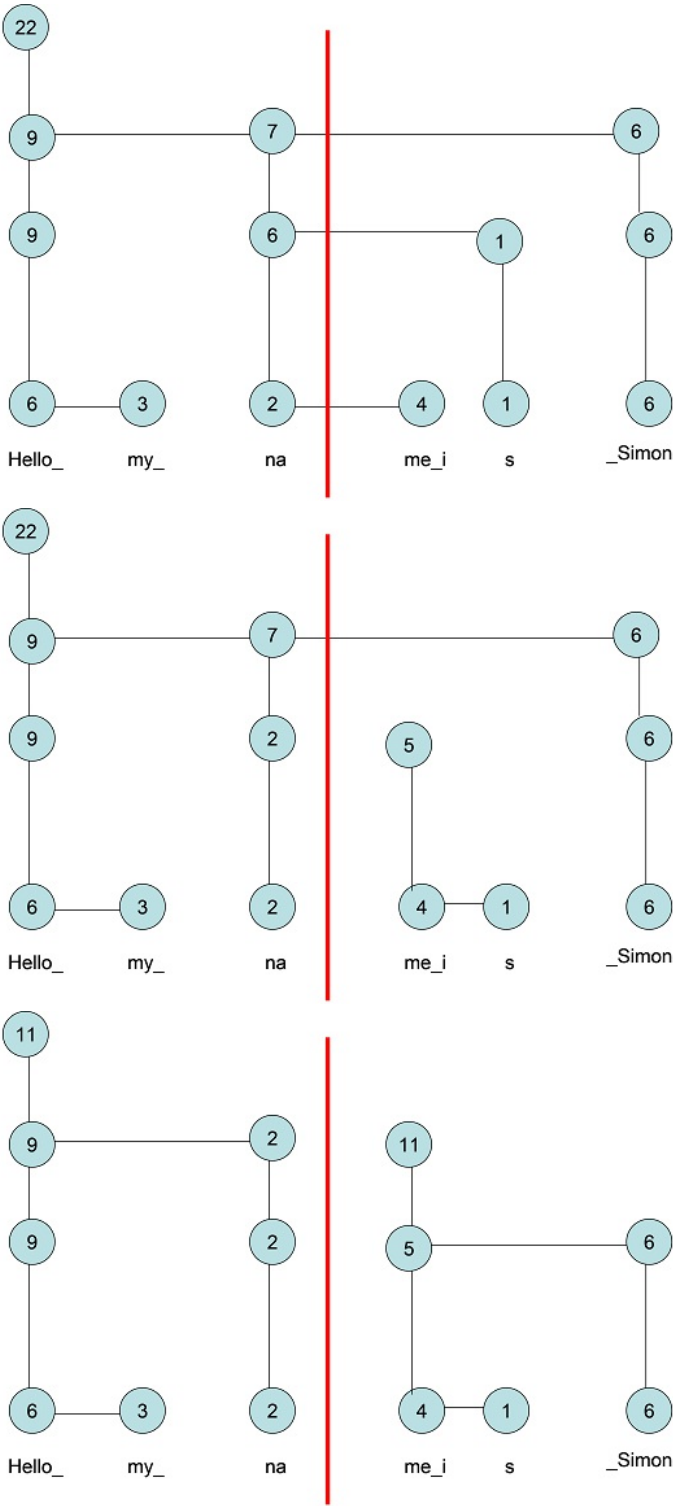


## Split

Definition: Split ( $i, S$ ): split the string  $S$  into two new strings  $S1$  and  $S2$ ,  $S1 = C1, \dots, Ci$  and  $S2 = Ci + 1, \dots, Cm$ .  
Time complexity:  $O(\log N)$

There are two cases: in the first, the  $i$ -th character is at the end of an array such as in the following picture; in the second, the character is in the middle of an array. The second case reduces to the first case as follows: split the node at the character into two nodes each with part of the array and make the second node the right child of the first node.

For example, to split the pictured rope into two parts, query the  $i$ -th character to locate the node  $v$  at the bottom level. Remove the link between  $v$  and the right child of  $v$ , or  $v'$ . Go to the parent  $u$  and subtract the weight of  $v'$  from the weight of  $u$ . Since the parent has the right child of  $u$ , now  $u'$ , modify  $u'$  to link to  $v'$  and increase the weight of  $u'$  by the weight of  $v'$ . The former left child of  $u'$  becomes the right child of  $v'$ , creating the second picture. Continue up to the parent of  $u$ ,  $w$ . Subtract the weight of  $v'$  from the weight of  $w$ . Then modify the right child of  $w$ , now  $w'$ , to link to  $u'$ . The former child of  $w'$  becomes the right child of  $u'$ . Increase the weight of  $w'$  by the weight of  $v'$ . Move to the parent of  $w$ ,  $x$ . Since  $w$  is already the right child of  $x$ , there is no change. Then go to the parent of  $x$ ,  $y$ , and reduce the weight of  $y$  by the weight of  $w'$ .



## Concat

Definition:  $\text{Concat}(S1, S2)$ : concatenate two ropes  $S1, S2$  into a single rope. Time complexity:  $O(\log N)$

This operation is the reverse of split. Alternatively, create a new root node with  $\text{left}=S1$  and  $\text{right}=S2$ . This is constant time, but can lead to an unbalanced tree.

## Insert

Definition:  $\text{Insert}(i, S')$ : insert the string  $S'$  beginning at position  $i$  in the string  $s$ , to form a new string  $C1, \dots, Ci, S', Ci + 1, \dots, Cm$ . Time complexity:  $O(\log N)$ .

This operation can be done by a  $\text{split}()$  and a  $\text{concat}()$ . The cost is the sum of the two.

## Delete

Definition:  $\text{Delete}(i, j)$ : delete the substring  $Ci, \dots, Ci + j - 1$ , from  $s$  to form a new string  $C1, \dots, Ci - 1, Ci + j, \dots, Cm$ . Time complexity:  $O(\log N)$ .

This operation can be done by two  $\text{split}()$  and a  $\text{concat}()$ . First, split the rope in three, divided by  $i$ -th and  $j$ -th character respectively, putting the string to delete in a separate node. Then concatenate the other two nodes.

## Report

Definition:  $\text{Report}(i, j)$ : output the string  $Ci, \dots, Ci + j - 1$ . Time complexity:  $O(j + \log N)$

To report the string  $Ci, \dots, Ci + j - 1$ , find the node  $u$  that contains  $ci$  and  $\text{weight}(u) \geq j$ , and then traverse  $T$  starting at node  $u$ . Output  $Ci, \dots, Ci + j - 1$  by doing an in-order traversal of  $T$  starting at node  $u$ .

## Comparison with monolithic arrays

Advantages:

- Ropes enable much faster insertion and deletion of text than monolithic string arrays, on which operations have time complexity  $O(n)$ .
- Ropes don't require the extra  $O(n)$  memory that arrays need for copying operations, and ropes don't require large contiguous memory spaces.

Disadvantages:

- Slightly greater overall space usage (mainly to store parent nodes)
- Increase in time to manage the extra storage

This table compares the *algorithmic* characteristics of string and rope implementations, not their "raw speed". Array-based strings have smaller overhead, so (for example) concatenation and split operations are faster on small datasets. However, when array-based strings are used for longer strings, time complexity and memory usage for insertion and deletion of characters become unacceptably large. A rope data structure, on the other hand, has stable performance regardless of data size. Moreover, the space complexity for ropes and arrays are both  $O(n)$ . In summary, ropes are better suited when the data is large and frequently modified.

## Performance

Operation	Rope	String
Index <sup>[1]</sup>	$O(\log n)$	$O(1)$
Split <sup>[1]</sup>	$O(\log n)$	$O(1)$
Concatenate	$O(\log n)$	$O(n)$
Iterate over each character <sup>[1]</sup>	$O(n)$	$O(n)$
Insert	$O(\log n)$	$O(n)$
Delete	$O(\log n)$	$O(n)$
Report	$O(j + \log n)$	$O(j)$
Build	$O(n)$	$O(n)$

## References

- [1] Boehm, Hans-J; Atkinson, Russ; and Plass, Michael (December 1995). "Ropes: an Alternative to Strings" (<http://citeseer.ist.psu.edu/viewdoc/download?doi=10.1.1.14.9450&rep=rep1&type=pdf>) (PDF). *Software—Practice & Experience* (New York, NY, USA: John Wiley & Sons, Inc.) **25** (12): 1315–1330. doi:10.1002/spe.4380251203. .

## External links

- SGI's implementation of ropes for C++ (<http://www.sgi.com/tech/stl/Rope.html>)
- libstdc++ support for ropes (<http://gcc.gnu.org/onlinedocs/libstdc++/libstdc++-html-USERS-4.3/a00223.html>)
- Ropes for Java (<http://ahmadsoft.org/ropes/>)
- Ropes (<http://rope.forge.ocamlcore.org/doc/Rope.html>) for OCaml
- ropes (<https://github.com/Ramarren/ropes>) for Common Lisp

# Top Trees

A **Top tree** is a data structure based on a binary tree for unrooted dynamic trees that is used mainly for various path-related operations. It allows simple divide-and-conquer algorithms. It has since been augmented to maintain dynamically various properties of a tree such as diameter, center and median.

A Top tree  $\mathfrak{T}$  is defined for an *underlying tree*  $\mathcal{T}$  and a pair of vertices  $\partial\mathcal{T}$  called as External Boundary Vertices

## Glossary

### Boundary Node

See Boundary Vertex

### Boundary Vertex

A vertex in a connected subtree is a *Boundary Vertex* if it is connected to a vertex outside the subtree by an edge.

### External Boundary Vertices

Up to a pair of vertices in the Top Tree  $\mathfrak{T}$  can be called as External Boundary Vertices, they can be thought of as Boundary Vertices of the cluster which represents the entire Top Tree.

### Cluster

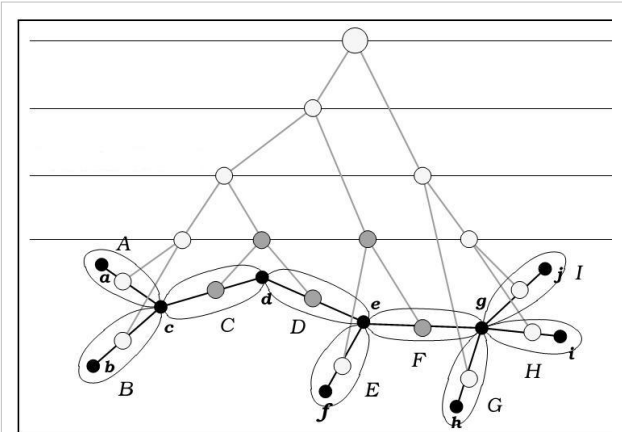
A *cluster* is a connected subtree with at most two Boundary Vertices. The set of Boundary Vertices of a given cluster  $\mathcal{C}$  is denoted as  $\partial\mathcal{C}$ . With each cluster  $\mathcal{C}$  the user may associate some meta information  $Info(\mathcal{C})$ , and give methods to maintain it under the various internal operations.

### Path Cluster

If  $\pi(\mathcal{C})$  contains at least one edge then  $\mathcal{C}$  is called a *Path Cluster*.

### Point Cluster

See Leaf Cluster



An image depicting a Top tree built on an underlying tree (black nodes) A tree divided into edge clusters and the complete top-tree for it. Filled nodes in the top-tree are path-clusters, while small circle nodes are leaf-clusters. The big circle node is the root. Capital letters denote clusters, non-capital letters are nodes.

**Leaf Cluster**

If  $\pi(\mathcal{C})$  does not contain any edge i.e.  $\mathcal{C}$  has only one Boundary Vertex then  $\mathcal{C}$  is called a *Leaf Cluster*.

**Edge Cluster**

A Cluster containing a single edge is called an *Edge Cluster*.

**Leaf Edge Cluster**

A Leaf in the original Cluster is represented by a Cluster with just a single Boundary Vertex and is called a *Leaf Edge Cluster*.

**Path Edge Cluster**

Edge Clusters with two Boundary Nodes are called *Path Edge Cluster*.

**Internal Node**

A node in  $\mathcal{C} \setminus \partial\mathcal{C}$  is called an *Internal Node* of  $\mathcal{C}$ .

**Cluster Path**

The path between the Boundary Vertices of  $\mathcal{C}$  is called the *cluster path* of  $\mathcal{C}$  and it is denoted by  $\pi(\mathcal{C})$ .

**Mergeable Clusters**

Two Clusters  $\mathcal{A}$  and  $\mathcal{B}$  are *Mergeable* if  $\mathcal{A} \cap \mathcal{B}$  is a singleton set (they have exactly one node in common) and  $\mathcal{A} \cup \mathcal{B}$  is a Cluster.

**Introduction**

*Top Trees* are used for maintaining a Dynamic forest (set of trees) under link and cut operations.

The basic idea is to maintain a balanced Binary tree  $\mathfrak{R}$  of logarithmic height in the number of nodes in the original tree  $\mathcal{T}$  ( i.e. in  $\mathcal{O}(\log n)$  time) ; the **Top Tree** essentially represents the recursive subdivision of the original tree  $\mathcal{T}$  into clusters.

In general the tree  $\mathcal{T}$  may have weight on its edges.

There is a one to one correspondence with the edges of the original tree  $\mathcal{T}$  and the leaf nodes of the Top Tree  $\mathfrak{R}$  and each internal node of  $\mathfrak{R}$  represents a cluster that is formed due to the union of the clusters that are its children.

The Top Tree data structure can be initialized in  $\mathcal{O}(n)$  time.  $\mathcal{T}$

Therefore the Top Tree  $\mathfrak{R}$  over  $(\mathcal{T}, \partial\mathcal{T})$  is a binary tree such that

- The nodes of  $\mathfrak{R}$  are clusters of  $(\mathcal{T}, \partial\mathcal{T})$ ;
- The leaves of  $\mathfrak{R}$  are the edges of  $\mathcal{T}$ ;
- Sibling clusters are neighbours in the sense that they intersect in a single vertex, and then their parent cluster is their union.
- Root of  $\mathfrak{R}$  is the tree  $\mathcal{T}$  itself, with a set of at most two External Boundary Vertices.

A tree with a single node has an empty top tree, and one with just an edge is just a single node.

These trees are freely augmentable allowing the user a wide variety of flexibility and productivity without going into the details of the internal workings of the data structure, something which is also referred to as the *Black Box*.

## Dynamic Operations

The following two are the user allowable Forest Updates.

- **Link( $v, w$ ):** Where  $v$  and  $w$  are nodes in different trees  $T_1$  and  $T_2$ . It returns a single top tree representing  $\mathfrak{R}_v \cap \mathfrak{R}_w \cap (v, w)$
- **Cut( $v, w$ ):** Removes the Edge  $(v, w)$  from a tree  $T$  with Top Tree  $\mathfrak{R}$ , thereby turning it into two trees  $T_v$  and  $T_w$  and returning two Top Trees  $\mathfrak{R}_v$  and  $\mathfrak{R}_w$ .

**Expose( $v, w$ ):** Is called as a subroutine for implementing most of the path related queries on a Top Tree. It makes  $v$  and  $w$  the External Boundary Vertices of the Top Tree and returns the new Root cluster.

## Internal Operations

The Forest updates are all carried out by a sequence of at most  $\mathcal{O}(\log n)$  Internal Operations, the sequence of which is computed in further  $\mathcal{O}(\log n)$  time.

- **Merge( $\mathcal{A}, \mathcal{B}$ ):** Here  $\mathcal{A}$  and  $\mathcal{B}$  are *Mergeable Clusters*, it returns  $\mathcal{C}$  as the parent cluster of  $\mathcal{A}$  and  $\mathcal{B}$  and with boundary vertices as the boundary vertices of  $\mathcal{A} \cup \mathcal{B}$ . Updates to  $\text{Info}(\mathcal{C})$  are carried out accordingly.
- **Split( $\mathcal{C}$ ):** Here  $\mathcal{C}$  is  $\mathcal{A} \cup \mathcal{B}$ . This deletes the cluster  $\mathcal{C}$  from  $\mathfrak{R}$  and methods are then called to update  $\text{Info}(\mathcal{A})$  and  $\text{Info}(\mathcal{B})$ .

The next two functions are analogous to the above two and are used for base clusters.

- **Create( $v, w$ ):** Creates a cluster  $\mathcal{C}$  for the edge  $(v, w)$ . Sets  $\partial\mathcal{C} = \partial(v, w)$ . Methods are then called to compute  $\text{Info}(\mathcal{C})$ .
- **Eradicate( $\mathcal{C}$ ):**  $\mathcal{C}$  is the edge cluster  $(v, w)$ . It deletes the cluster  $\mathcal{C}$  from the top tree. The  $\text{Info}(\mathcal{C})$  is stored by calling a user defined function, as it may also happen that during a tree update, a leaf cluster may change to a path cluster and the converse.

## Interesting Results and Applications

A number of interesting applications have been derived for these Top Trees some of them include

- ([SLEATOR AND TARJAN 1983]). We can maintain a dynamic collection of weighted trees in  $\mathcal{O}(\log n)$  time per link and cut, supporting queries about the maximum edge weight between any two vertices in  $\mathcal{O}(\log n)$  time.
  - Proof outline: It involves maintaining at each node the maximum weight ( $\text{max\_wt}$ ) on its cluster path, if it is a point cluster then  $\text{max\_wt}(\mathcal{C})$  is initialised as  $-\infty$ . When a cluster is a union of two clusters then it is the maximum value of the two merged clusters. If we have to find the max wt between  $v$  and  $w$  then we do  $\mathcal{C} = \text{Expose}(v, w)$ , and report  $\text{max\_wt}(\mathcal{C})$ .
- ([SLEATOR AND TARJAN 1983]). In the scenario of the above application we can also add a common weight  $x$  to all edges on a given path  $v \cdots w$  in  $\mathcal{O}(\log n)$  time.
  - Proof outline: We introduce a weight called  $\text{extra}(\mathcal{C})$  to be added to all the edges in  $\pi(\mathcal{C})$ . Which is maintained appropriately;  $\text{split}(\mathcal{C})$  requires that, for each path child  $\mathcal{A}$  of  $\mathcal{C}$ , we set  $\text{max\_wt}(\mathcal{A}) := \text{max\_wt}(\mathcal{A}) + \text{extra}(\mathcal{C})$  and  $\text{extra}(\mathcal{A}) := \text{extra}(\mathcal{A}) + \text{extra}(\mathcal{C})$ . For  $\mathcal{C} := \text{join}(\mathcal{A}, \mathcal{B})$ , we set  $\text{max\_wt}(\mathcal{C}) := \max\{\text{max\_wt}(\mathcal{A}), \text{max\_wt}(\mathcal{B})\}$  and  $\text{extra}(\mathcal{C}) := 0$ . Finally, to find the maximum weight on the path  $v \cdots w$ , we set  $\mathcal{C} := \text{Expose}(v, w)$  and return  $\text{max\_wt}(\mathcal{C})$ .
- ([GOLDBERG ET AL. 1991]). We can ask for the maximum weight in the underlying tree containing a given vertex  $v$  in  $\mathcal{O}(\log n)$  time.
  - Proof outline: This requires maintaining additional information about the maximum weight non cluster path edge in a cluster under the Merge and Split operations.
- The distance between two vertices  $v$  and  $w$  can be found in  $\mathcal{O}(\log n)$  time as  $\text{length}(\text{Expose}(v, w))$ .

- Proof outline: We will maintain the length  $\text{length}(\mathcal{C})$  of the cluster path. The length is maintained as the maximum weight except that, if  $\mathcal{C}$  is created by a join(Merge),  $\text{length}(\mathcal{C})$  is the sum of lengths stored with its path children.
- Queries regarding diameter of a tree and its subsequent maintenance takes  $\mathcal{O}(\log n)$  time.
- The Center and Median can be maintained under Link(Merge) and Cut(Split) operations in  $\mathcal{O}(\log n)$  time.

## Implementation

Top Trees have been implemented in a variety of ways, some of them include implementation using a *Multilevel Partition* (Top-trees and dynamic graph algorithms Jacob Holm and Kristian de Lichtenberg. Technical Report), and even by using Sleator-Tarjan s-t trees, Fredericksons Topology Trees (Alstrup et al. Maintaining Information in Fully Dynamic Trees with Top Trees).

## Using Multilevel Partitioning

Any partitioning of clusters of a tree  $\mathcal{T}$  can be represented by a Cluster Partition Tree CPT ( $\mathcal{T}$ ), by replacing each cluster in the tree  $\mathcal{T}$  by an edge. If we use a strategy  $P$  for partitioning  $\mathcal{T}$  then the CPT would be  $\text{CPT}_P \mathcal{T}$ . This is done recursively till only one edge remains.

We would notice that all the nodes of the corresponding Top Tree  $\mathfrak{R}$  are uniquely mapped into the edges of this multilevel partition. There may be some edges in the multilevel partition that do not correspond to any node in the Top tree, these are the edges which represent only a single child in the level below it, i.e. a simple cluster. Only the edges that correspond to composite clusters correspond to nodes in the Top Tree  $\mathfrak{R}$ .

A Partitioning Strategy is important while we partition the Tree  $\mathcal{T}$  into clusters. Only a careful strategy ensures that we end up in an  $\mathcal{O}(\log n)$  height Multilevel Partition (and therefore the Top Tree).

- The number of edges in subsequent levels should decrease by a constant factor.
- If a lower level is changed by an update then we should be able to update the one immediately above it using at most a constant number of insertions and deletions.

The above partitioning strategy ensures the maintenance of the Top Tree in  $\mathcal{O}(\log n)$  time.

## References

- Stephen Alstrup, Jacob Holm, Kristian De Lichtenberg, and Mikkel Thorup, *Maintaining information in fully dynamic trees with top trees*, ACM Transactions on Algorithms (TALG), Vol. 1 (2005), 243–264, doi:10.1145/1103963.1103966
- Donald Knuth. *The Art of Computer Programming: Fundamental Algorithms*, Third Edition. Addison-Wesley, 1997. ISBN 0-201-89683-4 . Section 2.3: Trees, pp. 308–423.
- Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7 . Section 10.4: Representing rooted trees, pp. 214–217. Chapters 12–14 (Binary Search Trees, Red-Black Trees, Augmenting Data Structures), pp. 253–320.



## External links

- Maintaining Information in Fully Dynamic Trees with Top Trees. Alstrup et al <sup>[1]</sup>
- Self Adjusting Top Trees. Tarjan and Werneck <sup>[2]</sup>
- Self-Adjusting Top Trees. Tarjan and Werneck, Proc. 16th SoDA, 2005 <sup>[3]</sup>

## References

- [1] <http://arxiv.org/abs/cs.DS/0310065>  
 [2] <http://www.cs.princeton.edu/~rwerneck/docs/TW05.htm>  
 [3] <http://portal.acm.org/citation.cfm?id=1070547&dl=&coll=&CFID=15151515&CFTOKEN=6184618>

# Tango Trees

A **Tango tree** is an online binary search tree that is  $O(\log \log n)$ -competitive proposed by Erik D. Demaine, Dion Harmon, John Iacono, and Mihai Patrascu in 2004.

## Overview

Tango trees were designed to surpass the usual  $O(\log n)$  binary search tree cost of operations. They perform basic operations such as searches in  $O(\log \log n)$  time. This optimization is achieved dynamically by adjusting the search tree structure after each search. They are similar in their dynamic behaviour to other types of structure like a Splay tree however the competitive ratio is dramatically improved.

The approach is similar to the Greedy BST algorithm that while searching for an element rearranges the nodes on the search path to minimize the cost of future searches.

For Tango Trees the approach is a classic divide and conquer approach combined with a bring to top approach.

The main divide and conquer idea behind this data structure is to extract from the original tree a number of virtual smaller subtrees all with a normal  $O(\log \text{ number of subtree elements})$  cost of access. These subtrees are dynamically balanced to offer the usual  $O(\log n)$  performance for data retrieval.

The bring to top approach is not done at the node level as much as at the subtree level which further improve competitiveness. Once the original tree has been adjusted to include a collection of these subtrees, it is possible to greatly improve the cost of access of these subtrees. Both the Tango tree and these subtrees are a type of Self-balancing binary search tree.

Tango tree achieves this outstanding competitive ratio by using a combination of augmentation of attributes in the data structure, a more elaborated algorithm and the use of other type of trees for some of its structure.

## Example

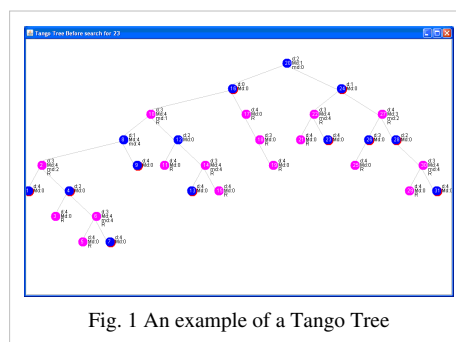


Fig. 1 An example of a Tango Tree

## Similar Data Structures

- Red Black tree, introduced by Bayer in 1972, having a  $O(\log n)$  competitive ratio
- Splay tree, introduced by Sleator and Tarjan in 1985, having a  $O(\log n)$  competitive ratio
- AVL tree, introduced by Adelson and Landis in 1962, having a  $O(\log n)$  competitive ratio
- Multi-splay tree, introduced by Sleator and Wang in 2006, having a  $O(\log \log n)$  competitive ratio

## Advantages

Tango Trees offer unsurpassed competitive ratio retrieval for online data. Online data means that operations that are not known in advance before the data structure is created.

Outstanding search performance for a Tango tree relies on the fact that accessing nodes constantly updates the structure of the search trees. That way the searches are rerouted to searches in much shallower balanced trees.

Obviously, significantly faster access time constitutes an advantage for nearly all practical applications that offer searches as a use case. Dictionary searches like telephone directory search would be just one of the possible examples.

## Disadvantages

The Tango tree focuses on data searches on static data structures, and does not support deletions or insertions, so it might not be appropriate in every situation.

The Tango tree uses augmentation, meaning storing more data in a node than in a node of a plain binary search tree. Tango trees use  $O(\log \log n)$  bits. Although that is not a significant increase, it results in a bigger memory footprint.

It is a complex algorithm to implement like for instance splay tree, and it also makes use of rarely used operations of Red-Black Tree.

Tango trees change when they are accessed in a 'read-only' manner (i.e. by *find* operations). This complicates the use of Tango trees in a multi-threaded environment.

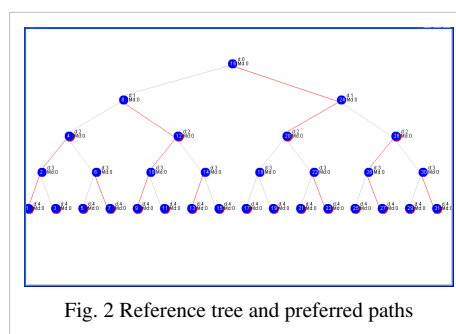
It is believed that Tango Tree would work in a practical situation where a very large data set with strong spatial and temporal coherence fits in the memory.

## Terminology and Concepts

There are several types of trees besides the Red-Black trees (RB) used as a base for all Tree structures:

### Reference Trees

Example of reference tree:

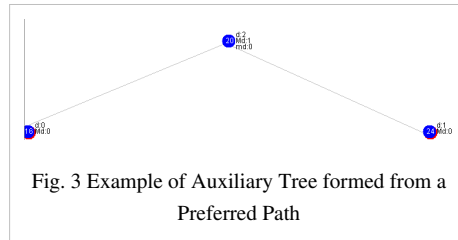


## Tango Trees

See Fig 1 for an example of Tango tree

## Auxiliary Trees

Example of auxiliary tree:



As all trees are derived from RB trees so they are also [Binary Search Trees] with all their inherent behaviour.

Auxiliary trees can be considered sub-trees of the Tango Tree. Tango Trees are the actual employed trees while in production mode.

Reference Trees are used only initial set-up and for illustration of the concepts.

Any search in the Reference Tree creates a path from root to the searched node. We call that a Preferred Path and the Preferred Child attribute specific to the Reference Tree indicates if the preferred path of a node goes to the left or right child if any. A Preferred Path is determined by the longest path formed by preferred children. Any new search in the Reference Tree will carve new paths and modify the existing paths. Correspondingly, the preferred children change too.

Any switch from right to left or vice versa is called an Interleave. Interleaves changes are the basis for analysis of expected performance.

## Operations

As we stated Tango Trees are static so they support only searches. That also means that there is a construction phase where the elements are inserted in the Tango Tree. That start-up cost and any search performance during the construction period is not considered part of the operational part of Tango trees therefore the performance is not  $O(\log \log n)$  competitive. The outstanding idea behind Tango Trees is to collect the nodes belonging to a Preferred Path as a balanced tree of height  $O(\log \log n)$  called auxiliary tree and then assemble them in a tree of trees where higher trees contain the mostly accessed preferred paths elements.

## Search

To search for a node  $x$  in a Tango tree, we search for the element in the collection of Auxiliary Trees that make up the Tango Tree like in any ordinary binary search tree. Simultaneously we adjust the corresponding affected Auxiliary Trees in the Tango Tree. This will preserve the ideal structure that will give us this unprecedented search performance. We could achieve this ideal structure by updating the Reference Tree  $P$  after every search and recreating the Tango tree however this would be very expensive and nonperforming. The reasons why such a structure is  $O(\log \log n)$  competitive is explained in the Analysis There is a direct way to update the structure and that is shown in the [Algorithm]

## Tango Tree Life Cycle

The main phases are Construction and Operation

### Construction

First create the reference tree and insert the desired data in it. Update the attributes of depth for each node. After this phase the data and the value of the depth for the nodes will remain unchanged. Let's call that field  $d$  for further reference and understand that it always refers to the Reference tree not to the Tango tree as that can cause confusions. While in principle the reference tree can be any balanced tree that is augmented with the depth of each node in the tree the TODO [Demaine et al. 2004] uses [red-black tree]. Secondly we will perform some warm-up searches with the goal of creating a decent distribution of Preferred Paths in the Reference Tree. Remember there is no Tango tree yet and all this is done on line. This means that performance is not critical at this point.

After this begins the phase of collecting the preferred paths. Out of each Preferred Path we create a new Auxiliary Tree which is just an ordinary RedBlack Tree where the nodes inherit the value of field  $d$ . That value will stay unchanged forever because it stays with the node even if the node is moved in the trees. There is no Tango Tree at this point. We add auxiliary trees in such a way that the largest one is the top of a tree and the rest 'hung' below it. This way we effectively create a forest where each tree is an Auxiliary tree. See Fig. 1 where the roots of the composing auxiliary tree are depicted by magenta nodes. It's after this step that the Tango tree becomes operational. See [Construction Algorithms And Pseudo-code] for main algorithms and pseudo-code.

### Operation

The operation phase is the main phase where we perform searches in the Tango tree. See [Operation Algorithms And Pseudo-code] for main algorithms And pseudo-code.

## Data Augmentation

### Reference Tree augmentation

Besides the regular RB Tree fields we introduce two more fields:

- `isPreferredChild` representing the direction the last search traversing the node took. It could be either Boolean or a pointer to another node. In the figures it is represented by a red line between a node and its preferred child.
- $d$  representing the depth of the node in the Reference Tree. (See value of  $d$  in Fig 2.)

### Tango Tree Augmentation

For each nodes in Tango or Auxiliary Tree we also introduce several new fields:

`isRoot` that will be false for all nodes except for the root. This is depicted in magenta for nodes where `isRoot` is true.

`maxD` that is the Maximum value of  $d$  for all the children of the node. This is depicted as  $Md$  in the figures. For some nodes this value is undefined and the field is not depicted in the figure

`minD` that is the Minimum value of  $d$  for all the children of the node. This is depicted as  $md$  in the figures. For some nodes this value is undefined and the field is not depicted in the figure

`isRoot`, `maxD` and `minD` will change in time when nodes are moved around in the structure.

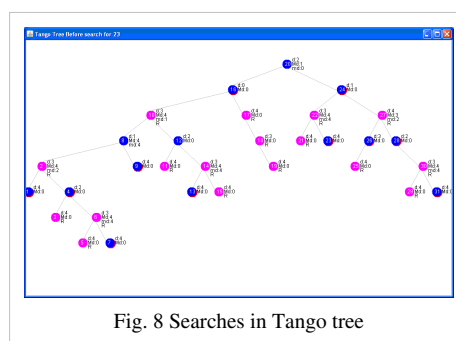
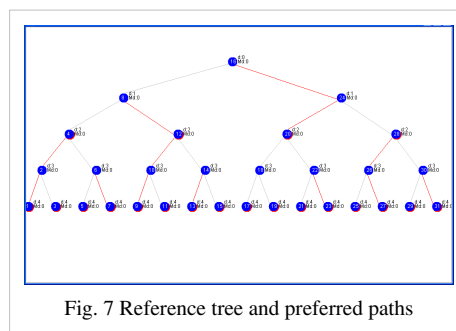
## Algorithms

The main task in a Tango Tree is to maintain an 'ideal' structure mirroring changes that occur in the reference tree. As recreating the Tango tree from the reference tree would not be performing the algorithms will have only use and modify the Tango tree. That means that after the construction phase the reference tree could and should be destroyed. After this we would refer to it as virtual meaning 'as if it existed'. As described in [Demaine et al. 2004], the main goal is to maintain the ideal Tango structure that would mimic preferred paths changes on the virtual reference tree. So the purpose of the Tango algorithm is the constructing of a new state  $T_i$  of the Tango Tree based on the previous state  $T_{i-1}$  or the Tango Tree and the new search of  $x_i$ .

## Tango Search

During the search for  $x_i$ , one or many auxiliary trees could be traversed. During this Tango walk phase of the Tango search every time when we are crossing from an auxiliary tree in a new auxiliary tree we perform exactly one cut operation on the tree just traversed.

The result of the Cut is then Joined with the newly entered auxiliary tree set of data and repeats for each new auxiliary tree encountered creating a snowball effect. Note that in this analogy, the snowball casts off some nodes and collects other nodes in its trajectory towards the final auxiliary tree that contains the searched element. Before performing a cut, the new auxiliary tree is queried to obtain a cut depth which is used in processing the previous auxiliary tree. Starting from a Reference tree in Fig. 7 obtained after performing searches on 17, 7, 1, 18, 23, 31, 27, 13, 9, 20, we see the corresponding Tango Tree as in Fig. 8. On this tree during a walk towards element  $x_i = 23$  we would have traversed the top auxiliary tree 20 and entered auxiliary tree 22. Note that preferred paths are marked in red and auxiliary tree roots in magenta. Remember that all auxiliary trees are actually RB trees so the red shadows on some nodes represent the red nodes in the RB trees.



In case we would have searched for element 7 during the Tango Walk we would have crossed the top auxiliary tree rooted at 20, the auxiliary tree rooted at 10, auxiliary tree rooted at 2 and auxiliary tree rooted at 6. That means that we would have to perform the Tango Cut-And-Join several times. For each entrance to a new auxiliary tree, the tree that was just crossed is processed by a Tango cut algorithm. In order to perform the Tango cut we need to determine the cut range. A specific  $d$  value is the input data in the algorithm to determine the cut range.

## Determining d

This specific  $d$  value is determined every time when we cross the boundary to a new auxiliary tree and it is determined from the soon-to-be-traversed auxiliary tree. We know when we cross the boundary by looking at the `isRoot` attribute of each node en route to  $x_i$ . We ignore the `isRoot` attribute of the root of Tango tree. To simplify the code we don't even set it on Tango root and that is why top nodes are not colored in magenta in any of the figures.

The value of  $d$  is determined by subtracting 1 from the minimum between the root of the new auxiliary tree  $\text{minD}$  value and its current  $d$  value.

So in search for 23 we reach auxiliary tree rooted at 22 we calculate minimum between its  $\text{minD}$  value and its  $d$  value and we subtract 1 and we get this special value of  $d=2$ . Please observe by looking in the reference tree in fig. 7 that, that is exactly the depth value where the new search will change the value of a preferred child and create a new interleave. See node 20 on the reference tree in fig 7.

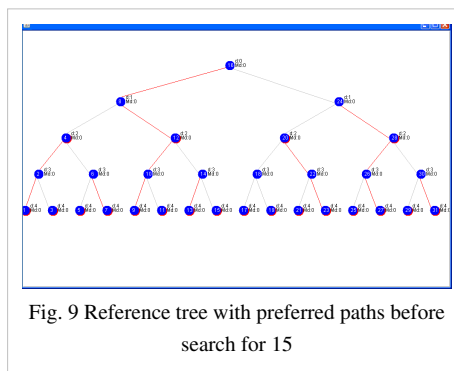
There are some particular cases when either  $\text{minD}$  or  $d$  are out of range or unknown. In the implementation that is represented by a -1, a value which denotes + or -infinity if the value denotes the right side of the range or left side of the range. In all the figures the nodes where the value of  $\text{minD}$  or  $\text{maxD}$  is -1 do not show the corresponding value(s) for brevity reasons.

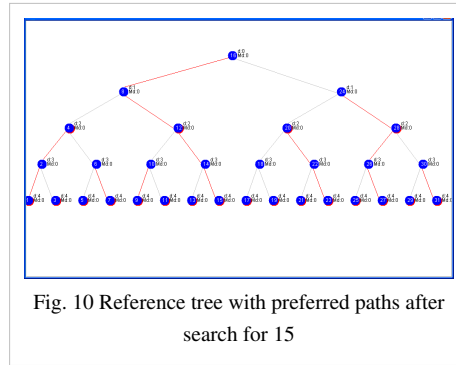
The value of -1 is screened out during the determination of a minimum so it does not mask legitimate  $d$  values. This is not the case for  $\text{maxD}$ .

## Determining the cutting range

Considering the auxiliary tree A in Fig 9, we need to determine nodes  $l$  and  $r$ . According to [Demaine et al. 2004] The node  $l$  is the node with the minimum key value (1) and the depth greater than  $d$  (2) and can be found in  $O(\log k)$  time where  $k$  is the number of nodes in A.

We observe by looking at the virtual reference tree that all nodes to be cut are actually under the interleave created by the search at the node corresponding to the auxiliary tree the tango search is about to enter. In Fig. 9 we show the reference tree at state  $t_i$  after a search for 9 and in Fig. 10 we show the reference tree at state  $t_{i+1}$  after a search for 15. The interleave appears at node 12 and we want to 'cut' all the nodes in the original path that are below 12. We observe their keys are all in a range of values and their  $d$  values are higher than the cutting  $d$  which is the depth of the interleave node ( $d=2$ ).





Of course we need to find the nodes to be cut not in the reference tree but in the corresponding auxiliary tree. We use the value of depth obtained during the Tango search as input to calculate the cut range.

We need to determine the minimum key of a node  $l$  that has the depth greater than  $d$ . This node can be found in  $O(\log k)$  time and is the left boundary of our interval. In the same way, a node  $r$  with the maximum key value and the depth greater than  $d$  can also be found in  $O(\log k)$ , where  $k$  is the number of nodes in  $A$ .

We already observed that the keys of the nodes to cut are in a key range so instead to take all the nodes in the corresponding auxiliary tree and check if their  $d$  is greater than the input  $d$  we can just find the left side of the range and the right side of the range.

### Finding $l$

We first find  $l$  which is the leftmost element to be cut (using `getL`). We determine it by walking to the leftmost child whose sub-tree has either its  $d$  value or maximum depth greater than  $d$ .

No nodes meeting this criteria result in  $l$  being  $-\infty$  (or `NIL` in the implementation). This means that the cut interval extends all the way to the left so during the cut or the join less split and concatenate operations have to be performed.

### Finding $r$

For finding  $r$  we walk right to the rightmost node whose  $d$  value is greater than  $d$ . No nodes meeting this criteria result in  $r$  being  $+\infty$  (or `NIL` in the implementation). This means that the cut interval extends all the way to the left so during the cut or the join less split and concatenate operations have to be performed.

Algorithms 7 and 8 describe the implementation of `getL` and `getR`.

### Finding $l'$ and $r'$

Following the determination of  $l$  and  $r$ , the predecessor  $l'$  of the node  $l$  and the successor  $r'$  of  $r$  are found. These new nodes are the first nodes that 'survive' the cut. A `NIL` in  $l$  will also result in a `NIL`  $l'$  and a `NIL` in  $r$  will result in a `NIL` in  $r'$ .

Both nodes being `NIL` take a new meaning signifying that all the nodes will survive the cut so practically we can skip the Tango cut and takes the set of nodes directly to Tango join.

During the Tango Search when we finally reached the searched node the tango cut algorithm is run once more however its input provided by Tango search is changed. The input value of  $d$  is now the  $d$  of the searched node and the tree to be cut is the auxiliary tree that contains the searched node.

During the Tango Search after encountering any new auxiliary tree (NAT) the Tango cut is normally applied on the tree above (except for final cut as described above) and results in a new structure of auxiliary trees. Let's call this result of the cut  $A$ .

The Join operation will join a set of nodes with the nodes in the NAT. There is an exception that applies when we reached the searched node, in which case we join A with the auxiliary tree rooted at the preceding marked node of the searched node.

The Tango cut algorithm consists of a sequence of tree split, mark and concatenate algorithms. The Tango join algorithm consists of a different sequence of tree split, un-mark and concatenate algorithms. Marking a node is just setting the isRoot attribute to true and un-marking setting the same attribute to false.

## RB Split and RB Concatenate

Red-black trees support search, split and concatenate in  $O(\log n)$  time. The split and concatenate operations are the following:

**Split:** A red-black tree is split by rearranging the tree so that  $x$  is at the root, the left sub-tree of  $x$  is a red-black tree on the nodes with keys less than  $x$ , and the right sub-tree of  $x$  is a red-black tree on the nodes with keys greater than  $x$ .

**Concatenate:** A red-black tree is concatenated by rearranging  $x$ 's sub-tree to form a red-black tree on  $x$  and the nodes in its sub-tree. As condition for this operation all the nodes in one of the trees have to have lower values than the nodes in the other tree.

Note that the behavior for split and concatenate in [Demaine et al. 2004] differs slightly from the standard functionality of these operations as the signature of the operations differ in terms of number and type of input and output parameters.

The two operations describe above apply only to an auxiliary tree and do not cross into other auxiliary trees. We use the isRoot information for each node to avoid wandering in other trees.

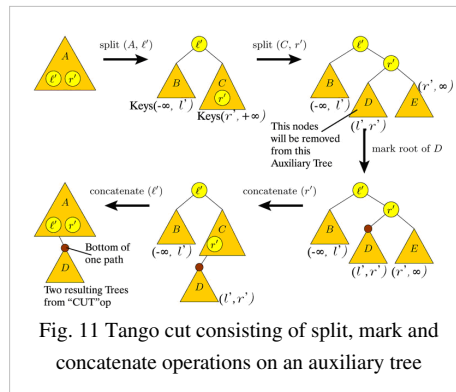
## Tango Cut algorithm

The main purpose of the cut is to separate a tree in two sets of nodes. Nodes need to be pushed to the top because the search path in the virtual reference tree traversed them and nodes that become less important (cut nodes) and are pushed downwards to the next auxiliary tree.

We already determined the values of  $l'$  and  $r'$  and we want to cut the input auxiliary tree (A in Fig. 11) in the range  $l'$  to  $r'$ . We need to split A in three trees and then cut the middle tree and re-assemble the remaining parts. We start with a split at  $l'$  that creates tree B and tree C. Tree B will be free of nodes to cut but tree C will contain nodes to cut and nodes to keep. We do a second split at  $r'$  and obtain trees D and E. D contains all nodes to be cut and E contains all the nodes to keep. We then mark the root of D as isRoot therefore logically pushing it to a lower level. We then concatenate at  $r'$ . Note: this is not really a standard RB concatenate but rather a merge between a node and a tree. As a result we obtained C. The last operation is to concatenate C tree, B tree and node  $l'$  obtaining the nodes we want to keep in the new tree A. Note: this is not really a standard RB concatenate but rather a merge between a node and a tree and then a standard two tree concatenation.

The resulting new node A is actually composed of two auxiliary trees: the top one that contains nodes we want to favor and the lower 'hung' D which contains nodes that get pushed downwards via this operation. The nodes being pushed downwards are exactly the nodes that were on the old preferred path corresponding to the auxiliary tree being processed but not in the new preferred path. The nodes being pushed upwards (now in A) are exactly the nodes that became part of the new preferred path due to the performed search. Fig. 11 shows this flow of operations.





The following special cases may occur:

- $l'$  and  $r'$  are NIL so cut is performed.
- $r'$  is NIL we just do a split at  $l'$  and C becomes the result while we hang B under C via a mark operation.
- $l'$  is NIL we just do a split at  $r'$  and E becomes the result and we hang the left resulting tree under E.

The result will be then joined with the content of the next auxiliary tree via the Tango Join algorithm.

### Tango Join algorithm

During the Tango Search after encountering any new auxiliary tree (NAT) the Tango Cut being applied on the tree above the NAT results in a new structure of auxiliary trees. We can think of that as a Tango sub tree. It will normally contain at least two connected auxiliary trees. The top tree A containing the nodes we want to keep close to the surface of the Tango Tree (so we can achieve the 'bring to top' approach) and 'hang' to it is auxiliary tree D which was pushed downwards. In case of search for 23 all of the nodes from previous auxiliary tree should be kept close to surface and the set of nodes destined to be moved in D is empty so we have no D. Regardless of the situation the result of the Tango cut will contain at least one node (the Tango root) in Tree A. Let's call this result of the cut A.

The Join will join A set of nodes with the nodes in the NAT. That is done via two splits, an un-mark and two concatenates.

Fig. 12 shows the high level sequence where A is coming from the previous cut tree and B is the NAT. We observe that NAT is actually hung under the tree that was just cut therefore the values of its keys are all in a range of two adjacent keys (a key and its successor) in the tree that was just cut. That is normal for any BST. If the NAT is hanging as a left tree the parent node marks the right side of the range while its predecessor (in the tree that was just cut universe) marks the left side of the range. So in order to join the two trees we just have to wedge B under to the left of its parent in A. Similarly for the case where B happens to hang to the right of its parent where we wedge the content of B to the right of its parent. In order to find the wedge insertion point we can just search in the A for the root of NAT. Of course that value is not in NAT but it will find a close value and by taking its predecessor or successor (depending on the search algorithm and if the close value was before or after the value) we find the two nodes between where B should be wedged. Let's call these values  $lPrime$  and  $rPrime$ . Next is to split A first at  $lPrime$  and then at  $rPrime$  therefore creating three trees, a First part (FP), middle part (MP) and last part (LP). While these operations were done in the A universe they also need to carry all the other auxiliary trees as in the Tango universe. In the Tango (forest) universe we discover that MD is actually is B however is severed logically from  $rPrime$  because its root is marked as  $isRoot$  and it appears like a hung auxiliary tree. Since we want that wedged, we "un-mark" it by resetting its  $isRoot$  attribute and making it logically part of  $rPrime$ . Now we have the final structure in place but we still need to concatenate it first at  $rPrime$  and then at  $lPrime$  to absorb all the nodes under the same Joined resulting tree.

The difficulty in doing this is the fact that the standard concatenates do not take nodes but just trees. A two tree and a node operation could be constructed and then repeated to obtain a Tango concatenate; however, it is hard to preserve

the RB integrity and is dependant on the order of operations so the resulting structure is different even if it contains the same node. That is an issue because we can not control the exact reproduction of the ideal structure as if generated from the reference tree. It can be done in such a way to contain all nodes and preserve the correct RB tree structure however the geometry is dictated by the RB concatenates and is not necessarily the ideal geometry mirroring perfectly the reference tree.

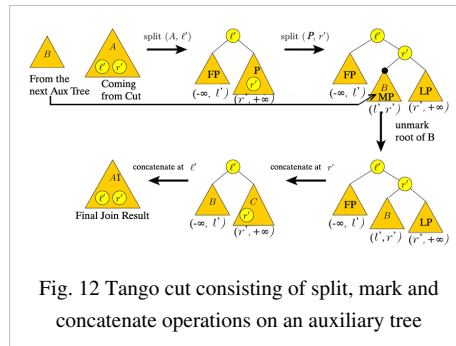


Fig. 12 Tango cut consisting of split, mark and concatenate operations on an auxiliary tree

So for example let's say search for 23. We obtain A as the result on the first Tango cut on the top auxiliary tree. See Fig. 13 where 22 is the root of NAT.

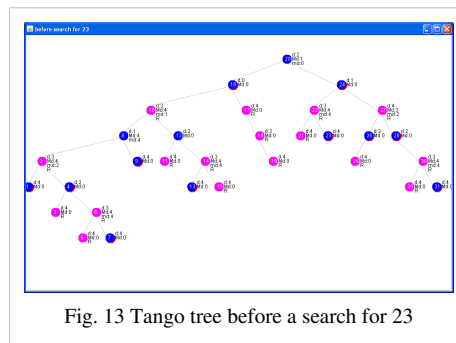


Fig. 13 Tango tree before a search for 23

We use the value of NAT (22) to search in the tree above and we obtain 20 and 24 as the lPrime and rPrime nodes. We split at lPrime (20) and we obtain FP as in Fig. 14.

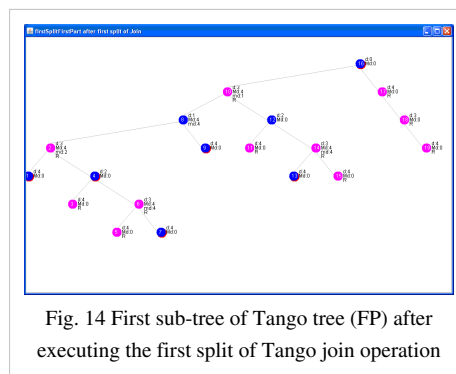


Fig. 14 First sub-tree of Tango tree (FP) after executing the first split of Tango join operation

We then split for the second time at rPrime (24) to get the last tree LP as in Fig. 15.

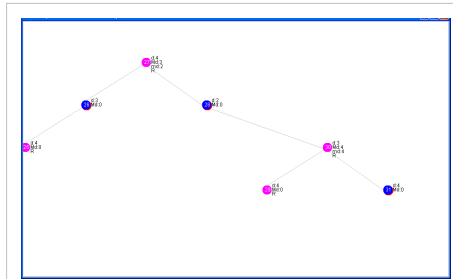


Fig. 15 Last sub-tree of Tango tree (LP) after executing the second split of the Tango join operation

Next we unmark B which is rooted at 22 and we obtain the result in Fig 16. As you can see 22 now is part of the top of the structure. That makes sense if you look at Fig. 10 representing the ‘virtual’ reference tree. To reach 23 which is our target we would have had to go through 22.

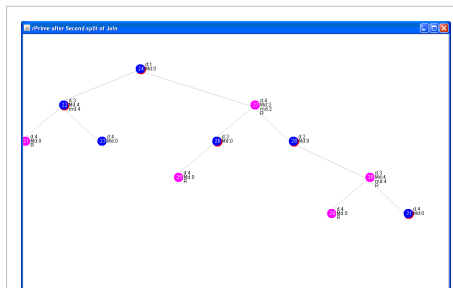


Fig. 16 Sub-tree rooted at rPrime after executing the second split of the Tango join operation

We then concatenate with rPrime and obtain the result in presented in Fig. 17:

Second concatenation takes place and in this particular example will not result in rearranging of the nodes so Fig. 17 is the final result of the Join operation.

## Construction Algorithms And Pseudo-code

Construct the reference tree and perform warm-up searches.

**Function:** constructReferenceTree

**Input:** None

**Output:** ReferenceTree p

```
ReferenceTree p = new ReferenceTree()
insertDataInReferenceTree()
p.setAllDepths()
p.warmUpSearches()
ArrayList<PreferredPath> paths = p.collectAndSortPreferredPaths()
assert paths.size() > 0
PreferredPath top = p.collectTopNodePath(p.root)
TangoTree tangoTree = new TangoTree(top)
tangoTree.updateMinMaxD()
while (takeNext PreferredPath path in paths) do
    if (path.top == p.root) then
        continue; // skip the top path as it was already added
    else
```

```

        RBTree auxTree = new RBTree(path)
        auxTree.updateMinMaxD()
        auxTree.root.isRoot = true
        tangoTree.hang(auxTree)
    return p

```

Construct an Auxiliary tree out of a Preferred Path. Used in the construction phase.

**Function:** constructAuxiliaryTree

**Input:** Preferred Path path

**Output:** AuxiliaryTree **this**

```

RBTree(PreferredPath path)
    this()
    RefTreeNode refNode = path.top
    while (Next PreferredPath path in paths exists) do
        RBNode n = new RBNode(refNode.value, RedBlackColor.RED,
null, null)
        n.d = refNode.d
        this.insert(n)
        refNode = refNode.getPreferredChild()

```

Construct a Tango tree from the Reference tree.

**Function:** constructTangoTree

**Input:** ReferenceTree p

**Output:** TangoTree tangoTree

```

ArrayList<PreferredPath> paths = p.collectAndSortPreferredPaths()
assert paths.size() > 0
PreferredPath top = p.collectTopNodePath(p.root)
TangoTree tangoTree = new TangoTree(top)
tangoTree.updateMinMaxD()
while (Next PreferredPath path in paths exists) do
    if (path.top = p.root) then
        continue; // skip the top path as it was already added
    else
        RBTree auxTree = new RBTree(path)
        auxTree.updateMinMaxD()
        auxTree.root.isRoot = true
return tangoTree

```

Set the depths of all noded in the tree. Used only once **for** setting the depths of all nodes in the reference tree.

**Function:** setAllDepths

**Input:** None

setAllDepthsHelper(root)

Algorithm 14. Set the depths of all noded in the tree.

**Function:** setAllDepthsHelper

**Input:** RefTreeNode n

```

if (n == NIL) then
    return
n.setD(getDepth(n))
if (n.right != null)
    setAllDepthsHelper((RefTreeNode) n.right)
if (n.left != null) then
    setAllDepthsHelper((RefTreeNode) (n.left))

```

## Operation Algorithms And Pseudo-code

There is just one operation: search that calls a number of algorithms to rearrange the data structure.

Tango search pseudocode. Used by the main operation on a Tango tree.

**Function:** tangoSearch

**Input:** int vKey

Node currentRoot = root

Node currentNode = root

Boolean found = **false**

**while** (**true**) **do**

**if** (currentNode == NIL) **then**

        found = **false**

**break**

**if** (currentNode.isRoot && currentNode != root) **then**

        cutAndJoin(minIgnoreMinusOne(currentNode.minD,

currentNode.d) - 1, currentRoot, currentNode)

        currentRoot = currentNode

**if** (currentNode.value == vKey) **then**

        found = **true**

**if** (currentNode != currentRoot) **then**

            cutAndJoin(currentNode.d, currentRoot, currentRoot)

**break**

**if** (currentNode.value < vKey) **then**

        currentNode = (RNode) currentNode.right

**else**

        currentNode = (RNode) currentNode.left

**if** (found) **then**

**return** currentNode

**else**

**return** NIL

Tango Cut and Join used by Tango Search

**Function:** cutAntJoin

**Input:** int d, Node currentRoot, Node newlyFoundRoot

RNode n = currentRoot

RNode l = NIL // l is the last node that gets cut

RNode lPrime = NIL // l' is the first node that escapes from cutting

RNode r = NIL

RNode rPrime = NIL

**if** (currentRoot.maxD <= d) // no nodes to be cut besides maybe the root

```

    if (currentRoot.d > d)
        l = currentRoot
        r = currentRoot
        lPrime = getPredecessor(l, currentRoot)
        rPrime = getSuccessor(r, currentRoot)

    else // there are nodes to be cut underneath
        l = getL(d, n, currentRoot)
        // determine lPrime as predecessor of l or NIL if it is the last
        if (l != NIL)
            lPrime = getPredecessor(l, currentRoot)
        else
            lPrime = NIL // - infinity maybe redundant

        // end calculating l and l prime
        // find r the right side node of the cutting range based on value
        n = currentRoot
        r = getR(d, n, currentRoot)
        if (r != NIL) // the root is not to be cut
            rPrime = getSuccessor(r, currentRoot)

    checkLandR(d, l, r, currentRoot)
    RBTREE aTree = NIL
    if (lPrime == NIL && rPrime == NIL) // nothing to cut therefore so
    aTree is the whole
        aTree = new RBTREE()
        aTree.root = currentRoot
    else
        RBTreesPair aAndDtreePair = new RBTreesPair()
        aTree = tangoCut(lPrime, rPrime, currentRoot)
    RBTREE afterCutAndJoin = tangoJoin(aTree,
    newlyFoundRootOrCurrentIfWeFound)

```

Tango Cut used by Tango cut And Join to separate nodes that need to be pushed to top from the rest of nodes.

**Function:** tangoCut

**Input:** RBNODE lPrime, RBNODE rPrime, RBNODE aRoot

**Output:** RBTREE

saveShadowAuxTrees(aRoot)

```

if (lPrime == null || rPrime == null) { // just one splitAnd Concatenate
    return simplifiedTangoCut(lPrime, rPrime, aRoot)
}

```

RBTREE a = new RBTREE()

a.root = aRoot

RBTreesPair firstPartAndLastPart = new RBTreesPair()

split(lPrime, a, firstPartAndLastPart)

RBTREE b = firstPartAndLastPart.firstPart

```

RBTree c = firstPartAndLastPart.lastPart
firstPartAndLastPart.firstPart.verifyProperties()
firstPartAndLastPart.lastPart.verifyProperties()
firstPartAndLastPart = new RBTreesPair()
split(rPrime, c, firstPartAndLastPart) // problem
firstPartAndLastPart.firstPart.verifyProperties()
firstPartAndLastPart.lastPart.verifyProperties()
RBTree d = firstPartAndLastPart.firstPart
RBTree e = firstPartAndLastPart.lastPart
// disconnect d
rPrime.left = NIL
d.root.parent = NIL

```

Tango Join used by Tango Cut and Join to join the result of Tango cut to auxiliary trees.

**Function:** tangoJoin

**Input:** RBTree a, RBNode newlyFoundRoot

**Output:** RBTree finalJoinResult

```
RBTree bPrevOp = new RBTree()
```

```
RBTree d = new RBTree()
```

```
order(bPrevOp, d, a, newlyFoundRoot)
```

```
RBNodePair lAndR = bPrevOp.searchBoundedLPrimeAndRPrime(d.root.value)
```

```
if (lPrime == null || rPrime == null) // just one split and one
concatenate
```

```
    return simplifiedTangoJoin(lPrime, rPrime, bPrevOp, d.root)
```

```
RBNode lPrime = lAndR.lPrime
```

```
RBNode rPrime = lAndR.rPrime
```

```
RBTreesPair firstPartAndLastPart = new RBTreesPair()
```

```
split(lPrime, bPrevOp, firstPartAndLastPart)
```

```
RBTree b = firstPartAndLastPart.firstPart
```

```
RBTree c = firstPartAndLastPart.lastPart
```

```
firstPartAndLastPart.firstPart.verifyProperties()
```

```
firstPartAndLastPart.lastPart.verifyProperties()
```

```
firstPartAndLastPart = new RBTreesPair()
```

```
split(rPrime, c, firstPartAndLastPart) //
```

```
firstPartAndLastPart.firstPart.verifyProperties()
```

```
firstPartAndLastPart.lastPart.verifyProperties()
```

```
RBTree e = firstPartAndLastPart.lastPart
```

```
// reconnect d which is normally newlyFoundRoot
```

```
d.root.isRoot = false // un-mark, a difference from tangoCut
```

```
// concatenate part
```

```
rPrime.parent = NIL // avoid side effects
```

```
RBTree res1 = concatenate(rPrime, d, e)
```

```
lPrime.parent = NIL // avoid side effects
```

```
RBTree res2 = concatenate(lPrime, b, res1)
```

```
return res2
```

Check **if** a node *n* is in an auxiliary tree defined by *currentRoot*. Used to verify wandering.

**Function:** *isInThisTree*

**Input:** *RBNode n*, *RBNode currentRoot*

**Output:** Boolean *v*

```
if (n.isRoot and n != currentRoot) then
    return false
else
    return true
```

Find node *l* as left of range used by Tango Cut, different from the [Demaine et al. 2004] paper.

**Function:** *getL*

**Input:** *int d*, *Node n*, *Node currentRoot*

**Output:** *Node l*

*Node l* = *n*

```
if (left[n] != NIL) and (not(isRoot(n) or n == currentRoot) and
((left(n).maxD > d) or (left(n).d > d)) then
    l=getL(d, left(n), currentRoot)
else
    if (n.d > d)
        l = n
    else
        l=getL(d, right(n), currentRoot)
return l
```

Find node *r* as the right limit of the range used by Tango Cut.

**Function:** *getR*

**Input:** *int d*, *Node n*, *Node currentRoot*

**Output:** *Node r*

*Node r* = *n*

```
if (right[n] != NIL) and (not(isRoot(n) or n == currentRoot) and
((right(n).maxD > d) or (right(n).d > d)) then
    r=getR(d, right(n), currentRoot)
else
    if (n.d > d)
        r = n
    else
        r=getR(d, right(n), currentRoot)
return r
```

Return Sibling. Within enclosing auxiliary tree boundary

**Function:** *siblingBound*

**Input:** *RBNode n*, *RBNode boundingRoot*

**Output:** *Node p*

```
if (n == left[parent[n]] && isInThisTreeAcceptsNull(left[parent[n]]),
boundingRoot)) then
```



```

    if (isInThisTreeAcceptsNull(right[parent[n]], boundingRoot) then
        return right[parent[n]]
    else
        return NIL
else
    if (isInThisTreeAcceptsNull(left[parent[n]], boundingRoot)) then
        return left[parent[n]]
    else
        return NIL

```

Return Uncle. Within enclosing auxiliary tree boundary

Function: uncleBound

Input: RBNODE n, RBNODE boundingRoot

Output: Node p

```

if (isInThisTreeAcceptsNull(parent[n], boundingRoot)) then
    return siblingBound(boundingRoot)
else
    return NIL

```

Update Min Max D values in red black tree augmented node. Used to update Tango Tree node attributes.

Function: updateMinMaxD

Input: RBNODE n

int minCandidate

```

if (n.left != NIL) then
    updateMinMaxD(n.left)
if (n.right != NIL) {
    updateMinMaxD(n.right)
if (n.left != NIL) then
    int maxFromLeft = max(n.left.d, n.left.maxD)
    n.maxD = maxFromLeft > n.maxD ? maxFromLeft : n.maxD
    if (n.left.minD != -1) {
        minCandidate = min(n.left.d, n.left.minD)
    else
        minCandidate = n.left.d
    if (n.minD != -1) then
        n.minD = min(n.minD, minCandidate)
    else
        n.minD = minCandidate
if (n.right != NIL) then
    int maxFromRight = max(n.right.d, n.right.maxD)
    n.maxD = maxFromRight > n.maxD ? maxFromRight : n.maxD
    if (n.right.minD != -1) then
        minCandidate = min(n.right.d, n.right.minD)
    else
        minCandidate = n.right.d
    if (n.minD != -1) then

```

```

        n.minD = min(n.minD, minCandidate)
    else
        n.minD = minCandidate

```

Search Bounded lPrime and rPrime used by Tango Join.

**Function:** searchBoundedLPrimeAndRPrime

**Input:** RBTREE rbTree, int value

**Output:** NodePair p

RBNODE n = root

RBNODE prevPrevN = NIL

RBNODE prevN = NIL

RBNODEPAIR lAndR

```

while (n != NIL && isInThisTree(n, rbTree.root)) do
    int compResult = value.compareTo(n.value)
    if (key(n) == value) then
        lAndR = new RBNODEPAIR(n, n)
        return lAndR
    else
        if (key(n) < value) then
            prevPrevN = prevN
            prevN = n
            if (isInThisTree(n.left, rbTree.root)) then
                n = n.left
            else
                n = NIL
        else
            prevPrevN = prevN
            prevN = n
            if (isInThisTree(n.right, rbTree.root)) then
                n = n.right
            else
                n = NIL
lAndR = new RBNODEPAIR(prevPrevN, prevN)
return lAndR

```

The minimum in a binary search tree is always located **if** the left side path is traversed down to the leaf in  $O(\log n)$  time:

Minimum Value Tree pseudocode. Used by successor.

**Function:** min\_val\_tree

**Input:** Node x

**Output:** Node x

```

while left(x) != NIL do
    x = left(x)
return x

```

Maximum Value Tree pseudocode used by predecessor.

**Function:** max\_val\_tree

```

Input: Node x
Output: Node x
while right(x) != NIL do
    x = right(x)
return x

```

The next two algorithms describe how to compute the predecessor and successor of a node. The predecessor of a node  $x$  is a node with the greatest value smaller the key[ $x$ ].

Predecessor computing pseudocode used to find lPrime.

Function: predecessor

Input: RBNODE  $x$

Output: RBNODE  $y$

RBNODE  $y = \text{null}$

```

if (n.left != null && isInThisTree((RBNODE) (n.left)), root))
    return getMaximum((RBNODE) (n.left), root)

```

$y = (\text{RBNODE}) (n.\text{parent})$

```

if (y == currentRoot) // don't let it escape above its own root
    return NIL//

```

----->

```

while (y != NIL && (y != currentRoot) && n == ((RBNODE) (y.left))) do
    n = y
    if (isInThisTree((RBNODE) (y.parent)), root))
        y = (RBNODE) (y.parent)
    else
        y = null
return (RBNODE) y

```

Successor computing algorithm used to find rPrime.

Function: successor

Input: Node  $x$

Output: Node  $y$

RBNODE  $y = \text{NIL}$

```

if (n.right != NIL && isInThisTree((RBNODE) (n.right)), currentRoot))

    return getMinimum((RBNODE) (n.right), currentRoot)

```

$y = (\text{RBNODE}) (n.\text{parent})$

```

if (y == currentRoot) // don't let it escape above its own root
    return NIL//

```

----->

```

while (y != NIL && isInThisTree(y, currentRoot) && n == ((RBNODE)
(y.right))) do
    n = y
    y = (RBNODE) (y.parent)

```

```
return (RBNode) y
```

Traverse Tree pseudocode used by Tango Search.

Function: traverse\_tree

Input: Node x

Output: None

```
if x != NIL and InTheAuxiliaryTree(x) then
    traverse_tree (left(x))
    traverse_tree (right(x))
```

Search Tree algorithm used to find lPrime and rPrime for Tango Join.

Function: search\_tree

Input: Node x, value

Output: Node x

```
if x = NIL or k = key[x] then
    return x
if k < key[x] then
    return search_tree(left[x]; value)
else
    return search_tree(right[x]; value)
```

Find minimum by ignoring specific values used for the calculation of d.

Function: minIgnoreMinusOne

Input: int minD, int d

Output: int d

```
if (minD == -1) then
    return d
if (d == -1) then
    return minD
return min(minD, d)
```

RedBlack split and Red Black concatenate algorithms are described in the RON WEIN, 2005, Efficient Implementation of Red-Black Trees with Split and Catenate Operations, [http://www.cs.tau.ac.il/~wein/publications/pdfs/rb\\_tree.pdf](http://www.cs.tau.ac.il/~wein/publications/pdfs/rb_tree.pdf)

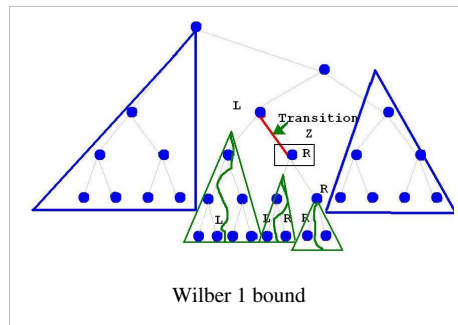
## Analysis

Here are some elements necessary to understand why the Tango Tree achieve such an amazing performance and become  $O(\log \log n)$  competitive.

Wilber's 1st Lower Bound [Wil89]

Fix an arbitrary static lower bound tree  $P$  with no relation to the actual BST  $T$ , but over the same keys. In the application that we consider,  $P$  is a perfect binary tree. For each node  $y$  in  $P$ , we label each access  $X1$   $L$  if key  $X1$  is in  $y$ 's left subtree in  $P$ ,  $R$  if key  $X1$  is in  $y$ 's right subtree in  $P$ , or leave it blank if otherwise. For each  $y$ , we count the number of interleaves (the number of alterations) between accesses to the left and right subtrees:  $\text{interleave}(y) = ?$  of alternations  $L ? R$ .

Wilber's 1st Lower Bound [Wil89] states that the total number of interleaves is a lower bound for all BST data structures serving the access sequence  $x$ . The lower bound tree  $P$  must remain static. Proof.



We define the transition point of  $y$  in  $P$  to be the highest node  $z$  in the BST  $T$  such that the root-to- $z$  path in  $T$  includes a node from the left and right subtrees if  $y$  in  $P$ . Observe that the transition point is well defined, and does not change until we touch  $z$ . In addition, the transition point is unique for every node  $y$ .

Lemma 1

The running time of an access  $x_i$  is  $O((k + 1)(1 + \log \log n))$ , where  $k$  is the number of nodes whose preferred child changes during access  $x_i$ .

Lemma 2

The number of nodes whose preferred child changes from left to right or from right to left during an access  $x_i$  is equal to the interleave bound ( $IB_i(X)$ ) of access  $x_i$ .

Theorem 1

The running time of the Tango BST on an sequence  $X$  of  $m$  accesses over the universe  $1, 2, \dots, n$  is  $O((OPT(X) + n)(1 + \log \log n))$  where  $OPT(X)$  is the cost of the offline optimal BST servicing  $X$ .

Corollary 1.1

When  $m = (n)$ , the running time of the Tango BST is ;  $O(OPT(X)(1 + \log \log n))$

## Bibliography

- Erik D. Demaine, Dion Harmon, John Iacono and Mihai Patrascu, Dynamic optimality - almost [competitive online binary search tree], In Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science, pages 484-490, Rome, Italy, October 2004, <http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=9430>
- Allen, Brian; and Munro, Ian (1978), "Self-organizing search trees", Journal of the ACM 25 (4): 526–535, doi:10.1145/322092.322094
- Knuth, Donald. *The Art of Computer Programming*, Volume 3: *Sorting and Searching*, Third Edition. Addison-Wesley, 1997. ISBN 0-201-89685-0. Page 478 of section 6.2.3.
- DANIEL DOMINIC SLEATOR AND ROBERT ENDRE TARJAN, 1985, Self-adjusting binary search trees. Journal of the ACM, 32(3):652-686, July 1985

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.95.1380&rep=rep1&type=pdf>

- ERIK D. DEMAINE, DION HARMON, JOHN IACONO, AND MIHAI PATRASCU, 2004, Dynamic optimality-almost, FOCS 2004
- ERIK D. DEMAINE, DION HARMON, JOHN IACONO, AND MIHAI PATRASCU, Dynamic optimality-almost. SIAM J. Computers., 37(1):240-251, 2007

R. BAYER 1972, Symmetric Binary B-trees: Data Structure and Maintenance Algorithms, Acta Informatica 1:290-306

<http://www.springerlink.com/content/qh51m2014673513j/>

- L. J. GUIBAS AND R. SEDGEWICK, 1978, A dichromatic framework for balanced trees. Nineteenth Annual IEEE Symposium on Foundations of Computer Science, pages 8–12, 1978

<http://www.cs.princeton.edu/~sssix/papers/rb-trees.pdf>

- CHENGWEN CHRIS WANG, JONATHAN DERRYBERRY, DANIEL DOMINIC SLEATOR 2006,  $O(\log \log n)$ -Competitive Dynamic Binary Search Trees, SODA '06, January 22–26, Miami, FL 2006 SIAM ISBN 0-89871-605-5/06/01

<http://books.google.ca/books?id=R3WyVR4nqzGC&pg=PA374&lpg=PA374&dq=O%28log+log+n%29-Competitive+Dynamic+Binary+Search#v=onepage&q=O%28log%20log%20n%29-Competitive%20Dynamic%20Binary%20Search&f=false>

- ROBERT WILBER, 1989, Lower bounds for accessing binary search trees with rotations, SIAM Journal on Computing, 18(1):56-67, 1989

<http://www.informatik.uni-trier.de/~ley/db/journals/siamcomp/siamcomp18.html>

- ROBERT ENDRE TARJAN, 1983, Linking and cutting trees, In Data Structures and Network Algorithms, chapter 5, pages 59–70. Society for Industrial and Applied Mathematics, 1983

<http://www.cambridge.org/aus/catalogue/catalogue.asp?isbn=9780898711875>

- DANIEL DOMINIC SLEATOR, Open Source top-down splay tree implementation, <http://www.link.cs.cmu.edu/splay/>

## External links

- ERIK D. DEMAINE, DION HARMON, JOHN IACONO, AND MIHAI PATRASCU, 2007, Dynamic optimality-almost. <sup>[1]</sup>
- RON WEIN, 2005, Efficient Implementation of Red-Black Trees with Split and Catenate Operations <sup>[2]</sup>
- DANIEL DOMINIC SLEATOR, Open Source top-down splay tree implementation <sup>[3]</sup>

## References

- [1] [http://epubs.siam.org/sicomp/resource/1/smjcat/v37/i1/p240\\_s1?isAuthorized=no](http://epubs.siam.org/sicomp/resource/1/smjcat/v37/i1/p240_s1?isAuthorized=no)  
 [2] [http://www.cs.tau.ac.il/~wein/publications/pdfs/rb\\_tree.pdf?isAuthorized=no](http://www.cs.tau.ac.il/~wein/publications/pdfs/rb_tree.pdf?isAuthorized=no)  
 [3] <http://www.link.cs.cmu.edu/splay/?isAuthorized=no>

# Van Emde Boas tree

Van Emde Boas tree	
Type	Non-binary tree
Invented	1975
Invented by	Peter van Emde Boas
Asymptotic complexity in big O notation	
Space	$O(M)$
Search	$O(\log \log M)$
Insert	$O(\log \log M)$
Delete	$O(\log \log M)$

A **Van Emde Boas tree** (or **Van Emde Boas priority queue**), also known as a **vEB tree**, is a tree data structure which implements an associative array with  $m$ -bit integer keys. It performs all operations in  $O(\log m)$  time. Notice that  $m$  is the *size* of the keys — therefore  $O(\log m)$  is  $O(\log \log n)$  in a tree where every key below  $n$  is set, exponentially better than a full self-balancing binary search tree. They also have good space efficiency when they contain a large number of elements, as discussed below. They were invented by a team led by Peter van Emde Boas in 1975.<sup>[1]</sup>

## Supported operations

The operations supported by a vEB tree are those of an *ordered associative array*, which includes the usual associative array operations along with two more *order* operations, *FindNext* and *FindPrevious*:<sup>[2]</sup>

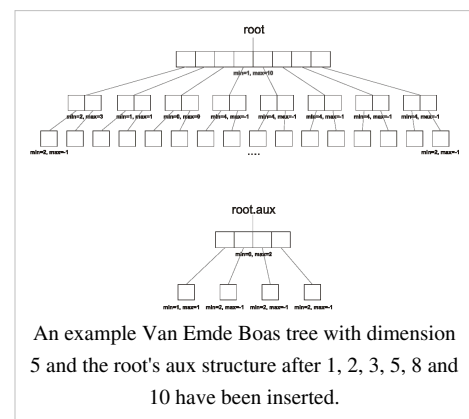
- *Insert*: insert a key/value pair with an  $m$ -bit key
- *Delete*: remove the key/value pair with a given key
- *Lookup*: find the value associated with a given key
- *FindNext*: find the key/value pair with the smallest key at least a given  $k$
- *FindPrevious*: find the key/value pair with the largest key at most a given  $k$

## How it works

For the sake of simplicity, let  $\log_2 m = k$  for some integer  $k$ . Define  $M=2^m$ . A vEB tree  $T$  over the universe  $\{0, \dots, M-1\}$  has a root node that stores an array  $T.children$  of length  $M^{1/2}$ .  $T.children[i]$  is a pointer to a vEB tree that is responsible for the values  $\{iM^{1/2}, \dots, (i+1)M^{1/2}-1\}$ . Additionally,  $T$  stores two values  $T.min$  and  $T.max$  as well as an auxiliary vEB tree  $T.aux$ .

Data is stored in a vEB tree as follows: The smallest value currently in the tree is stored in  $T.min$  and largest value is stored in  $T.max$ . These two values are not stored anywhere else in the vEB tree. If  $T$  is empty then we use the convention that  $T.max=-1$  and  $T.min=M$ . Any other value  $x$  is stored in the subtree  $T.children[i]$  where  $i = \lfloor x/M^{1/2} \rfloor$ .

The auxiliary tree  $T.aux$  keeps track of which children are non-empty, so  $T.aux$  contains the value  $j$  if and only if  $T.children[j]$  is non-empty.



## FindNext

The operation  $\text{FindNext}(T, x)$  that searches for the successor of an element  $x$  in a vEB tree proceeds as follows: If  $x \leq T.\text{min}$  then the search is complete, and the answer is  $T.\text{min}$ . If  $x > T.\text{max}$  then the next element does not exist, return  $M$ . Otherwise, let  $i = \lfloor x/M^{1/2} \rfloor$ . If  $x \leq T.\text{children}[i].\text{max}$  then the value being searched for is contained in  $T.\text{children}[i]$  so the search proceeds recursively in  $T.\text{children}[i]$ . Otherwise, We search for the value  $i$  in  $T.\text{aux}$ . This gives us the index  $j$  of the first subtree that contains an element larger than  $x$ . The algorithm then returns  $T.\text{children}[j].\text{min}$ . The element found on the children level needs to be composed with the high bits to form a complete next element.

```
function FindNext(T, x)
  if x ≤ T.min then
    return T.min
  if x > T.max then // no next element
    return M
  i = floor(x /  $\sqrt{M}$ )
  lo = x %  $\sqrt{M}$ 
  hi = x - lo
  if lo ≤ T.children[i].max then
    return hi + FindNext(T.children[i], lo)
  return hi + T.children[FindNext(T.aux, i+1)].min
end
```

Note that, in any case, the algorithm performs  $O(1)$  work and then possibly recurses on a subtree over a universe of size  $M^{1/2}$  (an  $m/2$  bit universe). This gives a recurrence for the running time of  $T(m) = T(m/2) + O(1)$ , which resolves to  $O(\log m) = O(\log \log M)$ .

## Insert

The call  $\text{insert}(T, x)$  that inserts a value  $x$  into a vEB tree  $T$  operates as follows:

If  $T$  is empty then we set  $T.\text{min} = T.\text{max} = x$  and we are done.

Otherwise, if  $x < T.\text{min}$  then we insert  $T.\text{min}$  into the subtree  $i$  responsible for  $T.\text{min}$  and then set  $T.\text{min} = x$ . If  $T.\text{children}[i]$  was previously empty, then we also insert  $i$  into  $T.\text{aux}$

Otherwise, if  $x > T.\text{max}$  then we insert  $T.\text{max}$  into the subtree  $i$  responsible for  $T.\text{max}$  and then set  $T.\text{max} = x$ . If  $T.\text{children}[i]$  was previously empty, then we also insert  $i$  into  $T.\text{aux}$

Otherwise,  $T.\text{min} < x < T.\text{max}$  so we insert  $x$  into the subtree  $i$  responsible for  $x$ . If  $T.\text{children}[i]$  was previously empty, then we also insert  $i$  into  $T.\text{aux}$ .

In code:

```
function Insert(T, x)
  if T.min > T.max then // T is empty
    T.min = T.max = x;
    return
  if T.min == T.max then
    if x < T.min then
      T.min = x
    if x > T.max then
      T.max = x
    return
  if x < T.min then
```



```

        swap(x, T.min)
    if x > T.max then
        swap(x, T.max)
    i = floor(x /  $\sqrt{M}$ )
    Insert(T.children[i], x %  $\sqrt{M}$ )
    if T.children[i].min == T.children[i].max then
        Insert(T.aux, i)
end

```

The key to the efficiency of this procedure is that inserting an element into an empty vEB tree takes  $O(1)$  time. So, even though the algorithm sometimes makes two recursive calls, this only occurs when the first recursive call was into an empty subtree. This gives the same running time recurrence of  $T(m)=T(m/2) + O(1)$  as before.

## Delete

Deletion from vEB trees is the trickiest of the operations. The call *Delete*(*T*, *x*) that deletes a value *x* from a vEB tree *T* operates as follows:

If  $T.min = T.max = x$  then *x* is the only element stored in the tree and we set  $T.min = M$  and  $T.max = -1$  to indicate that the tree is empty.

Otherwise, if  $x = T.min$  then we need to find the second-smallest value *y* in the vEB tree, delete it from its current location, and set  $T.min=y$ . The second-smallest value *y* is either  $T.max$  or  $T.children[T.aux.min].min$ , so it can be found in  $O(1)$  time. In the latter case we delete *y* from the subtree that contains it.

Similarly, if  $x = T.max$  then we need to find the second-largest value *y* in the vEB tree, delete it from its current location, and set  $T.max=y$ . The second-largest value *y* is either  $T.min$  or  $T.children[T.aux.max].max$ , so it can be found in  $O(1)$  time. In the latter case, we delete *y* from the subtree that contains it.

In case where *x* is not  $T.min$  or  $T.max$ , and *T* has no other elements, we know *x* is not in *T* and return without further operations.

Otherwise, we have the typical case where  $x \neq T.min$  and  $x \neq T.max$ . In this case we delete *x* from the subtree  $T.children[i]$  that contains *x*.

In any of the above cases, if we delete the last element *x* or *y* from any subtree  $T.children[i]$  then we also delete *i* from  $T.aux$ .

In code:

```

function Delete(T, x)
    if T.min == T.max == x then
        T.min = M
        T.max = -1
        return
    if x == T.min then
        if T.aux is empty then
            T.min = T.max
            return
        else
            x = T.children[T.aux.min].min
            T.min = x
    if x == T.max then
        if T.aux is empty then
            T.max = T.min

```

```

        return
    else
        x = T.children[T.aux.max].max
        T.max = x
    if T.aux is empty then
        return
    i = floor(x /  $\sqrt{M}$ )
    Delete(T.children[i], x %  $\sqrt{M}$ )
    if T.children[i] is empty then
        Delete(T.aux, i)
end

```

Again, the efficiency of this procedure hinges on the fact that deleting from a vEB tree that contains only one element takes only constant time. In particular, the last line of code only executes if  $x$  was the only element in  $T.children[i]$  prior to the deletion.

## Discussion

The assumption that  $\log m$  is an integer is unnecessary. The operations  $x / \sqrt{M}$  and  $x \% \sqrt{M}$  can be replaced by taking only higher-order  $\text{ceil}(m/2)$  and the lower-order  $\text{floor}(m/2)$  bits of  $x$ , respectively. On any existing machine, this is more efficient than division or remainder computations.

The implementation described above uses pointers and occupies a total space of  $O(M) = O(2^m)$ . This can be seen as follows. The recurrence is  $S(M) = O(\sqrt{M}) + (\sqrt{M} + 1) \cdot S(\sqrt{M})$ . Resolving that would lead to  $S(M) \in (1 + \sqrt{M})^{\log \log M} + \log \log M \cdot O(\sqrt{M})$ . One can, fortunately, also show that  $S(M) = M - 2$  by induction.<sup>[3]</sup>

In practical implementations, especially on machines with *shift-by-k* and *find first zero* instructions, performance can further be improved by switching to a bit array once  $m$  equal to the word size (or a small multiple thereof) is reached. Since all operations on a single word are constant time, this does not affect the asymptotic performance, but it does avoid the majority of the pointer storage and several pointer dereferences, achieving a significant practical savings in time and space with this trick.

An obvious optimization of vEB trees is to discard empty subtrees. This makes vEB trees quite compact when they contain many elements, because no subtrees are created until something needs to be added to them. Initially, each element added creates about  $\log(m)$  new trees containing about  $m/2$  pointers all together. As the tree grows, more and more subtrees are reused, especially the larger ones. In a full tree of  $2^m$  elements, only  $O(2^m)$  space is used. Moreover, unlike a binary search tree, most of this space is being used to store data: even for billions of elements, the pointers in a full vEB tree number in the thousands.

However, for small trees the overhead associated with vEB trees is enormous: on the order of  $2^{m/2}$ . This is one reason why they are not popular in practice. One way of addressing this limitation is to use only a fixed number of bits per level, which results in a trie. Other structures, including y-fast tries and x-fast tries have been proposed that have comparable update and query times but use only  $O(n)$  or  $O(n \log M)$  space where  $n$  is the number of elements stored in the data structure.

## References

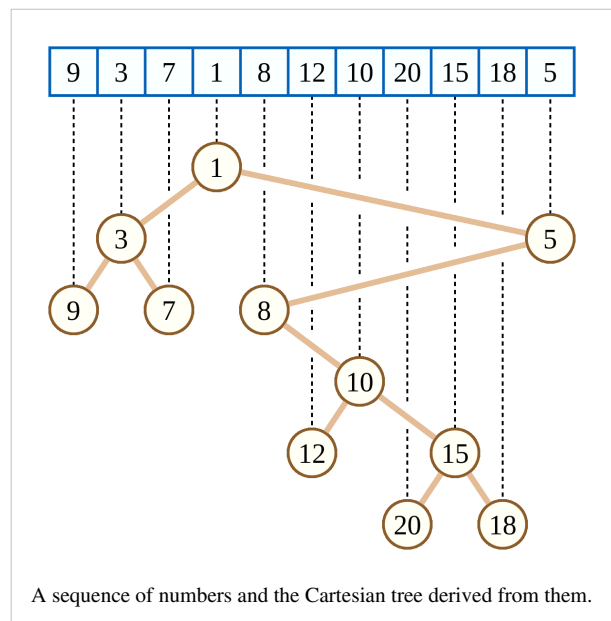
- [1] Peter van Emde Boas: *Preserving order in a forest in less than logarithmic time* (*Proceedings of the 16th Annual Symposium on Foundations of Computer Science* 10: 75-84, 1975)
- [2] Gudmund Skovbjerg Frandsen: *Dynamic algorithms: Course notes on van Emde Boas trees (PDF)* (<http://www.daimi.au.dk/~gudmund/dynamicF04/vEB.pdf>) (University of Aarhus, Department of Computer Science)
- [3] Rex, A. "Determining the space complexity of van Emde Boas trees" (<http://mathoverflow.net/questions/2245/determining-the-space-complexity-of-van-emde-boas-trees>). . Retrieved 2011-05-27.

## Further reading

- Erik Demaine, Shantonu Sen, and Jeff Lindy. Massachusetts Institute of Technology. 6.897: Advanced Data Structures (Spring 2003). Lecture 1 notes: Fixed-universe successor problem, van Emde Boas ([http://theory.csail.mit.edu/classes/6.897/spring03/scribe\\_notes/L1/lecture1.pdf](http://theory.csail.mit.edu/classes/6.897/spring03/scribe_notes/L1/lecture1.pdf)). Lecture 2 notes: More van Emde Boas, ... ([http://theory.csail.mit.edu/classes/6.897/spring03/scribe\\_notes/L2/lecture2.pdf](http://theory.csail.mit.edu/classes/6.897/spring03/scribe_notes/L2/lecture2.pdf)).
- van Emde Boas, Peter; Kaas, R.; Zijlstra, E. (1976). "Design and implementation of an efficient priority queue" ([http://ervain.cpsc.ucalgary.ca/Link to Documents/Design and Implementation of an Efficient Priority Queue.pdf](http://ervain.cpsc.ucalgary.ca/Link%20to%20Documents/Design%20and%20Implementation%20of%20an%20Efficient%20Priority%20Queue.pdf)). *Mathematical Systems Theory* **10**: 99–127. doi:10.1007/BF01683268.

## Cartesian tree

In computer science, a **Cartesian tree** is a binary tree derived from a sequence of numbers; it can be uniquely defined from the properties that it is heap-ordered and that a symmetric (in-order) traversal of the tree returns the original sequence. Introduced by Vuillemin (1980) in the context of geometric range searching data structures, Cartesian trees have also been used in the definition of the treap and randomized binary search tree data structures for binary search problems. The Cartesian tree for a sequence may be constructed in linear time using a stack-based algorithm for finding all nearest smaller values in a sequence.



## Definition

The Cartesian tree for a sequence of distinct numbers can be uniquely defined by the following properties:

1. The Cartesian tree for a sequence has one node for each number in the sequence. Each node is associated with a single sequence value.
2. A symmetric (in-order) traversal of the tree results in the original sequence. That is, the left subtree consists of the values earlier than the root in the sequence order, while the right subtree consists of the values later than the root, and a similar ordering constraint holds at each lower node of the tree.
3. The tree has the heap property: the parent of any non-root node has a smaller value than the node itself.<sup>[1]</sup>

Based on the heap property, the root of the tree must be the smallest number in the sequence. From this, the tree itself may also be defined recursively: the root is the minimum value of the sequence, and the left and right subtrees are the Cartesian trees for the subsequences to the left and right of the root value. Therefore, the three properties above uniquely define the Cartesian tree.

If a sequence of numbers contains repetitions, the Cartesian tree may be defined by determining a consistent tie-breaking rule (for instance, determining that the first of two equal elements is treated as the smaller of the two) before applying the above rules.

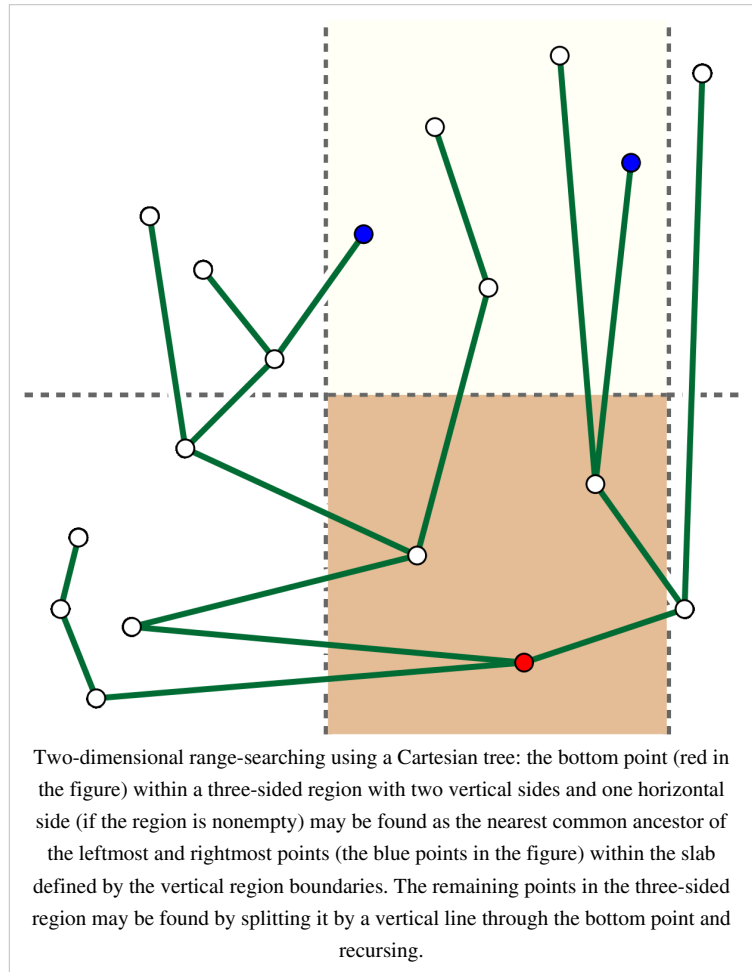
An example of a Cartesian tree is shown in the figure above.

## Range searching and lowest common ancestors

Cartesian trees may be used as part of an efficient data structure for range minimum queries, a range searching problem involving queries that ask for the minimum value in a contiguous subsequence of the original sequence.<sup>[2]</sup> In a Cartesian tree, this minimum value may be found at the lowest common ancestor of the leftmost and rightmost values in the subsequence. For instance, in the subsequence (12,10,20,15) of the sequence shown in the first illustration, the minimum value of the subsequence (10) forms the lowest common ancestor of the leftmost and rightmost values (12 and 15). Because lowest common ancestors may be found in constant time per query, using a data structure that takes linear space to store and that may be constructed in linear time,<sup>[3]</sup> the same bounds hold for the range minimization problem.

Bender & Farach-Colton (2000) reversed this relationship between the two data structure problems by showing that lowest common ancestors in an input tree could be solved efficiently applying a non-tree-based technique for range minimization. Their data structure uses an Euler tour technique to transform the input tree into a sequence and then finds range minima in the resulting sequence. The sequence resulting from this transformation has a special form (adjacent numbers, representing heights of adjacent nodes in the tree, differ by  $\pm 1$ ) which they take advantage of in their data structure; to solve the range minimization problem for sequences that do not have this special form, they use Cartesian trees to transform the range minimization problem into a lowest common ancestor problem, and then apply the Euler tour technique to transform the problem again into one of range minimization for sequences with this special form.

The same range minimization problem may also be given an alternative interpretation in terms of two dimensional range searching. A collection of finitely many points in the Cartesian plane may be used to form a Cartesian tree, by sorting the points by their  $x$ -coordinates and using the  $y$ -coordinates in this order as the sequence of values from which this tree is formed. If  $S$  is the subset of the input points within some vertical slab defined by the inequalities  $L \leq x \leq R$ ,  $p$  is the leftmost point in  $S$  (the one with minimum  $x$ -coordinate), and  $q$  is the rightmost point in  $S$  (the one with maximum  $x$ -coordinate) then the lowest common ancestor of  $p$  and  $q$  in the Cartesian tree is the bottommost point in the slab. A three-sided range query, in which the task is to list all points within a region bounded by the three



inequalities  $L \leq x \leq R$  and  $y \leq T$ , may be answered by finding this bottommost point  $b$ , comparing its  $y$ -coordinate to  $T$ , and (if the point lies within the three-sided region) continuing recursively in the two slabs bounded between  $p$  and  $b$  and between  $b$  and  $q$ . In this way, after the leftmost and rightmost points in the slab are identified, all points within the three-sided region may be listed in constant time per point.<sup>[4]</sup>

The same construction, of lowest common ancestors in a Cartesian tree, makes it possible to construct a data structure with linear space that allows the distances between pairs of points in any ultrametric space to be queried in constant time per query. The distance within an ultrametric is the same as the minimax path weight in the minimum spanning tree of the metric.<sup>[5]</sup> From the minimum spanning tree, one can construct a Cartesian tree, the root node of which represents the heaviest edge of the minimum spanning tree. Removing this edge partitions the minimum spanning tree into two subtrees, and Cartesian trees recursively constructed for these two subtrees form the children of the root node of the Cartesian tree. The leaves of the Cartesian tree represent points of the metric space, and the lowest common ancestor of two leaves in the Cartesian tree is the heaviest edge between those two points in the minimum spanning tree, which has weight equal to the distance between the two points. Once the minimum spanning tree has been found and its edge weights sorted, the Cartesian tree may be constructed in linear time.<sup>[6]</sup>

## Treaps

*Main article: Treap*

Because a Cartesian tree is a binary tree, it is natural to use it as a binary search tree for an ordered sequence of values. However, defining a Cartesian tree based on the same values that form the search keys of a binary search tree does not work well: the Cartesian tree of a sorted sequence is just a path, rooted at its leftmost endpoint, and binary searching in this tree degenerates to sequential search in the path. However, it is possible to generate more-balanced search trees by generating *priority* values for each search key that are different than the key itself, sorting the inputs by their key values, and using the corresponding sequence of priorities to generate a Cartesian tree. This construction may equivalently be viewed in the geometric framework described above, in which the  $x$ -coordinates of a set of points are the search keys and the  $y$ -coordinates are the priorities.

This idea was applied by Seidel & Aragon (1996), who suggested the use of random numbers as priorities. The data structure resulting from this random choice is called a treap, due to its combination of binary search tree and binary heap features. An insertion into a treap may be performed by inserting the new key as a leaf of an existing tree, choosing a priority for it, and then performing tree rotation operations along a path from the node to the root of the tree to repair any violations of the heap property caused by this insertion; a deletion may similarly be performed by a constant amount of change to the tree followed by a sequence of rotations along a single path in the tree.

If the priorities of each key are chosen randomly and independently once whenever the key is inserted into the tree, the resulting Cartesian tree will have the same properties as a random binary search tree, a tree computed by inserting the keys in a randomly chosen permutation starting from an empty tree, with each insertion leaving the previous tree structure unchanged and inserting the new node as a leaf of the tree. Random binary search trees had been studied for much longer, and are known to behave well as search trees (they have logarithmic depth with high probability); the same good behavior carries over to treaps. It is also possible, as suggested by Aragon and Seidel, to reprioritize frequently-accessed nodes, causing them to move towards the root of the treap and speeding up future accesses for the same keys.

## Efficient construction

A Cartesian tree may be constructed in linear time from its input sequence. One method is to simply process the sequence values in left-to-right order, maintaining the Cartesian tree of the nodes processed so far, in a structure that allows both upwards and downwards traversal of the tree. To process each new value  $x$ , start at the node representing the value prior to  $x$  in the sequence and follow the path from this node to the root of the tree until finding a value  $y$  smaller than  $x$ . This node  $y$  is the parent of  $x$ , and the previous right child of  $y$  becomes the new left child of  $x$ . The total time for this procedure is linear, because the time spent searching for the parent  $y$  of each new node  $x$  can be charged against the number of nodes that are removed from the rightmost path in the tree.<sup>[4]</sup>

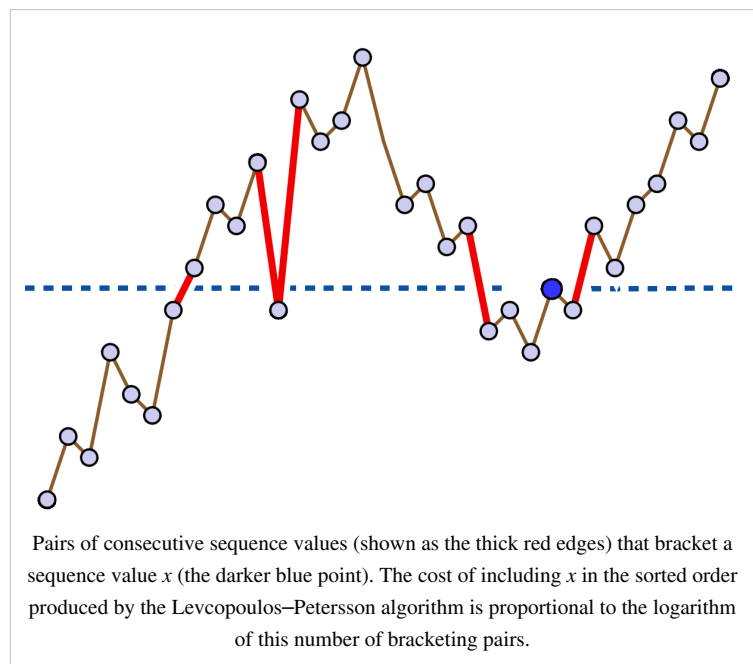
An alternative linear-time construction algorithm is based on the all nearest smaller values problem. In the input sequence, one may define the *left neighbor* of a value  $x$  to be the value that occurs prior to  $x$ , is smaller than  $x$ , and is closer in position to  $x$  than any other smaller value. The *right neighbor* is defined symmetrically. The sequence of left neighbors may be found by an algorithm that maintains a stack containing a subsequence of the input. For each new sequence value  $x$ , the stack is popped until it is empty or its top element is smaller than  $x$ , and then  $x$  is pushed onto the stack. The left neighbor of  $x$  is the top element at the time  $x$  is pushed. The right neighbors may be found by applying the same stack algorithm to the reverse of the sequence. The parent of  $x$  in the Cartesian tree is either the left neighbor of  $x$  or the right neighbor of  $x$ , whichever exists and has a larger value. The left and right neighbors may also be constructed efficiently by parallel algorithms, so this formulation may be used to develop efficient parallel algorithms for Cartesian tree construction.<sup>[7]</sup>

## Application in sorting

Levcopoulos & Petersson (1989) describe a sorting algorithm based on Cartesian trees. They describe the algorithm as based on a tree with the maximum at the root, but it may be modified straightforwardly to support a Cartesian tree with the convention that the minimum value is at the root. For consistency, it is this modified version of the algorithm that is described below.

The Levcopoulos–Petersson algorithm can be viewed as a version of selection sort or heap sort that maintains a priority queue of candidate minima, and that at each step finds and removes the minimum value in this queue, moving this value to the end of an output sequence. In their algorithm, the priority queue consists only of elements whose parent in the Cartesian tree has already been found and removed. Thus, the algorithm consists of the following steps:

1. Construct a Cartesian tree for the input sequence
2. Initialize a priority queue, initially containing only the tree root
3. While the priority queue is non-empty:
  - Find and remove the minimum value  $x$  in the priority queue
  - Add  $x$  to the output sequence
  - Add the Cartesian tree children of  $x$  to the priority queue



As Levcopoulos and Petersson show, for input sequences that are already nearly sorted, the size of the priority queue will remain small, allowing this method to take advantage of the nearly-sorted input and run more quickly. Specifically, the worst-case running time of this algorithm is  $O(n \log k)$ , where  $k$  is the average, over all values  $x$  in the sequence, of the number of consecutive pairs of sequence values that bracket  $x$ . They also prove a lower bound stating that, for any  $n$  and  $k = \omega(1)$ , any comparison-based sorting algorithm must use  $\Omega(n \log k)$  comparisons for some inputs.

## History

Cartesian trees were introduced and named by Vuillemin (1980). The name is derived from the Cartesian coordinate system for the plane: in Vuillemin's version of this structure, as in the two-dimensional range searching application discussed above, a Cartesian tree for a point set has the sorted order of the points by their  $x$ -coordinates as its symmetric traversal order, and it has the heap property according to the  $y$ -coordinates of the points. Gabow, Bentley & Tarjan (1984) and subsequent authors followed the definition here in which a Cartesian tree is defined from a sequence; this change generalizes the geometric setting of Vuillemin to allow sequences other than the sorted order of  $x$ -coordinates, and allows the Cartesian tree to be applied to non-geometric problems as well.

## Notes

- [1] In some references, the ordering is reversed, so the parent of any node always has a larger value and the root node holds the maximum value.
- [2] Gabow, Bentley & Tarjan (1984); Bender & Farach-Colton (2000).
- [3] Harel & Tarjan (1984); Schieber & Vishkin (1988).
- [4] Gabow, Bentley & Tarjan (1984).
- [5] Hu (1961); Leclerc (1981)
- [6] Demaine, Landau & Weimann (2009).
- [7] Berkman, Schieber & Vishkin (1993).

## References

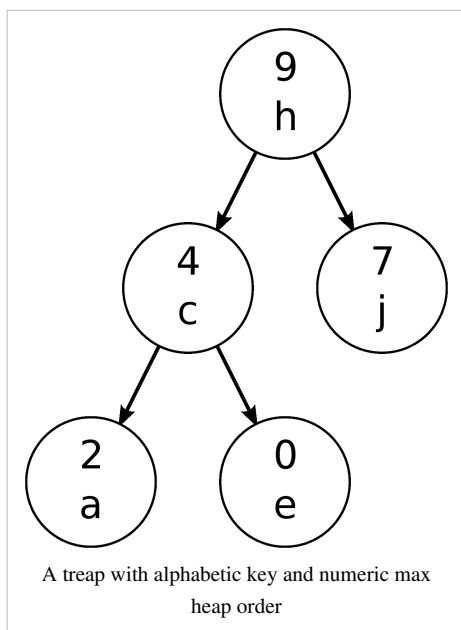
- Bender, Michael A.; Farach-Colton, Martin (2000), "The LCA problem revisited" (<http://www.cs.sunysb.edu/~bender/pub/lca.ps>), *Proceedings of the 4th Latin American Symposium on Theoretical Informatics*, Springer-Verlag, Lecture Notes in Computer Science 1776, pp. 88–94.
- Berkman, Omer; Schieber, Baruch; Vishkin, Uzi (1993), "Optimal doubly logarithmic parallel algorithms based on finding all nearest smaller values", *Journal of Algorithms* **14** (3): 344–370, doi:10.1006/jagm.1993.101.
- Demaine, Erik D.; Landau, Gad M.; Weimann, Oren (2009), "On cartesian trees and range minimum queries", *Automata, Languages and Programming, 36th International Colloquium, ICALP 2009, Rhodes, Greece, July 5-12, 2009*, Lecture Notes in Computer Science, **5555**, pp. 341–353, doi:10.1007/978-3-642-02927-1\_29, ISBN 978-3-642-02926-4.
- Fischer, Johannes; Heun, Volker (2006), "Theoretical and Practical Improvements on the RMQ-Problem, with Applications to LCA and LCE", *Proceedings of the 17th Annual Symposium on Combinatorial Pattern Matching*, Lecture Notes in Computer Science, **4009**, Springer-Verlag, pp. 36–48, doi:10.1007/11780441\_5, ISBN 978-3-540-35455-0
- Fischer, Johannes; Heun, Volker (2007), "A New Succinct Representation of RMQ-Information and Improvements in the Enhanced Suffix Array.", *Proceedings of the International Symposium on Combinatorics, Algorithms, Probabilistic and Experimental Methodologies*, Lecture Notes in Computer Science, **4614**, Springer-Verlag, pp. 459–470, doi:10.1007/978-3-540-74450-4\_41, ISBN 978-3-540-74449-8
- Gabow, Harold N.; Bentley, Jon Louis; Tarjan, Robert E. (1984), "Scaling and related techniques for geometry problems", *STOC '84: Proc. 16th ACM Symp. Theory of Computing*, New York, NY, USA: ACM, pp. 135–143, doi:10.1145/800057.808675, ISBN 0-89791-133-4.

- Harel, Dov; Tarjan, Robert E. (1984), "Fast algorithms for finding nearest common ancestors", *SIAM Journal on Computing* **13** (2): 338–355, doi:10.1137/0213024.
- Hu, T. C. (1961), "The maximum capacity route problem", *Operations Research* **9** (6): 898–900, doi:10.1287/opre.9.6.898, JSTOR 167055.
- Leclerc, Bruno (1981), "Description combinatoire des ultramétries" (in French), *Centre de Mathématique Sociale. École Pratique des Hautes Études. Mathématiques et Sciences Humaines* (73): 5–37, 127, MR623034.
- Levkopoulos, Christos; Petersson, Ola (1989), "Heapsort - Adapted for Presorted Files", *WADS '89: Proceedings of the Workshop on Algorithms and Data Structures*, Lecture Notes in Computer Science, **382**, London, UK: Springer-Verlag, pp. 499–509, doi:10.1007/3-540-51542-9\_41.
- Seidel, Raimund; Aragon, Cecilia R. (1996), "Randomized Search Trees" (<http://citeseer.ist.psu.edu/seidel96randomized.html>), *Algorithmica* **16** (4/5): 464–497, doi:10.1007/s004539900061.
- Schieber, Baruch; Vishkin, Uzi (1988), "On finding lowest common ancestors: simplification and parallelization", *SIAM Journal on Computing* **17** (6): 1253–1262, doi:10.1137/0217079.
- Vuillemin, Jean (1980), "A unifying look at data structures", *Commun. ACM* (New York, NY, USA: ACM) **23** (4): 229–239, doi:10.1145/358841.358852.

## Treap

In computer science, the **treap** and the **randomized binary search tree** are two closely related forms of binary search tree data structures that maintain a dynamic set of ordered keys and allow binary searches among the keys. After any sequence of insertions and deletions of keys, the shape of the tree is a random variable with the same probability distribution as a random binary tree; in particular, with high probability its height is proportional to the logarithm of the number of keys, so that each search, insertion, or deletion operation takes logarithmic time to perform.

### Treap



The treap was first described by Cecilia R. Aragon and Raimund Seidel in 1989;<sup>[1][2]</sup> its name is a portmanteau of tree and heap. It is a Cartesian tree<sup>[3]</sup> in which each key is given a (randomly chosen) numeric priority. As with any binary search tree, the inorder traversal order of the nodes is the same as the sorted order of the keys. The structure of the tree is determined by the requirement that it be heap-ordered: that is, the priority number for any non-leaf node must be greater than or equal to the priority of its children. Thus, as with Cartesian trees more generally, the root node is the maximum-priority node, and its left and right subtrees are formed in the same manner from the subsequences of the sorted order to the left and right of that node.

An equivalent way of describing the treap is that it could be formed by inserting the nodes highest-priority-first into a binary search tree without doing any rebalancing. Therefore, if the priorities are independent random numbers (from a distribution over a large enough

space of possible priorities to ensure that two nodes are very unlikely to have the same priority) then the shape of a treap has the same probability distribution as the shape of a random binary search tree, a search tree formed by inserting the nodes without rebalancing in a randomly chosen insertion order. Because random binary search trees



are known to have logarithmic height with high probability, the same is true for treaps.

Specifically, the treap supports the following operations:

- To search for a given key value, apply a standard binary search algorithm in a binary search tree, ignoring the priorities.
- To insert a new key  $x$  into the treap, generate a random priority  $y$  for  $x$ . Binary search for  $x$  in the tree, and create a new node at the leaf position where the binary search determines a node for  $x$  should exist. Then, as long as  $x$  is not the root of the tree and has a larger priority number than its parent  $z$ , perform a tree rotation that reverses the parent-child relation between  $x$  and  $z$ .
- To delete a node  $x$  from the treap, if  $x$  is a leaf of the tree, simply remove it. If  $x$  has a single child  $z$ , remove  $x$  from the tree and make  $z$  be the child of the parent of  $x$  (or make  $z$  the root of the tree if  $x$  had no parent). Finally, if  $x$  has two children, swap its position in the tree with the position of its immediate successor  $z$  in the sorted order, resulting in one of the previous cases. In this final case, the swap may violate the heap-ordering property for  $z$ , so additional rotations may need to be performed to restore this property.
- To split a treap into two smaller treaps, those smaller than key  $x$ , and those larger than key  $x$ , insert  $x$  into the treap with maximum priority—larger than the priority of any node in the treap. After this insertion,  $x$  will be the root node of the treap, all values less than  $x$  will be found in the left subtree, and all values greater than  $x$  will be found in the right subtree. This costs as much as a single insertion into the treap.
- Merging two treaps (assumed to be the product of a former split), one can safely assume that the greatest value in the first treap is less than the smallest value in the second treap. Insert a value  $x$ , such that  $x$  is larger than this max-value in the first treap, and smaller than the min-value in the second treap, and assign it the minimum priority. After insertion it will be a leaf node, and can easily be deleted. The result is one treap merged from the two original treaps. This is effectively "undoing" a split, and costs the same.

Aragon and Seidel also suggest assigning higher priorities to frequently accessed nodes, for instance by a process that, on each access, chooses a random number and replaces the priority of the node with that number if it is higher than the previous priority. This modification would cause the tree to lose its random shape; instead, frequently accessed nodes would be more likely to be near the root of the tree, causing searches for them to be faster.

Blelloch and Reid-Miller<sup>[4]</sup> describe an application of treaps to a problem of maintaining sets of items and performing set union, set intersection, and set difference operations, using a treap to represent each set. Naor and Nissim<sup>[5]</sup> describe another application, for maintaining authorization certificates in public-key cryptosystems.

## Randomized binary search tree

The randomized binary search tree, introduced by Martínez and Roura subsequently to the work of Aragon and Seidel on treaps<sup>[6]</sup>, stores the same nodes with the same random distribution of tree shape, but maintains different information within the nodes of the tree in order to maintain its randomized structure.

Rather than storing random priorities on each node, the randomized binary search tree stores at each node a small integer, the number of its descendants (counting itself as one); these numbers may be maintained during tree rotation operations at only a constant additional amount of time per rotation. When a key  $x$  is to be inserted into a tree that already has  $n$  nodes, the insertion algorithm chooses with probability  $1/(n + 1)$  to place  $x$  as the new root of the tree, and otherwise it calls the insertion procedure recursively to insert  $x$  within the left or right subtree (depending on whether its key is less than or greater than the root). The numbers of descendants are used by the algorithm to calculate the necessary probabilities for the random choices at each step. Placing  $x$  at the root of a subtree may be performed either as in the treap by inserting it at a leaf and then rotating it upwards, or by an alternative algorithm described by Martínez and Roura that splits the subtree into two pieces to be used as the left and right children of the new node.

The deletion procedure for a randomized binary search tree uses the same information per node as the insertion procedure, and like the insertion procedure it makes a sequence of  $O(\log n)$  random decisions in order to join the two

subtrees descending from the left and right children of the deleted node into a single tree. If the left or right subtree of the node to be deleted is empty, the join operation is trivial; otherwise, the left or right child of the deleted node is selected as the new subtree root with probability proportional to its number of descendants, and the join proceeds recursively.

## Comparison

The information stored per node in the randomized binary tree is simpler than in a treap (a small integer rather than a high-precision random number), but it makes a greater number of calls to the random number generator ( $O(\log n)$  calls per insertion or deletion rather than one call per insertion) and the insertion procedure is slightly more complicated due to the need to update the numbers of descendants per node. A minor technical difference is that, in a treap, there is a small probability of a collision (two keys getting the same priority), and in both cases there will be statistical differences between a true random number generator and the pseudo-random number generator typically used on digital computers. However, in any case the differences between the theoretical model of perfect random choices used to design the algorithm and the capabilities of actual random number generators are vanishingly small.

Although the treap and the randomized binary search tree both have the same random distribution of tree shapes after each update, the history of modifications to the trees performed by these two data structures over a sequence of insertion and deletion operations may be different. For instance, in a treap, if the three numbers 1, 2, and 3 are inserted in the order 1, 3, 2, and then the number 2 is deleted, the remaining two nodes will have the same parent-child relationship that they did prior to the insertion of the middle number. In a randomized binary search tree, the tree after the deletion is equally likely to be either of the two possible trees on its two nodes, independently of what the tree looked like prior to the insertion of the middle number.

## References

- [1] Aragon, Cecilia R.; Seidel, Raimund (1989), "Randomized Search Trees" (<http://faculty.washington.edu/aragon/pubs/rst89.pdf>), *Proc. 30th Symp. Foundations of Computer Science (FOCS 1989)*, Washington, D.C.: IEEE Computer Society Press, pp. 540–545, doi:10.1109/SFCS.1989.63531, ISBN 0-8186-1982-1,
- [2] Seidel, Raimund; Aragon, Cecilia R. (1996), "Randomized Search Trees" (<http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.30.8602>), *Algorithmica* **16** (4/5): 464–497, doi:10.1007/s004539900061,
- [3] Vuillemin, Jean (1980), "A unifying look at data structures", *Commun. ACM* (New York, NY, USA: ACM) **23** (4): 229–239, doi:10.1145/358841.358852.
- [4] Blelloch, Guy E.; Reid-Miller, Margaret, (1998), "Fast set operations using treaps", *Proc. 10th ACM Symp. Parallel Algorithms and Architectures (SPAA 1998)*, New York, NY, USA: ACM, pp. 16–26, doi:10.1145/277651.277660, ISBN 0-89791-989-0.
- [5] Naor, M.; Nissim, K. (April 2000), "Certificate revocation and certificate update" (<http://eprints.kfupm.edu.sa/29443/1/29443.pdf>), *IEEE Journal on Selected Areas in Communications* **18** (4): 561–570, doi:10.1109/49.839932, .
- [6] Martínez, Conrado; Roura, Salvador (1997), "Randomized binary search trees" (<http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.17.243>), *Journal of the ACM* **45** (2): 288–323, doi:10.1145/274787.274812,

## External links

- Collection of treap references and info (<http://faculty.washington.edu/aragon/treaps.html>) by Cecilia Aragon
- Treap Applet (<http://people.ksp.sk/~kuko/bak/index.html>) by Kubo Kovac
- Animated treap (<http://www.ibr.cs.tu-bs.de/lehre/ss98/audii/applets/BST/Treap-Example.html>)
- Randomized binary search trees (<http://www.cs.uiuc.edu/class/sp09/cs473/notes/08-treaps.pdf>). Lecture notes from a course by Jeff Erickson at UIUC. Despite the title, this is primarily about treaps and skip lists; randomized binary search trees are mentioned only briefly.
- A high performance key-value store based on treap (<http://code.google.com/p/treapdb/>) by Junyi Sun
- VB6 implementation of treaps (<http://www.fernando-rodriguez.com/a-high-performance-alternative-to-dictionary>). Visual basic 6 implementation of treaps as a COM object.

- ActionScript3 implementation of a treap (<http://code.google.com/p/as3-commons/source/browse/trunk/as3-commons-collections/src/main/actionsript/org/as3commons/collections/Treap.as>)
  - Pure Python and Cython in-memory treap and dupltreap (<http://stromberg.dnsalias.org/~dstromberg/treap/>)
  - Treaps in C# (<http://www.codeproject.com/Articles/8184/Treaps-in-C>). By Roy Clemmons
-

# B-trees

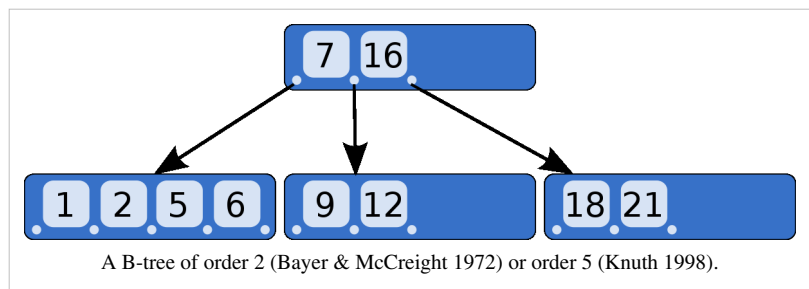
## B-tree

B-tree		
Type	Tree	
Invented	1972	
Invented by	Rudolf Bayer, Edward M. McCreight	
Time complexity in big O notation		
	Average	Worst case
Space	$O(n)$	$O(n)$
Search	$O(\log n)$	$O(\log n)$
Insert	$O(\log n)$	$O(\log n)$
Delete	$O(\log n)$	$O(\log n)$

In computer science, a **B-tree** is a tree data structure that keeps data sorted and allows searches, sequential access, insertions, and deletions in logarithmic time. The B-tree is a generalization of a binary search tree in that a node can have more than two children. (Comer 1979, p. 123) Unlike self-balancing binary search trees, the B-tree is optimized for systems that read and write large blocks of data. It is commonly used in databases and filesystems.

### Overview

In B-trees, internal (non-leaf) nodes can have a variable number of child nodes within some pre-defined range. When data is inserted or removed from a node, its number of child nodes changes. In order to maintain the pre-defined range, internal nodes may be joined or split. Because a range of



child nodes is permitted, B-trees do not need re-balancing as frequently as other self-balancing search trees, but may waste some space, since nodes are not entirely full. The lower and upper bounds on the number of child nodes are typically fixed for a particular implementation. For example, in a 2-3 B-tree (often simply referred to as a **2-3 tree**), each internal node may have only 2 or 3 child nodes.

Each internal node of a B-tree will contain a number of keys. Usually, the number of keys is chosen to vary between  $d$  and  $2d$ . In practice, the keys take up the most space in a node. The factor of 2 will guarantee that nodes can be split or combined. If an internal node has  $2d$  keys, then adding a key to that node can be accomplished by splitting the  $2d$  key node into two  $d$  key nodes and adding the key to the parent node. Each split node has the required minimum number of keys. Similarly, if an internal node and its neighbor each have  $d$  keys, then a key may be deleted from the internal node by combining with its neighbor. Deleting the key would make the internal node have  $d - 1$  keys; joining the neighbor would add  $d$  keys plus one more key brought down from the neighbor's parent.

The result is an entirely full node of  $2d$  keys.

The number of branches (or child nodes) from a node will be one more than the number of keys stored in the node. In a 2-3 B-tree, the internal nodes will store either one key (with two child nodes) or two keys (with three child nodes). A B-tree is sometimes described with the parameters  $(d + 1) - (2d + 1)$  or simply with the highest branching order,  $(2d + 1)$ .

A B-tree is kept balanced by requiring that all leaf nodes be at the same depth. This depth will increase slowly as elements are added to the tree, but an increase in the overall depth is infrequent, and results in all leaf nodes being one more node further away from the root.

B-trees have substantial advantages over alternative implementations when node access times far exceed access times within nodes, because then the cost of accessing the node may be amortized over multiple operations within the node. This usually occurs when the nodes are in secondary storage such as disk drives. By maximizing the number of child nodes within each internal node, the height of the tree decreases and the number of expensive node accesses is reduced. In addition, rebalancing the tree occurs less often. The maximum number of child nodes depends on the information that must be stored for each child node and the size of a full disk block or an analogous size in secondary storage. While 2-3 B-trees are easier to explain, practical B-trees using secondary storage want a large number of child nodes to improve performance.

## Variants

The term **B-tree** may refer to a specific design or it may refer to a general class of designs. In the narrow sense, a B-tree stores keys in its internal nodes but need not store those keys in the records at the leaves. The general class includes variations such as the  $B^+$ -tree and the  $B^*$ -tree.

- In the  $B^+$ -tree, copies of the keys are stored in the internal nodes; the keys and records are stored in leaves; in addition, a leaf node may include a pointer to the next leaf node to speed sequential access. (Comer 1979, p. 129)
- The  $B^*$ -tree balances more neighboring internal nodes to keep the internal nodes more densely packed. (Comer 1979, p. 129) This variant requires non-root nodes to be at least  $2/3$  full instead of  $1/2$ . (Knuth 1998, p. 488) To maintain this, instead of immediately splitting up a node when it gets full, its keys are shared with a node next to it. When both nodes are full, then the two nodes are split into three.
- Counted B-trees store, with each pointer within the tree, the number of nodes in the subtree below that pointer.<sup>[1]</sup> This allows rapid searches for the Nth record in key order, or counting the number of records between any two records, and various other related operations.

## Etymology unknown

Rudolf Bayer and Ed McCreight invented the B-tree while working at Boeing Research Labs in 1971 (Bayer & McCreight 1972), but they did not explain what, if anything, the *B* stands for. Douglas Comer explains:

The origin of "B-tree" has never been explained by the authors. As we shall see, "balanced," "broad," or "bushy" might apply. Others suggest that the "B" stands for Boeing. Because of his contributions, however, it seems appropriate to think of B-trees as "Bayer"-trees. (Comer 1979, p. 123 footnote 1)

Donald Knuth speculates on the etymology of B-trees in his May, 1980 lecture on the topic "CS144C classroom lecture about disk storage and B-trees", suggesting the "B" may have originated from Boeing or from Bayer's name.<sup>[2]</sup>

## The database problem

### Time to search a sorted file

Usually, sorting and searching algorithms have been characterized by the number of comparison operations that must be performed using order notation. A binary search of a sorted table with  $N$  records, for example, can be done in  $O(\log_2 N)$  comparisons. If the table had 1,000,000 records, then a specific record could be located with about 20 comparisons:  $\log_2 1,000,000 = 19.931\dots$

Large databases have historically been kept on disk drives. The time to read a record on a disk drive can dominate the time needed to compare keys once the record is available. The time to read a record from a disk drive involves a seek time and a rotational delay. The seek time may be 0 to 20 or more milliseconds, and the rotational delay averages about half the rotation period. For a 7200 RPM drive, the rotation period is 8.33 milliseconds. For a drive such as the Seagate ST3500320NS, the track-to-track seek time is 0.8 milliseconds and the average reading seek time is 8.5 milliseconds.<sup>[3]</sup> For simplicity, assume reading from disk takes about 10 milliseconds.

Naively, then, the time to locate one record out of a million would take 20 disk reads times 10 milliseconds per disk read, which is 0.2 seconds.

The time won't be that bad because individual records are grouped together in a disk **block**. A disk block might be 16 kilobytes. If each record is 160 bytes, then 100 records could be stored in each block. The disk read time above was actually for an entire block. Once the disk head is in position, one or more disk blocks can be read with little delay. With 100 records per block, the last 6 or so comparisons don't need to do any disk reads—the comparisons are all within the last disk block read.

To speed the search further, the first 13 to 14 comparisons (which each required a disk access) must be sped up.

### An index speeds the search

A significant improvement can be made with an index. In the example above, initial disk reads narrowed the search range by a factor of two. That can be improved substantially by creating an auxiliary index that contains the first record in each disk block (sometimes called a sparse index). This auxiliary index would be 1% of the size of the original database, but it can be searched more quickly. Finding an entry in the auxiliary index would tell us which block to search in the main database; after searching the auxiliary index, we would have to search only that one block of the main database—at a cost of one more disk read. The index would hold 10,000 entries, so it would take at most 14 comparisons. Like the main database, the last 6 or so comparisons in the aux index would be on the same disk block. The index could be searched in about 8 disk reads, and the desired record could be accessed in 9 disk reads.

The trick of creating an auxiliary index can be repeated to make an auxiliary index to the auxiliary index. That would make an aux-aux index that would need only 100 entries and would fit in one disk block.

Instead of reading 14 disk blocks to find the desired record, we only need to read 3 blocks. Reading and searching the first (and only) block of the aux-aux index identifies the relevant block in aux-index. Reading and searching that aux-index block identifies the relevant block in the main database. Instead of 150 milliseconds, we need only 30 milliseconds to get the record.

The auxiliary indices have turned the search problem from a binary search requiring roughly  $\log_2 N$  disk reads to one requiring only  $\log_b N$  disk reads where  $b$  is the blocking factor (the number of entries per block:  $b = 100$  entries per block;  $\log_b 1,000,000 = 3$  reads).

In practice, if the main database is being frequently searched, the aux-aux index and much of the aux index may reside in a disk cache, so they would not incur a disk read.

## Insertions and deletions cause trouble

If the database does not change, then compiling the index is simple to do, and the index need never be changed. If there are changes, then managing the database and its index becomes more complicated.

Deleting records from a database doesn't cause much trouble. The index can stay the same, and the record can just be marked as deleted. The database stays in sorted order. If there are a lot of deletions, then the searching and storage become less efficient.

Insertions are a disaster in a sorted sequential file because room for the inserted record must be made. Inserting a record before the first record in the file requires shifting all of the records down one. Such an operation is just too expensive to be practical.

A trick is to leave some space lying around to be used for insertions. Instead of densely storing all the records in a block, the block can have some free space to allow for subsequent insertions. Those records would be marked as if they were "deleted" records.

Now, both insertions and deletions are fast as long as space is available on a block. If an insertion won't fit on the block, then some free space on some nearby block must be found and the auxiliary indices adjusted. The hope is that enough space is nearby such that a lot of blocks do not need to be reorganized. Alternatively, some out-of-sequence disk blocks may be used.

## The B-tree uses all those ideas

The B-tree uses all of the above ideas:

- It keeps records in sorted order for sequential traversing
- It uses a hierarchical index to minimize the number of disk reads
- It uses partially full blocks to speed insertions and deletions
- The index is elegantly adjusted with a recursive algorithm

In addition, a B-tree minimizes waste by making sure the interior nodes are at least  $\frac{1}{2}$  full. A B-tree can handle an arbitrary number of insertions and deletions.

## Technical description

### Terminology

Unfortunately, the literature on B-trees is not uniform in its use of terms relating to B-Trees. (Folk & Zoellick 1992, p. 362)

Bayer & McCreight (1972), Comer (1979), and others define the **order** of B-tree as the minimum number of keys in a non-root node. Folk & Zoellick (1992) points out that terminology is ambiguous because the maximum number of keys is not clear. An order 3 B-tree might hold a maximum of 6 keys or a maximum of 7 keys. (Knuth 1998, p. 483) avoids the problem by defining the **order** to be maximum number of children (which is one more than the maximum number of keys).

The term **leaf** is also inconsistent. Bayer & McCreight (1972) considered the leaf level to be the lowest level of keys, but Knuth considered the leaf level to be one level below the lowest keys. (Folk & Zoellick 1992, p. 363) There are many possible implementation choices. In some designs, the leaves may hold the entire data record; in other designs, the leaves may only hold pointers to the data record. Those choices are not fundamental to the idea of a B-tree.<sup>[4]</sup>

There are also unfortunate choices like using the variable  $k$  to represent the number of children when  $k$  could be confused with the number of keys.

For simplicity, most authors assume there are a fixed number of keys that fit in a node. The basic assumption is the key size is fixed and the node size is fixed. In practice, variable length keys may be employed. (Folk & Zoellick

1992, p. 379)

## Definition

According to Knuth's definition, a B-tree of order  $m$  (the maximum number of children for each node) is a tree which satisfies the following properties:

1. Every node has at most  $m$  children.
2. Every node (except root) has at least  $\lceil m/2 \rceil$  children.
3. The root has at least two children if it is not a leaf node.
4. A non-leaf node with  $k$  children contains  $k-1$  keys.
5. All leaves appear in the same level, and carry information.

Each internal node's elements act as separation values which divide its subtrees. For example, if an internal node has 3 child nodes (or subtrees) then it must have 2 separation values or elements:  $a_1$  and  $a_2$ . All values in the leftmost subtree will be less than  $a_1$ , all values in the middle subtree will be between  $a_1$  and  $a_2$ , and all values in the rightmost subtree will be greater than  $a_2$ .

## Internal nodes

Internal nodes are all nodes except for leaf nodes and the root node. They are usually represented as an ordered set of elements and child pointers. Every internal node contains a **maximum** of  $U$  children and a **minimum** of  $L$  children. Thus, the number of elements is always 1 less than the number of child pointers (the number of elements is between  $L-1$  and  $U-1$ ).  $U$  must be either  $2L$  or  $2L-1$ ; therefore each internal node is at least half full. The relationship between  $U$  and  $L$  implies that two half-full nodes can be joined to make a legal node, and one full node can be split into two legal nodes (if there's room to push one element up into the parent). These properties make it possible to delete and insert new values into a B-tree and adjust the tree to preserve the B-tree properties.

## The root node

The root node's number of children has the same upper limit as internal nodes, but has no lower limit. For example, when there are fewer than  $L-1$  elements in the entire tree, the root will be the only node in the tree, with no children at all.

## Leaf nodes

Leaf nodes have the same restriction on the number of elements, but have no children, and no child pointers.

A B-tree of depth  $n+1$  can hold about  $U$  times as many items as a B-tree of depth  $n$ , but the cost of search, insert, and delete operations grows with the depth of the tree. As with any balanced tree, the cost grows much more slowly than the number of elements.

Some balanced trees store values only at leaf nodes, and use different kinds of nodes for leaf nodes and internal nodes. B-trees keep values in every node in the tree, and may use the same structure for all nodes. However, since leaf nodes never have children, the B-trees benefit from improved performance if they use a specialized structure.



## Best case and worst case heights

Let  $h$  be the height of the classic B-tree. Let  $n > 0$  be the number of entries in the tree.<sup>[5]</sup> Let  $m$  be the maximum number of children a node can have. Each node can have at most  $m-1$  keys.

It can be shown (by induction for example) that a B-tree of height  $h$  with all its keys completely filled has  $n = m^h - 1$  keys. Hence, the best case height of a B-tree is:

$$\lceil \log_m(n+1) \rceil.$$

Let  $d$  be the minimum number of children an internal (non-root) node can have. For an ordinary B-tree,  $d = \lceil m/2 \rceil$ .

The worst case height of a B-tree is:

$$\lfloor \log_d((n+1)/2) + 1 \rfloor$$

Comer (1979, p. 127) and Cormen et al. (year, pp. 383–384) give a slightly different expression for the worst case height (perhaps because the root node is considered to have height 0).

$$h \leq \lfloor \log_d\left(\frac{n+1}{2}\right) \rfloor.$$

## Algorithms

### Search

Searching is similar to searching a binary search tree. Starting at the root, the tree is recursively traversed from top to bottom. At each level, the search chooses the child pointer (subtree) whose separation values are on either side of the search value.

Binary search is typically (but not necessarily) used within nodes to find the separation values and child tree of interest.

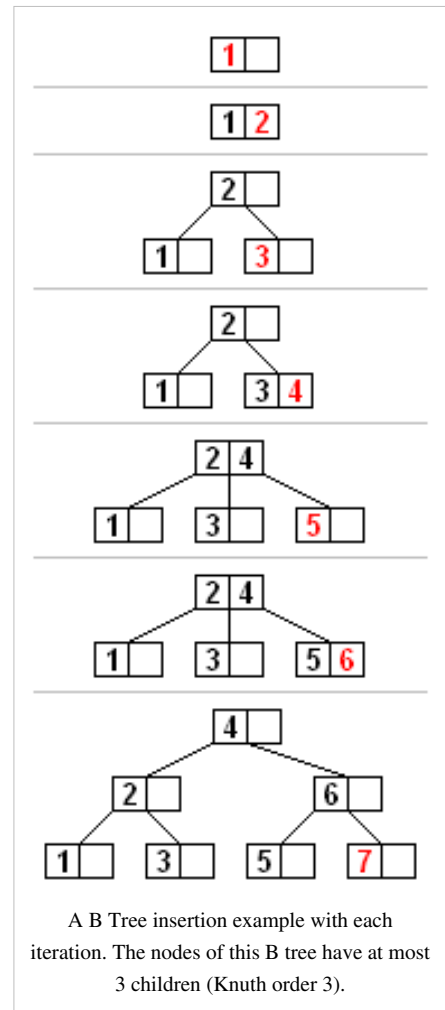
### Insertion

All insertions start at a leaf node. To insert a new element, search the tree to find the leaf node where the new element should be added. Insert the new element into that node with the following steps:

1. If the node contains fewer than the maximum legal number of elements, then there is room for the new element. Insert the new element in the node, keeping the node's elements ordered.
2. Otherwise the node is full, evenly split it into two nodes so:
  1. A single median is chosen from among the leaf's elements and the new element.
  2. Values less than the median are put in the new left node and values greater than the median are put in the new right node, with the median acting as a separation value.
3. The separation value is inserted in the node's parent, which may cause it to be split, and so on. If the node has no parent (i.e., the node was the root), create a new root above this node (increasing the height of the tree).

If the splitting goes all the way up to the root, it creates a new root with a single separator value and two children, which is why the lower bound on the size of internal nodes does not apply to the root. The maximum number of elements per node is  $U-1$ . When a node is split, one element moves to the parent, but one element is added. So, it must be possible to divide the maximum number  $U-1$  of elements into two legal nodes. If this number is odd, then  $U=2L$  and one of the new nodes contains  $(U-2)/2 = L-1$  elements, and hence is a legal node, and the other contains one more element, and hence it is legal too. If  $U-1$  is even, then  $U=2L-1$ , so there are  $2L-2$  elements in the node. Half of this number is  $L-1$ , which is the minimum number of elements allowed per node.

An improved algorithm (Mond & Raz 1985) supports a single pass down the tree from the root to the node where the insertion will take place, splitting any full nodes encountered on the way. This prevents the need to recall the parent nodes into memory, which may be expensive if the nodes are on secondary storage. However, to use this improved algorithm, we must be able to send one element to the parent and split the remaining  $U-2$  elements into two legal nodes, without adding a new element. This requires  $U = 2L$  rather than  $U = 2L-1$ , which accounts for why some textbooks impose this requirement in defining B-trees.



## Deletion

There are two popular strategies for deletion from a B-Tree.

1. Locate and delete the item, then restructure the tree to regain its invariants, **OR**
2. Do a single pass down the tree, but before entering (visiting) a node, restructure the tree so that once the key to be deleted is encountered, it can be deleted without triggering the need for any further restructuring

The algorithm below uses the former strategy.

There are two special cases to consider when deleting an element:

1. The element in an internal node is a separator for its child nodes
2. Deleting an element may put its node under the minimum number of elements and children

The procedures for these cases are in order below.

### Deletion from a leaf node

1. Search for the value to delete
2. If the value's in a leaf node, simply delete it from the node
3. If underflow happens, check siblings, and either transfer a key or fuse the siblings together
4. If deletion happened from right child, retrieve the max value of left child if it has no underflow
5. In vice-versa situation, retrieve the min element from right

### Deletion from an internal node

Each element in an internal node acts as a separation value for two subtrees, and when such an element is deleted, two cases arise.

In the first case, both of the two child nodes to the left and right of the deleted element have the minimum number of elements, namely  $L-1$ . They can then be joined into a single node with  $2L-2$  elements, a number which does not exceed  $U-1$  and so is a legal node. Unless it is known that this particular B-tree does not contain duplicate data, we must then also (recursively) delete the element in question from the new node.

In the second case, one of the two child nodes contains more than the minimum number of elements. Then a new separator for those subtrees must be found. Note that the largest element in the left subtree is still less than the separator. Likewise, the smallest element in the right subtree is the smallest element which is still greater than the separator. Both of those elements are in leaf nodes, and either can be the new separator for the two subtrees.

1. If the value is in an internal node, choose a new separator (either the largest element in the left subtree or the smallest element in the right subtree), remove it from the leaf node it is in, and replace the element to be deleted with the new separator
2. This has deleted an element from a leaf node, and so is now equivalent to the previous case

### Rebalancing after deletion

If deleting an element from a leaf node has brought it under the minimum size, some elements must be redistributed to bring all nodes up to the minimum. In some cases the rearrangement will move the deficiency to the parent, and the redistribution must be applied iteratively up the tree, perhaps even to the root. Since the minimum element count doesn't apply to the root, making the root be the only deficient node is not a problem. The algorithm to rebalance the tree is as follows:

1. If the right sibling has more than the minimum number of elements
  1. Add the separator to the end of the deficient node
  2. Replace the separator in the parent with the first element of the right sibling
  3. Append the first child of the right sibling as the last child of the deficient node
2. Otherwise, if the left sibling has more than the minimum number of elements

1. Add the separator to the start of the deficient node
2. Replace the separator in the parent with the last element of the left sibling
3. Insert the last child of the left sibling as the first child of the deficient node
3. If both immediate siblings have only the minimum number of elements
  1. Create a new node with all the elements from the deficient node, all the elements from one of its siblings, and the separator in the parent between the two combined sibling nodes
  2. Remove the separator from the parent, and replace the two children it separated with the combined node
  3. If that brings the number of elements in the parent under the minimum, repeat these steps with that deficient node, unless it is the root, since the root is permitted to be deficient

The only other case to account for is when the root has no elements and one child. In this case it is sufficient to replace it with its only child.

## Initial construction

In applications, it is frequently useful to build a B-tree to represent a large existing collection of data and then update it incrementally using standard B-tree operations. In this case, the most efficient way to construct the initial B-tree is not to insert every element in the initial collection successively, but instead to construct the initial set of leaf nodes directly from the input, then build the internal nodes from these. This approach to B-tree construction is called bulkloading. Initially, every leaf but the last one has one extra element, which will be used to build the internal nodes.

For example, if the leaf nodes have maximum size 4 and the initial collection is the integers 1 through 24, we would initially construct 4 leaf nodes containing 5 values each and 1 which contains 4 values:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

We build the next level up from the leaves by taking the last element from each leaf node except the last one. Again, each node except the last will contain one extra value. In the example, suppose the internal nodes contain at most 2 values (3 child pointers). Then the next level up of internal nodes would be:

5	10	15	20
---	----	----	----

1	2	3	4	6	7	8	9	11	12	13	14	16	17	18	19	21	22	23	24
---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----

This process is continued until we reach a level with only one node and it is not overfilled. In the example only the root level remains:

15
----

5	10	20
---	----	----

1	2	3	4	6	7	8	9	11	12	13	14	16	17	18	19	21	22	23	24
---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----

## In filesystems

In addition to its use in databases, the B-tree is also used in filesystems to allow quick random access to an arbitrary block in a particular file. The basic problem is turning the file block  $i$  address into a disk block (or perhaps to a cylinder-head-sector) address.

Some operating systems require the user to allocate the maximum size of the file when the file is created. The file can then be allocated as contiguous disk blocks. Converting to a disk block: the operating system just adds the file block address to the starting disk block of the file. The scheme is simple, but the file cannot exceed its created size.

Other operating systems allow a file to grow. The resulting disk blocks may not be contiguous, so mapping logical blocks to physical blocks is more involved.

MS-DOS, for example, used a simple File Allocation Table (FAT). The FAT has an entry for each disk block,<sup>[6]</sup> and that entry identifies whether its block is used by a file and if so, which block (if any) is the next disk block of the same file. So, the allocation of each file is represented as a linked list in the table. In order to find the disk address of file block  $i$ , the operating system (or disk utility) must sequentially follow the file's linked list in the FAT. Worse, to find a free disk block, it must sequentially scan the FAT. For MS-DOS, that was not a huge penalty because the disks and files were small and the FAT had few entries and relatively short file chains. In the FAT12 filesystem (used on floppy disks and early hard disks), there were no more than 4,080<sup>[7]</sup> entries, and the FAT would usually be resident in memory. As disks got bigger, the FAT architecture began to confront penalties. On a large disk using FAT, it may be necessary to perform disk reads to learn the disk location of a file block to be read or written.

TOPS-20 (and possibly TENEX) used a 0 to 2 level tree that has similarities to a B-Tree. A disk block was 512 36-bit words. If the file fit in a 512 ( $2^9$ ) word block, then the file directory would point to that physical disk block. If the file fit in  $2^{18}$  words, then the directory would point to an aux index; the 512 words of that index would either be NULL (the block isn't allocated) or point to the physical address of the block. If the file fit in  $2^{27}$  words, then the directory would point to a block holding an aux-aux index; each entry would either be NULL or point to an aux index. Consequently, the physical disk block for a  $2^{27}$  word file could be located in two disk reads and read on the third.

Apple's filesystem HFS+, Microsoft's NTFS,<sup>[8]</sup> AIX (jfs2) and some Linux filesystems, such as btrfs and Ext4, use B-trees.

B\*-trees are used in the HFS and Reiser4 file systems.

## Variations

### Access concurrency

Lehman and Yao<sup>[9]</sup> showed that all read locks could be avoided (and thus concurrent access greatly improved) by linking the tree blocks at each level together with a "next" pointer. This results in a tree structure where both insertion and search operations descend from the root to the leaf. Write locks are only required as a tree block is modified. This maximizes access concurrency by multiple users, an important consideration for databases and/or other B-Tree based ISAM storage methods. The cost associated with this improvement is that empty pages cannot be removed from the btree during normal operations. (However, see <sup>[10]</sup> for various strategies to implement node merging, and source code at.<sup>[11]</sup>)

United States Patent 5283894, granted In 1994, appears to show a way to use a 'Meta Access Method' <sup>[12]</sup> to allow concurrent B+Tree access and modification without locks. The technique accesses the tree 'upwards' for both searches and updates by means of additional in-memory indexes that point at the blocks in each level in the block cache. No reorganization for deletes is needed and there are no 'next' pointers in each block as in Lehman and Yao.

## Notes

- [1] Counted B-Trees (<http://www.chiark.greenend.org.uk/~sgtatham/algorithms/cbtree.html>), retrieved 2010-01-25
- [2] Knuth's video lectures from Stanford (<http://scpd.stanford.edu/knuth/index.jsp>)
- [3] Seagate Technology LLC, Product Manual: Barracuda ES.2 Serial ATA, Rev. F., publication 100468393, 2008 ([http://www.seagate.com/staticfiles/support/disc/manuals/NL35 Series &BC ES Series/Barracuda ES.2 Series/100468393f.pdf](http://www.seagate.com/staticfiles/support/disc/manuals/NL35%20Series%20&BC%20ES%20Series/Barracuda%20ES.2%20Series/100468393f.pdf)), page 6
- [4] Bayer & McCreight (1972) avoided the issue by saying an index element is a (physically adjacent) pair of  $(x, a)$  where  $x$  is the key, and  $a$  is some associated information. The associated information might be a pointer to a record or records in a random access, but what it was didn't really matter. Bayer & McCreight (1972) states, "For this paper the associated information is of no further interest."
- [5] If  $n$  is zero, then no root node is needed, so the height of an empty tree is not well defined.
- [6] For FAT, what is called a "disk block" here is what the FAT documentation calls a "cluster", which is fixed-size group of one or more contiguous whole physical disk sectors. For the purposes of this discussion, a cluster has no significant difference from a physical sector.
- [7] Two of these were reserved for special purposes, so only 4078 could actually represent disk blocks (clusters).
- [8] Mark Russinovich. "Inside Win2K NTFS, Part 1" (<http://msdn2.microsoft.com/en-us/library/ms995846.aspx>). Microsoft Developer Network. Archived (<http://web.archive.org/web/20080413181940/http://msdn2.microsoft.com/en-us/library/ms995846.aspx>) from the original on 13 April 2008. . Retrieved 2008-04-18.
- [9] "Efficient locking for concurrent operations on B-trees" (<http://portal.acm.org/citation.cfm?id=319663&dl=GUIDE&coll=GUIDE&CFID=61777986&CFTOKEN=74351190>). Portal.acm.org. doi:10.1145/319628.319663. . Retrieved 2012-06-28.
- [10] <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA232287&Location=U2&doc=GetTRDoc.pdf>
- [11] "Downloads - high-concurrency-btree - High Concurrency B-Tree code in C - Google Project Hosting" (<http://code.google.com/p/high-concurrency-btree/downloads/list>). Code.google.com. . Retrieved 2012-06-28.
- [12] Lockless Concurrent B+Tree (<http://www.freepatentsonline.com/5283894.html>)

## References

- Bayer, R.; McCreight, E. (1972), "Organization and Maintenance of Large Ordered Indexes" ([http://www.minet.uni-jena.de/dbis/lehre/ws2005/dbs1/Bayer\\_hist.pdf](http://www.minet.uni-jena.de/dbis/lehre/ws2005/dbs1/Bayer_hist.pdf)), *Acta Informatica* **1** (3): 173–189
- Comer, Douglas (June 1979), "The Ubiquitous B-Tree", *Computing Surveys* **11** (2): 123–137, doi:10.1145/356770.356776, ISSN 0360-0300.
- Cormen, Thomas; Leiserson, Charles; Rivest, Ronald; Stein, Clifford (2001), *Introduction to Algorithms* (Second ed.), MIT Press and McGraw-Hill, pp. 434–454, ISBN 0-262-03293-7. Chapter 18: B-Trees.
- Folk, Michael J.; Zoellick, Bill (1992), *File Structures* (2nd ed.), Addison-Wesley, ISBN 0-201-55713-4
- Knuth, Donald (1998), *Sorting and Searching*, The Art of Computer Programming, **Volume 3** (Second ed.), Addison-Wesley, ISBN 0-201-89685-0. Section 6.2.4: Multiway Trees, pp. 481–491. Also, pp. 476–477 of section 6.2.3 (Balanced Trees) discusses 2-3 trees.
- Mond, Yehudit; Raz, Yoav (1985), "Concurrency Control in B+-Trees Databases Using Preparatory Operations" (<http://www.informatik.uni-trier.de/~ley/db/conf/vldb/MondR85.html>), *VLDB'85, Proceedings of 11th International Conference on Very Large Data Bases*: 331–334.

## Original papers

- Bayer, Rudolf; McCreight, E. (July 1970), *Organization and Maintenance of Large Ordered Indices*, **Mathematical and Information Sciences Report No. 20**, Boeing Scientific Research Laboratories.
- Bayer, Rudolf (1971), "Binary B-Trees for Virtual Memory", Proceedings of 1971 ACM-SIGFIDET Workshop on Data Description, Access and Control, San Diego, California. November 11–12, 1971.

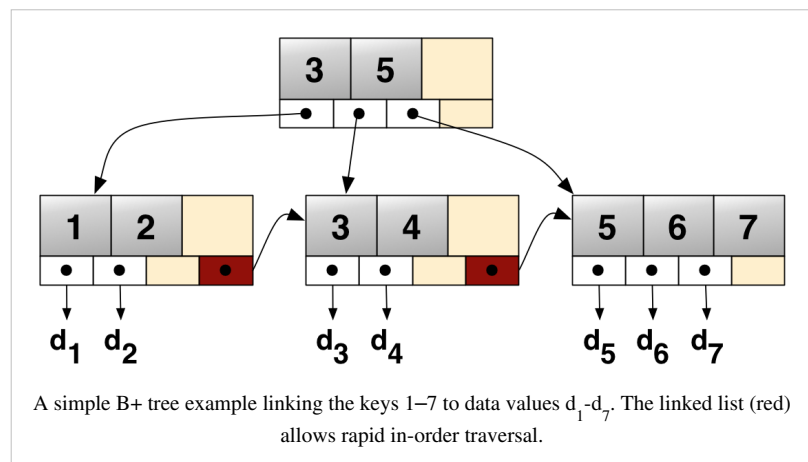
## External links

- B-Tree animation applet (<http://slady.net/java/bt/view.php>) by slady
- B-tree and UB-tree on Scholarpedia ([http://www.scholarpedia.org/article/B-tree\\_and\\_UB-tree](http://www.scholarpedia.org/article/B-tree_and_UB-tree)) Curator: Dr Rudolf Bayer
- B-Trees: Balanced Tree Data Structures (<http://www.bluerwhite.org/btree>)
- NIST's Dictionary of Algorithms and Data Structures: B-tree (<http://www.nist.gov/dads/HTML/btree.html>)

- B-Tree Tutorial (<http://cis.stvincent.edu/html/tutorials/swd/btree/btree.html>)
- The InfinityDB BTree implementation (<http://www.boilerbay.com/infinitydb/TheDesignOfTheInfinityDatabaseEngine.htm>)
- Cache Oblivious B(+)-trees (<http://supertech.csail.mit.edu/cacheObliviousBTree.html>)
- Dictionary of Algorithms and Data Structures entry for B\*-tree (<http://www.nist.gov/dads/HTML/bstartree.html>)
- Open Data Structures - Section 14.2 - B-Trees ([http://opendatastructures.org/versions/edition-0.1e/ods-java/14\\_2\\_B\\_Trees.html](http://opendatastructures.org/versions/edition-0.1e/ods-java/14_2_B_Trees.html))

## B+ tree

In computer science, a **B+ tree** is a type of tree which represents sorted data in a way that allows for efficient insertion, retrieval and removal of records, each of which is identified by a *key*. It is a dynamic, multilevel index, with maximum and minimum bounds on the number of keys in each index segment (usually called a "block" or "node"). In a B+ tree, in contrast to a B-tree, all records are stored at the leaf level of the tree; only keys are stored in interior nodes.



The primary value of a B+ tree is in storing data for efficient retrieval in a block-oriented storage context—in particular, filesystems. This is primarily because unlike binary search trees, B+ trees have very high fanout (typically on the order of 100 or more), which reduces the number of I/O operations required to find an element in the tree.

The NTFS, ReiserFS, NSS, XFS, JFS, and ReFS filesystems all use this type of tree for metadata indexing. Relational database management systems such as IBM DB2,<sup>[1]</sup> Informix,<sup>[1]</sup> Microsoft SQL Server,<sup>[1]</sup> Oracle 8,<sup>[1]</sup> Sybase ASE,<sup>[1]</sup> and SQLite<sup>[2]</sup> support this type of tree for table indices. Key-value database management systems such as CouchDB,<sup>[3]</sup> Tokyo Cabinet<sup>[4]</sup> support this type of tree for data access.

## Overview

The order, or branching factor  $b$  of a B+ tree measures the capacity of nodes (i.e. the number of children nodes) for internal nodes in the tree. The actual number of children for a node, referred to here as  $m$ , is constrained for internal nodes so that  $\lceil b/2 \rceil \leq m \leq b$ . The root is an exception: it is allowed to have as few as two children. For example, if the order of a B+ tree is 7, each internal node (except for the root) may have between 4 and 7 children; the root may have between 2 and 7. Leaf nodes have no children, but are constrained so that the number of keys must be at least  $\lceil b/2 \rceil$  and at most  $b - 1$ . In the situation where a B+ tree is nearly empty, it only contains one node, which is a leaf node. (The root is also the single leaf, in this case.) This node is permitted to have as little as one key if necessary, and at most  $b$ .

Node Type	Children Type	Min Children	Max Children	Example $b = 7$	Example $b = 100$
Root Node (when it is the only node in the tree)	Records	1	$b$	1 - 7	1 - 100
Root Node	Internal Nodes or Leaf Nodes	2	$b$	2 - 7	2 - 100
Internal Node	Internal Nodes or Leaf Nodes	$\lceil b/2 \rceil$	$b$	4 - 7	50 - 100
Leaf Node	Records	$\lfloor b/2 \rfloor$	$b - 1$	3 - 6	50 - 99

## Algorithms

### Search

The root of a  $B^+$ -Tree represents the whole range of values in the tree, where every internal node a subinterval.

We are looking for a value  $k$  in the  $B^+$ -Tree. Starting from the root, we are looking for the leaf which may contain the value  $k$ . At each node, we figure out which internal pointer we should to follow. An internal  $B^+$ -Tree node has at most  $d \leq b$  children, where every one of them represents a different sub-interval. We select the corresponding node by searching on the key values of the node.

```

Function: search (k)
    return tree_search (k, root);

Function: tree_search (k, node)
    if node is a leaf then
        return node;
    switch k do
        case  $k < k_0$ 
            return tree_search(k, p_0);
        case  $k_i \leq k < k_{i+1}$ 
            return tree_search(k, p_i);
        case  $k_d \leq k$ 
            return tree_search(k, p_d);

```

This pseudocode assumes that no repetition is allowed.

### Insertion

Perform a search to determine what bucket the new record should go into.

- If the bucket is not full (at most  $b - 1$  entries after the insertion), add the record.
- Otherwise, split the bucket.
  - Allocate new leaf and move half the bucket's elements to the new bucket.
  - Insert the new leaf's smallest key and address into the parent.
  - If the parent is full, split it too.
    - Add the middle key to the parent node.
    - Repeat until a parent is found that need not split.
- If the root splits, create a new root which has one key and two pointers.



B-trees grow at the root and not at the leaves.

Note that, for a non-leaf node split, we can simply push up the middle key (17). Contrast this with a leaf node split.

## Deletion

- Start at root, find leaf L where entry belongs.
- Remove the entry.
  - If L is at least half-full, done!
  - If L has fewer entries than it should,
    - Try to re-distribute, borrowing from sibling (adjacent node with same parent as L).
    - If re-distribution fails, merge L and sibling.
- If merge occurred, must delete entry (pointing to L or sibling) from parent of L.
- Merge could propagate to root, decreasing height.

## Bulk-loading

Given a collection of data records, we want to create a B+ tree index on some key field. One approach to insert each record into a empty tree. However, it is quite expensive, because each entry requires us to start from the root and go down to the appropriate leaf page. An efficient alternative is use bulk-loading.

- The first step is to sort the data entries according to a search key,.
- We allocate an empty page to serve as the root, and insert a pointer to the first page of entries into it.
- When the root is full, we split the root, and create a new root page.
- Keep insert entries to the right most index page just above the leaf level, till all entries are indexed.

Note (1) when the right-most index page above the leaf level fills up, it is split; (2) this action may, in turn, cause a split of the right-most index page on step closer to the root; and (3) splits only occur on the right-most path from the root to the leaf level.

## Characteristics

For a  $b$ -order B+ tree with  $h$  levels of index:

- The maximum number of records stored is  $n_{max} = b^h - b^{h-1}$
- The minimum number of records stored is  $n_{min} = 2 \left\lceil \frac{b}{2} \right\rceil^{h-1}$
- The minimum number of keys is  $n_{kmin} = 2 \left\lceil \frac{b}{2} \right\rceil^h - 1$
- The space required to store the tree is  $O(n)$
- Inserting a record requires  $O(\log_b n)$  operations
- Finding a record requires  $O(\log_b n)$  operations
- Removing a (previously located) record requires  $O(\log_b n)$  operations
- Performing a range query with  $k$  elements occurring within the range requires  $O(\log_b n + k)$  operations
- Performing a pagination query with page size  $s$  and page number  $p$  requires  $O(p * s)$  operations

## Implementation

The leaves (the bottom-most index blocks) of the B+ tree are often linked to one another in a linked list; this makes range queries or an (ordered) iteration through the blocks simpler and more efficient (though the aforementioned upper bound can be achieved even without this addition). This does not substantially increase space consumption or maintenance on the tree. This illustrates one of the significant advantages of a B+-tree over a B-tree; in a B-tree, since not all keys are present in the leaves, such an ordered linked list cannot be constructed. A B+ tree is thus particularly useful as a database system index, where the data typically resides on disk, as it allows the B+-tree to actually provide an efficient structure for housing the data itself (this is described in <sup>[5]</sup> as index structure "Alternative 1").

If a storage system has a block size of  $B$  bytes, and the keys to be stored have a size of  $k$ , arguably the most efficient B+ tree is one where  $b = \lceil B/k \rceil - 1$ . Although theoretically the one-off is unnecessary, in practice there is often a little extra space taken up by the index blocks (for example, the linked list references in the leaf blocks). Having an index block which is slightly larger than the storage system's actual block represents a significant performance decrease; therefore erring on the side of caution is preferable.

If nodes of the B+ tree are organized as arrays of elements, then it may take a considerable time to insert or delete an element as half of the array will need to be shifted on average. To overcome this problem, elements inside a node can be organized in a binary tree or a B+ tree instead of an array.

B+ trees can also be used for data stored in RAM. In this case a reasonable choice for block size would be the size of processor's cache line. However, some studies have proved that a block size a few times larger than the processor's cache line can deliver better performance if cache prefetching is used.

Space efficiency of B+ trees can be improved by using some compression techniques. One possibility is to use delta encoding to compress keys stored into each block. For internal blocks, space saving can be achieved by either compressing keys or pointers. For string keys, space can be saved by using the following technique: Normally the  $i$ th entry of an internal block contains the first key of block  $i+1$ . Instead of storing the full key, we could store the shortest prefix of the first key of block  $i+1$  that is strictly greater (in lexicographic order) than last key of block  $i$ . There is also a simple way to compress pointers: if we suppose that some consecutive blocks  $i, i+1 \dots i+k$  are stored contiguously, then it will suffice to store only a pointer to the first block and the count of consecutive blocks.

All the above compression techniques have some drawbacks. First, a full block must be decompressed to extract a single element. One technique to overcome this problem is to divide each block into sub-blocks and compress them separately. In this case searching or inserting an element will only need to decompress or compress a sub-block instead of a full block. Another drawback of compression techniques is that the number of stored elements may vary considerably from a block to another depending on how well the elements are compressed inside each block.

## History

The B tree was first described in the paper *Organization and Maintenance of Large Ordered Indices*. *Acta Informatica* 1: 173–189 (1972) by Rudolf Bayer and Edward M. McCreight. There is no single paper introducing the B+ tree concept. Instead, the notion of maintaining all data in leaf nodes is repeatedly brought up as an interesting variant. An early survey of B trees also covering B+ trees is Douglas Comer: "The Ubiquitous B-Tree <sup>[6]</sup>", *ACM Computing Surveys* 11(2): 121–137 (1979). Comer notes that the B+ tree was used in IBM's VSAM data access software and he refers to an IBM published article from 1973.

## References

- [1] Ramakrishnan Raghu, Gehrke Johannes - Database Management Systems, McGraw-Hill Higher Education (2000), 2nd edition (en) page 267
- [2] SQLite Version 3 Overview (<http://sqlite.org/version3.html>)
- [3] CouchDB Guide (see note after 3rd paragraph) (<http://guide.couchdb.org/draft/btree.html>)
- [4] Tokyo Cabinet reference (<http://1978th.net/tokyocabinet/>)
- [5] Ramakrishnan, R. and Gehrke, J. Database Management Systems, McGraw-Hill Higher Education (2002), 3rd edition
- [6] <http://doi.acm.org/10.1145/356770.356776>

## External links

- B+ tree in Python, used to implement a list (<http://pypi.python.org/pypi/blist>)
- Dr. Monge's B+ Tree index notes (<http://www.cecs.csulb.edu/~monge/classes/share/B+TreeIndexes.html>)
- Evaluating the performance of CSB+-trees on Mutithreaded Architectures (<http://blogs.ubc.ca/lrashid/files/2011/01/CCECE07.pdf>)
- Effect of node size on the performance of cache conscious B+-trees (<http://www.eecs.umich.edu/~jignesh/quickstep/publ/cci.pdf>)
- Fractal Prefetching B+-trees (<http://www.pittsburgh.intel-research.net/people/gibbons/papers/fpbptrees.pdf>)
- Towards pB+-trees in the field: implementations Choices and performance (<http://leo.saclay.inria.fr/events/EXPDB2006/PAPERS/Jonsson.pdf>)
- Cache-Conscious Index Structures for Main-Memory Databases (<https://oa.doria.fi/bitstream/handle/10024/2906/cachecon.pdf?sequence=1>)
- Cache Oblivious B(+)-trees (<http://supertech.csail.mit.edu/cacheObliviousBTree.html>)
- The Power of B-Trees: CouchDB B+ Tree Implementation (<http://books.couchdb.org/relax/appendix/btrees>)

## Implementations

- Interactive B+ Tree Implementation in C (<http://www.amittai.com/prose/bplustree.html>)
  - Memory based B+ tree implementation as C++ template library (<http://idlebox.net/2007/stx-btree/>)
  - Stream based B+ tree implementation as C++ template library (<http://gitorious.org/bp-tree/main>)
  - Open Source C++ B+ Tree Implementation ([http://www.scalingweb.com/bplus\\_tree.php](http://www.scalingweb.com/bplus_tree.php))
  - Open Source Javascript B+ Tree Implementation (<http://blog.conquex.com/?p=84>)
  - Perl implementation of B+ trees (<http://search.cpan.org/~hanenkamp/Tree-BPTree-1.07>)
  - Java/C#/Python implementations of B+ trees (<http://bplusdotnet.sourceforge.net>)
  - File based B+Tree in C# with threading and MVCC support ([http://csharptest.net/?page\\_id=563](http://csharptest.net/?page_id=563))
  - Javascript B+ Tree, MIT License (<http://prosehack.wordpress.com/2012/05/25/a-javascript-b-tree/>)
-

# Dancing tree

---

In computer science, a **dancing tree** is a tree data structure similar to B+ trees. It was invented by Hans Reiser, for use by the Reiser4 file system. As opposed to self-balancing binary search trees that attempt to keep their nodes balanced at all times, dancing trees only balance their nodes when flushing data to a disk (either because of memory constraints or because a transaction has completed).<sup>[1]</sup>

The idea behind this is to speed up file system operations by delaying optimization of the tree and only writing to disk when necessary, as writing to disk is thousands of times slower than writing to memory. Also, because this optimization is done less often than with other tree data structures, the optimization can be more extensive.

In some sense, this can be considered to be a self-balancing binary search tree that is optimized for storage on a slow medium, in that the on-disc form will always be balanced but will get no mid-transaction writes; doing so eases the difficulty (at the time) of adding and removing nodes, and instead performs these (slow) rebalancing operations at the same time as the (much slower) write to the storage medium.

However, a (negative) side effect of this behavior is witnessed in cases of unexpected shutdown, incomplete data writes, and other occurrences that may prevent the final (balanced) transaction from completing. In general, dancing trees will pose a greater difficulty for data recovery from incomplete transactions than a normal tree; though this can be addressed by either adding extra transaction logs or developing an algorithm to locate data on disk not previously present, then going through with the optimizations once more before continuing with any other pending operations/transactions.

## References

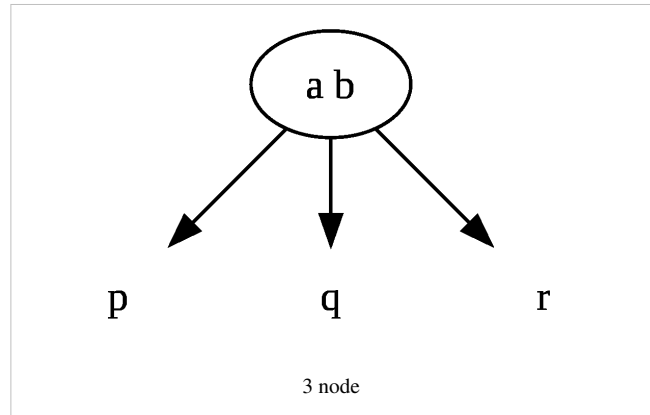
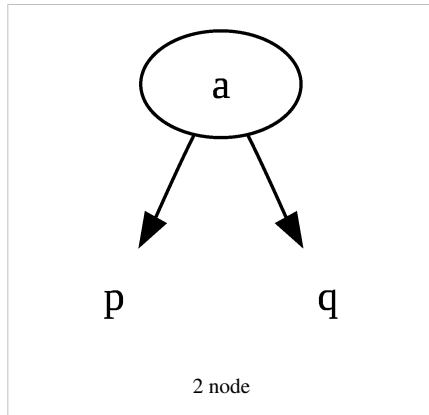
- [1] Hans Reiser. "Reiser4 release notes - Dancing Tree" ([http://web.archive.org/web/20071024001500/http://www.namesys.com/v4/v4.html#dancing\\_tree](http://web.archive.org/web/20071024001500/http://www.namesys.com/v4/v4.html#dancing_tree)). Archive.org, as Namesys.com is no longer accessible. Archived from the original ([http://www.namesys.com/v4/v4.html#dancing\\_tree](http://www.namesys.com/v4/v4.html#dancing_tree)) on 2007-10-24. . Retrieved 2009-07-22.

## External links

- Software Engineering Based Reiser4 Design Principles* ([http://www.namesys.com/v4/v4.html#dancing\\_tree](http://www.namesys.com/v4/v4.html#dancing_tree))
- Description of the Reiser4 internal tree (<http://nikitadanilov.blogspot.com/2006/03/reiser4-1-internal-tree.html>)

## 2-3 tree

In computer science, a **2-3 tree** is a type of data structure, a tree where every node with children (internal node) has either two children (2-node) and one data element or three children (3-nodes) and two data elements. Nodes on the outside of the tree (leaf nodes) have no children and one or two data elements.<sup>[1]</sup>



2-3 trees are an isometry of AA trees, meaning that they are equivalent data structures. In other words, for every 2-3 tree, there exists at least one AA tree with data elements in the same order. 2-3 trees are balanced, meaning that each right, center, and left subtree contains the same or close to the same amount of data.

### Properties

- Every non-leaf is a 2-node or a 3-node. A 2-node contains one data item and has two children. A 3-node contains two data items and has 3 children.
- All leaves are at the same level (the bottom level)
- All data are kept in sorted order
- Every leaf node will contain 1 or 2 fields.

### Non-leaf nodes

These contain one or two fields which indicate the range of values in its subtrees. If a node has two children, it will have one field; if the node has three children, it will have two fields. Each non-leaf node will contain a value in field 1 which is greater than the largest item in its left sub-tree, but less than or equal to the smallest item in its right sub-tree (or center sub-tree, if it has three children). If that node has three children, field 2 contains a value which is greater than the largest value in the center sub-tree, but less than or equal to the smallest item in its right sub-tree. The purpose of these values is to direct a search function to the correct sub-tree and eventually to the correct data node.

## References

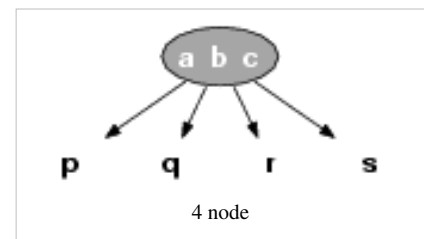
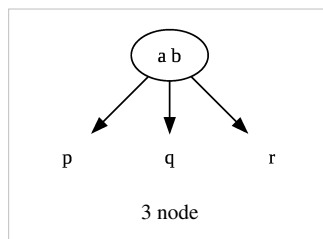
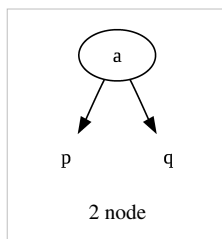
[1] Gross, R. Hernández, J. C. Lázaro, R. Dormido, S. Ros (2001). *Estructura de Datos y Algoritmos*. Prentice Hall. ISBN 84-205-2980-X

## External links

- 2-3 Trees Complete Description ([http://www.cs.ucr.edu/cs14/cs14\\_06win/slides/2-3\\_trees\\_covered.pdf](http://www.cs.ucr.edu/cs14/cs14_06win/slides/2-3_trees_covered.pdf))
- 2-3 Tree Java Applet (<http://www.cosc.canterbury.ac.nz/mukundan/dsal/TwoThreeTree.html>)
- 2-3 Tree In-depth description (<http://www.aihorizon.com/essays/basiccs/trees/twothree.htm>)
- 2-3 Tree in F# (<http://v2matveev.blogspot.com/2010/03/data-structures-2-3-tree.html>)
- 2-3 Tree in Python (<http://code.google.com/p/risboo6909/source/browse/trunk/23tree/tttree.py>)

## 2-3-4 tree

In computer science, a **2-3-4 tree** (also called a **2-4 tree**) is a self-balancing data structure that is commonly used to implement dictionaries. The numbers mean a tree where every node with children (internal node) has either two children (2-node) and one data element or three children (3-node) and two data elements or four children (4-node) and three data elements.



2-3-4 trees are B-trees of order 4; like B-trees in general, they can search, insert and delete in  $O(\log n)$  time. One property of a 2-3-4 tree is that all external nodes are at the same depth.

2-3-4 trees are an isometry of red-black trees, meaning that they are equivalent data structures. In other words, for every 2-3-4 tree, there exists at least one red-black tree with data elements in the same order. Moreover, insertion and deletion operations on 2-3-4 trees that cause node expansions, splits and merges are equivalent to the color-flipping and rotations in red-black trees. Introductions to red-black trees usually introduce 2-3-4 trees first, because they are conceptually simpler. 2-3-4 trees, however, can be difficult to implement in most programming languages because of the large number of special cases involved in operations on the tree. Red-black trees are simpler to implement, so tend to be used instead.

## Properties

- Every non-leaf is a 2-node, 3-node or a 4-node. A 2-node contains one data item and has two children. A 3-node contains two data items and has 3 children. A 4-node contains 3 data items and has 4 children.
- All leaves are at the same level (the bottom level)
- All data are kept in sorted order
- Every non-leaf node will contain 1, 2 or 3 fields.

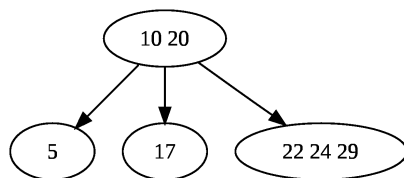
## Insertion

To insert a value, we start at the root of the 2-3-4 tree:

1. If the current node is a 4-node:
  - Remove and save the middle value to get a 3-node.
  - Split the remaining 3-node up into a pair of 2-nodes (the now missing middle value is handled in the next step).
  - If this is the root node (which thus has no parent):
    - the middle value becomes the new root 2-node and the tree height increases by 1. Ascend into the root.
    - Otherwise, push the middle value up into the parent node. Ascend into the parent node.
2. Find the child whose interval contains the value to be inserted.
3. If that child is a leaf, insert the value into current node and finish.
  - Otherwise, descend into the child and repeat from step 1.<sup>[1][2]</sup>

## Example

To insert the value "25" into this 2-3-4 tree:



- Begin at the root (10, 20) and descend towards the rightmost child (22, 24, 29). (Its interval (20,  $\infty$ ) contains 25.)
- Node (22, 24, 29) is a 4-node, so its middle element 24 is pushed up into the parent node.
- The remaining 3-node (22, 29) is split into a pair of 2-nodes (22) and (29). Ascend back into the new parent (10, 20, 24).
- Descend towards the rightmost child (29). (Its interval (24,  $\infty$ ) contains 25.)
- Node (29) has no leftmost child. (The child for interval ( $\infty$ , 29) is empty.) Stop here and insert value 25 into this node.

## Deletion

Consider just leaving the element there, marking it "deleted," possibly to be re-used for a future insertion.

Find the element to be deleted. If the element is not in a leaf node remember its location and continue searching until a leaf, which will contain the element's successor, is reached. Then swap the leaf element with the one to be deleted, and delete the element node. It is simplest to make adjustments to the tree from the top down, as the element to be deleted is pursued that guarantee that the leaf node found is not a two-node, so that we can delete something from it and leave it there.

The adjustments we make on the way to a leaf are as follows: Assume, without loss of generality, that the child we are about to go to is the leftmost.

If we're at the root

If the root and both children are two-nodes, combine all three elements into the root, making a 4-node and shortening the tree, Otherwise, if the root and left child are two-nodes, the right child isn't a two-node. Perform a left rotation to make the left sibling a 3-node, and move to the left child.

From now on, we can be sure that we're at a node which is not a 2-node.

If the leftmost child is not a 2-node, just move to it.  
If the adjacent sibling is not a 2-node, perform a left rotation using its leftmost element to make the left child a 3-node.

Otherwise, add the leftmost element of the parent and the single element of the sibling to the left node, making it a 4-node, and discard the empty sibling.

Go to the left-most child.

Deletion in a 2-3-4 tree is  $O(\log n)$ , assuming transfer and fusion run in constant time ( $O(1)$ ).<sup>[1][3]</sup>

## References

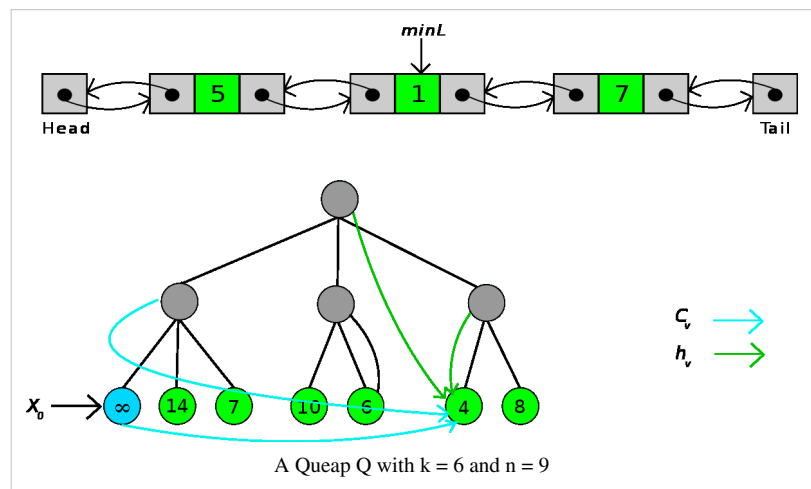
- [1] Ford, William; Topp, William (2002), *Data Structures with C++ Using STL* (2nd ed.), New Jersey: Prentice Hall, pp. 683, ISBN 0-13-085850-1
- [2] Goodrich, Michael T; Tamassia, Roberto; Mount, David M (2002), *Data Structures and Algorithms in C++*, Wiley, ISBN 0-471-20208-8
- [3] Grama, Ananth (2004). "(2,4) Trees" (<http://www.cs.purdue.edu/homes/ayg/CS251/slides/chap13a.pdf>). *CS251: Data Structures Lecture Notes*. Department of Computer Science, Purdue University. . Retrieved 2008-04-10.

## External links

- Animation of a 2-3-4 Tree (<http://www.cse.ohio-state.edu/~bondhugu/acads/234-tree/index.shtml>)
- Java Applet showing a 2-3-4 Tree (<http://www.cs.unm.edu/~rlpm/499/ttft.html>)
- Left-leaning Red-Black Trees - Princeton University, 2008 (<http://www.cs.princeton.edu/~rs/talks/LLRB/LLRB.pdf>)
- Open Data Structures - Section 9.1 - 2-4 Trees ([http://opendatastructures.org/versions/edition-0.1e/ods-java/9\\_1\\_2\\_4\\_Trees.html](http://opendatastructures.org/versions/edition-0.1e/ods-java/9_1_2_4_Trees.html))

## Queaps

In computer science, a **queap** is a priority queue data structure that points to the smallest stored item. Queap is composed of a doubly linked list and a 2-4 tree data structure. The data structure satisfies the queueish property, a complement of the working set property, which makes the search operations of some element  $x$  to run in  $O(\lg q(x))$  amortized time where  $q(x)$  is the number of items that has been in the priority queue longer than  $x$ .



The doubly linked list keeps a list of  $k$  inserted elements. When a deletion operation occurs, the  $k$  items are added to the 2-4 tree. The item is then deleted from the tree. Each structure points to the minimum one. The 2-4 tree has been modified for this structure in order to access the smallest element in constant time. The 'queap' was coined by John Iacono and Stefan Langerman<sup>[1]</sup>



## Description

**Queap** is a priority queue that inserts elements in  $O(1)$  amortized time, and removes the minimum element in  $O(\log(k + 2))$  if there are  $k$  items in the heap longer than the element to be extracted. The queap has a property called the queueish property: the time to search for element  $x$  is  $O(\lg q(x))$  where  $q(x)$  is equal to  $n - 1 - w(x)$  and  $w(x)$  is the number of distinct items that has been accessed by operations such as searching, inserting, or deleting.  $q(x)$  is defined as how many elements have not been accessed since  $x$ 's last access. Indeed, the queueish property is the complement of the splay tree working set property: the time to search for element  $x$  is  $O(\lg w(x))$ .

Queap can be represented by two data structures: a doubly linked list and a modified version of 2-4 tree. The doubly linked list,  $L$ , is used for a series of insert and locate-min operations. The queap keeps a pointer to the minimum element stored in the list. To add element  $x$  to list  $L$ , the element  $x$  is added to the end of the list and a bit variable in element  $x$  is set to one. This operation is done to determine if the element is either in the list or in a 2-4 tree.

A 2-4 tree is used when a delete operation occurs. If the item  $x$  is already in tree  $T$ , the item is removed using the 2-4 tree delete operation. Otherwise, the item  $x$  is in list  $L$  (done by checking if the bit variable is set). All the elements stored in list  $L$  are then added to the 2-4 tree, setting the bit variable of each element to zero.  $x$  is then removed from  $T$ .

Queap uses only the 2-4 tree structure properties, not a search tree. The modified 2-4 tree structure is as follows. Suppose list  $L$  has the following set of elements:  $x_1, x_2, x_3, \dots, x_k$ . When the deletion operation is invoked, the set of elements stored in  $L$  is then added to the leaves of the 2-4 tree in that order, preceded by a dummy leaf containing an infinite key. Each internal node of  $T$  has a pointer  $h_v$ , which points to the smallest item in subtree  $v$ . Each internal node on path  $P$  from the root to  $x_0$  has a pointer  $c_v$ , which points to the smallest key in  $T - T_v - \{r\}$ . The  $h_v$  pointers of each internal node on path  $P$  are ignored. The queap has a pointer to  $c_{x_0}$ , which points to the smallest element in  $T$ .

An application of queap includes a unique set of high priority events and extraction of the highest priority event for processing.

## Operations

Let  $minL$  be a pointer that points to the minimum element in the doubly linked list  $L$ ,  $c_{x_0}$  be the minimum element stored in the 2-4 tree,  $T$ ,  $k$  be the number of elements stored in  $T$ , and  $n$  be the total number of elements stored in queap  $Q$ . The operations are as follows:

**New( $Q$ ):** Initializes a new empty queap.

Initialize an empty doubly linked list  $L$  and 2-4 tree  $T$ . Set  $k$  and  $n$  to zero.

**Insert( $Q, x$ ):** Add the element  $x$  to queap  $Q$ .

Insert the element  $x$  in list  $L$ . Set the bit in element  $x$  to one to demonstrate that the element is in the list  $L$ .

Update the  $minL$  pointer if  $x$  is the smallest element in the list. Increment  $n$  by 1.

**Minimum( $Q$ ):** Retrieve a pointer to the smallest element from queap  $Q$ .

If  $key(minL) < key(c_{x_0})$ , return  $minL$ . Otherwise return  $c_{x_0}$ .

**Delete( $Q, x$ ):** Remove element  $x$  from queap  $Q$ .

If the bit of the element  $x$  is set to one, the element is stored in list  $L$ . Add all the elements from  $L$  to  $T$ , setting the bit of each element to zero. Each element is added to the parent of the right most child of  $T$  using the insert operation of the 2-4 tree.  $L$  becomes empty. Update  $h_v$  pointers for all the nodes  $v$  whose children are new/modified, and repeat the process with the next parent until the parent is equal to the root. Walk from the root to node  $x_0$ , and update the  $c_v$  values. Set  $k$  equal to  $n$ .

If the bit of the element  $x$  is set to zero,  $x$  is a leaf of  $T$ . Delete  $x$  using the 2-4 tree delete operation. Starting from node  $x$ , walk in  $T$  to node  $x_0$ , updating  $h_v$  and  $c_v$  pointers. Decrement  $n$  and  $k$  by 1.

**DeleteMin(Q):** Delete and return the smallest element from queap  $Q$ .

Invoke the *Minimum(Q)* operation. The operation returns *min*. Invoke the *Delete(Q, min)* operation. Return *min*.

**CleanUp(Q):** Delete all the elements in list  $L$  and tree  $T$ .

Starting from the first element in list  $L$ , traverse the list, deleting each node.

Starting from the root of the tree  $T$ , traverse the tree using the post-order traversal algorithm, deleting each node in the tree.

## Analysis

The running time is analyzed using the Amortized Analysis tool. The potential function for queap  $Q$  will be  $\phi(Q) = c|L|$  where  $Q = (T, L)$ .

**Insert(Q, x):** The cost of the operation is  $O(1)$ . The size of list  $L$  grows by one, the potential increases by some constant  $c$ .

**Minimum(Q):** The operation does not alter the data structure so the amortized cost is equal to its actual cost,  $O(1)$ .

**Delete(Q, x):** There are two cases.

### Case 1

If  $x$  is in tree  $T$ , then the amortized cost is not modified. The delete operation is  $O(1)$  amortized 2-4 tree. Since  $x$  was removed from the tree,  $h_v$  and  $c_v$  pointers may need updating. At most, there will be  $O(\lg q(x))$  updates.

### Case 2

If  $x$  is in list  $L$ , then all the elements from  $L$  are inserted in  $T$ . This has a cost of  $a|L|$  of some constant  $a$ , amortized over the 2-4 tree. After inserting and updating the  $h_v$  and  $c_v$  pointers, the total time spent is bounded by  $2a|L|$ .

The second operation is to delete  $x$  from  $T$ , and to walk on the path from  $x$  to  $x_0$ , correcting  $h_v$  and  $c_v$  values. The time is spent at most  $2a|L| + O(\lg q(x))$ . If  $c > 2a$ , then the amortized cost will be  $O(\lg q(x))$ . **Delete(Q, x):** is the addition of the amortized cost of **Minimum(Q)** and **Delete(Q, x)**, which is  $O(\lg q(x))$ .

## Code example

A small java implementation of a queap:

```
public class Queap
{
    public int n, k;
    public List<Element> l; //Element is a generic data type
    public QueapTree t;    //a 2-4 tree, modified for Queap purpose
    public Element minL;

    private Queap() {
        n = 0;
        k = 0;
        l = new LinkedList<Element>();
        t = new QueapTree();
    }

    public static Queap New() {
```

```
        return new Queap();
    }

    public static void Insert(Queap Q, Element x) {
        if (Q.n == 0)
            Q.minL = x;
        Q.l.add(x);
        x.inList = true;
        if (x.compareTo(Q.minL) < 0)
            Q.minL = x;
    }

    public static Element Minimum(Queap Q) {
        //t is a 2-4 tree and x0, cv are tree nodes.
        if (Q.minL.compareTo(Q.t.x0.cv.key) < 0)
            return Q.minL;

        return Q.t.x0.cv.key;
    }

    public static void Delete(Queap Q, QueapNode x) {
        Q.t.deleteLeaf(x);
        --Q.n;
        --Q.k;
    }

    public static void Delete(Queap Q, Element x) {
        QueapNode n;
        if (x.inList) {
            //set inList of all the elements in the list to false
            n = Q.t.insertList(Q.l, x);
            Q.k = Q.n;
            Delete(Q, n);
        }
        else if ((n = Q.t.x0.cv).key == x)
            Delete(Q, n);
    }

    public static Element DeleteMin(Queap Q) {
        Element min = Minimum(Q);
        Delete(Q, min);
        return min;
    }
}
```

## References

[1] \* Iacono, John & Langerman, Stefan: *Queaps*, Springer New York, Algorithmica 42(1): 49–56 (2005)

## Fusion tree

---

A **fusion tree** is a type of tree data structure that implements an associative array on  $w$ -bit integers. It uses  $O(n)$  space and performs searches in  $O(\log_w n)$  time, which is asymptotically faster than a traditional self-balancing binary search tree, and actually better than the van Emde Boas tree when  $w$  is large. It achieves this speed by exploiting certain constant-time operations that can be done on a machine word. Fusion trees were invented in 1990 by Michael Fredman and Dan Willard.<sup>[1]</sup>

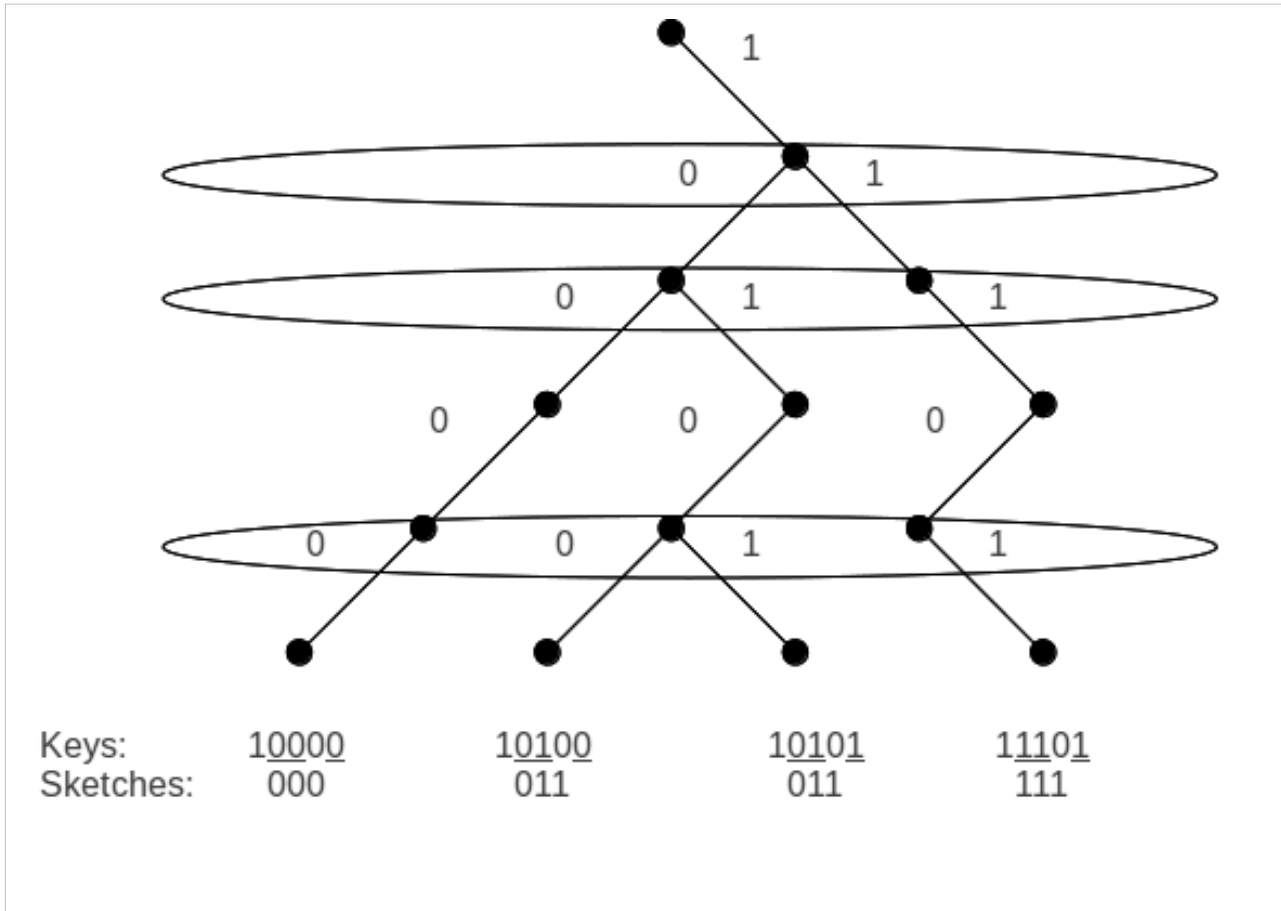
Several advances have been made since Fredman and Willard's original 1990 paper. In 1999<sup>[2]</sup> it was shown how to implement fusion trees under the  $AC^0$  model, in which multiplication no longer takes constant time. A dynamic version of fusion trees using Hash tables was proposed in 1996<sup>[3]</sup> which matched the  $O(\log_w n)$  runtime in expectation. Another dynamic version using Exponential tree was proposed in 2007<sup>[4]</sup> which yields worst-case runtimes of  $O(\log_w n + \log \log u)$  per operation, where  $u$  is the size of the largest key. It remains open whether dynamic fusion trees can achieve  $O(\log_w n)$  per operation with high probability.

## How it works

A fusion tree is essentially a B-tree with branching factor of  $w^{1/5}$  (any small exponent is also possible), which gives it a height of  $O(\log_w n)$ . To achieve the desired runtimes for updates and queries, the fusion tree must be able to search a node containing up to  $w^{1/5}$  keys in constant time. This is done by compressing ("sketching") the keys so that all can fit into one machine word, which in turn allows comparisons to be done in parallel. The rest of this article will describe the operation of a static Fusion Tree; that is, only queries are supported.

## Sketching

Sketching is the method by which each  $w$ -bit key at a node containing  $k$  keys is compressed into only  $k-1$  bits. Each key  $x$  may be thought of as a path in the full binary tree of height  $w$  starting at the root and ending at the leaf corresponding to  $x$ . To distinguish two paths, it suffices to look at their branching point (the first bit where the two keys differ). All  $k$  paths together have  $k-1$  branching points, so at most  $k-1$  bits are needed to distinguish any two of the  $k$  keys.



An important property of the sketch function is that it preserves the order of the keys. That is,  $\text{sketch}(x) < \text{sketch}(y)$  for any two keys  $x < y$ .

## Approximating the sketch

If the locations of the sketch bits are  $b_1 < b_2 < \dots < b_r$ , then the sketch of the key  $x_{w-1} \dots x_1 x_0$  is the  $r$ -bit integer  $x_{b_r} x_{b_{r-1}} \dots x_{b_1}$ .

With only standard word operations, such those of the C programming language, it is difficult to directly compute the sketch of a key in constant time. Instead, the sketch bits can be packed into a range of size at most  $r^4$ , using bitwise AND and multiplication. The bitwise AND operation serves to clear all non-sketch bits from the key, while the multiplication shifts the sketch bits into a small range. Like the "perfect" sketch, the approximate sketch preserves the order of the keys.

Some preprocessing is needed to determine the correct multiplication constant. Each sketch bit in location  $b_i$  will get shifted to  $b_i + m_i$  via a multiplication by  $m = \sum_{i=1}^r 2^{m_i}$ . For the approximate sketch to work, the following three properties must hold:

1.  $b_i + m_j$  are distinct for all pairs  $(i, j)$ . This will ensure that the sketch bits are uncorrupted by the multiplication.
2.  $b_i + m_j$  is a strictly increasing function of  $i$ . That is, the order of the sketch bits is preserved.

3.  $(b_r + m_r) - (b_1 - m_1) \leq r^4$ . That is, the sketch bits are packed into a range of size at most  $r^4$ .

An inductive argument shows how the  $m_t$  can be constructed. Let  $m_1 = w - b_1$ . Suppose that  $1 < t \leq r$  and that  $m_1, m_2, \dots, m_t$  have already been chosen. Then pick the smallest integer  $m_t$  such that both properties (1) and (2) are satisfied. Property (1) requires that  $m_t \neq b_i - b_j + m_l$  for all  $1 \leq i, j \leq r$  and  $1 \leq l \leq t-1$ . Thus, there are less than  $tr^2 \leq r^3$  values that  $m_t$  must avoid. Since  $m_t$  is chosen to be minimal,  $(b_t + m_t) \leq (b_{t-1} + m_{t-1}) + r^3$ . This implies Property (3).

The approximate sketch is thus computed as follows:

1. Mask out all but the sketch bits with a bitwise AND.
2. Multiply the key by the predetermined constant  $m$ . This operation actually requires two machine words, but this can still be done in constant time.
3. Mask out all but the shifted sketch bits. These are now contained in a contiguous block of at most  $r^4 < w^{4/5}$  bits.

For the rest of this article, sketching will be taken to mean approximate sketching.

## Parallel comparison

The purpose of the compression achieved by sketching is to allow all of the keys to be stored in one  $w$ -bit word. Let the *node sketch* of a node be the bit string

$$1\text{sketch}(x_1)1\text{sketch}(x_2)\dots 1\text{sketch}(x_k)$$

We can assume that the sketch function uses exactly  $b \leq r^4$  bits. Then each block uses  $1 + b \leq w^{4/5}$  bits, and since  $k \leq w^{1/5}$ , the total number of bits in the node sketch is at most  $w$ .

A brief notational aside: for a bit string  $s$  and nonnegative integer  $m$ , let  $s^m$  denote the concatenation of  $s$  to itself  $m$  times. If  $t$  is also a bit string  $st$  denotes the concatenation of  $t$  to  $s$ .

The node sketch makes it possible to search the keys for any  $b$ -bit integer  $y$ . Let  $z = (0y)^k$ , which can be computed in constant time (multiply  $y$  by the constant  $(0^b1)^k$ ). Note that  $1\text{sketch}(x_i) - 0y$  is always positive, but preserves its leading 1 iff  $\text{sketch}(x_i) \geq y$ . We can thus compute the smallest index  $i$  such that  $\text{sketch}(x_i) \geq y$  as follows:

1. Subtract  $z$  from the node sketch.
2. Take the bitwise AND of the difference and the constant  $(10^b)^k$ . This clears all but the leading bit of each block.
3. Find the most significant bit of the result.
4. Compute  $i$ , using the fact that the leading bit of the  $i$ -th block has index  $i(b+1)$ .

## Desketching

For an arbitrary query  $q$ , parallel comparison computes the index  $i$  such that

$$\text{sketch}(x_{i-1}) \leq \text{sketch}(q) \leq \text{sketch}(x_i)$$

Unfortunately, the sketch function is not in general order-preserving outside the set of keys, so it is not necessarily the case that  $x_{i-1} \leq q \leq x_i$ . What is true is that, among all of the keys, either  $x_{i-1}$  or  $x_i$  has the longest common prefix with  $q$ . This is because any key  $y$  with a longer common prefix with  $q$  would also have more sketch bits in common with  $q$ , and thus  $\text{sketch}(y)$  would be closer to  $\text{sketch}(q)$  than any  $\text{sketch}(x_j)$ .

The length longest common prefix between two  $w$ -bit integers  $a$  and  $b$  can be computed in constant time by finding the most significant bit of the bitwise XOR between  $a$  and  $b$ . This can then be used to mask out all but the longest common prefix.

Note that  $p$  identifies exactly where  $q$  branches off from the set of keys. If the next bit of  $q$  is 0, then the successor of  $q$  is contained in the  $p1$  subtree, and if the next bit of  $q$  is 1, then the predecessor of  $q$  is contained in the  $p0$  subtree. This suggests the following algorithm:

1. Use parallel comparison to find the index  $i$  such that  $\text{sketch}(x_{i-1}) \leq \text{sketch}(q) \leq \text{sketch}(x_i)$ .

2. Compute the longest common prefix  $p$  of  $q$  and either  $x_{i-1}$  or  $x_i$  (taking the longer of the two).
3. Let  $l-1$  be the length of the longest common prefix  $p$ .
  1. If the  $l$ -th bit of  $q$  is 0, let  $e = p10^{w-l}$ . Use parallel comparison to search for the successor of `sketch(e)`. This is the actual predecessor of  $q$ .
  2. If the  $l$ -th bit of  $q$  is 1, let  $e = p01^{w-l}$ . Use parallel comparison to search for the predecessor of `sketch(e)`. This is the actual successor of  $q$ .
4. Once either the predecessor or successor of  $q$  is found, the exact position of  $q$  among the set of keys is determined.

## References

- [1] M. L. Fredman and D. E. Willard. BLASTING through the information theoretic barrier with FUSION TREES. Proceedings of the twenty-second annual ACM symposium on Theory of Computing, 1-7, 1990.
  - [2] A. Andersson, P. B. Miltersen, and M. Thorup. Fusion trees can be implemented with AC0 instructions only. Theoretical Computer Science, 215:337-344, 1999.
  - [3] R. Raman. Priority queues: Small, monotone, and trans-dichotomous. Algorithms - ESA '96, 121-137, 1996.
  - [4] A. Andersson and M. Thorup. Dynamic ordered sets with exponential search trees. Journal of the ACM, 54:3:13, 2007.
- MIT CS 6.897: Advanced Data Structures: Lecture 4, Fusion Trees ([http://theory.csail.mit.edu/classes/6.897/spring03/scribe\\_notes/L4/lecture4.pdf](http://theory.csail.mit.edu/classes/6.897/spring03/scribe_notes/L4/lecture4.pdf)), Prof. Erik Demaine (Spring 2003)
  - MIT CS 6.897: Advanced Data Structures: Lecture 5, More fusion trees; self-organizing data structures, move-to-front, static optimality ([http://theory.csail.mit.edu/classes/6.897/spring03/scribe\\_notes/L5/lecture5.pdf](http://theory.csail.mit.edu/classes/6.897/spring03/scribe_notes/L5/lecture5.pdf)), Prof. Erik Demaine (Spring 2003)
  - MIT CS 6.851: Advanced Data Structures: Lecture 13, Fusion Tree notes (<http://courses.csail.mit.edu/6.851/spring07/scribe/lec13.pdf>), Prof. Erik Demaine (Spring 2007)
  - MIT CS 6.851: Advanced Data Structures: Lecture 12, Fusion Tree notes (<http://courses.csail.mit.edu/6.851/spring12/scribe/lec12.pdf>), Prof. Erik Demaine (Spring 2012)

# Bx-tree

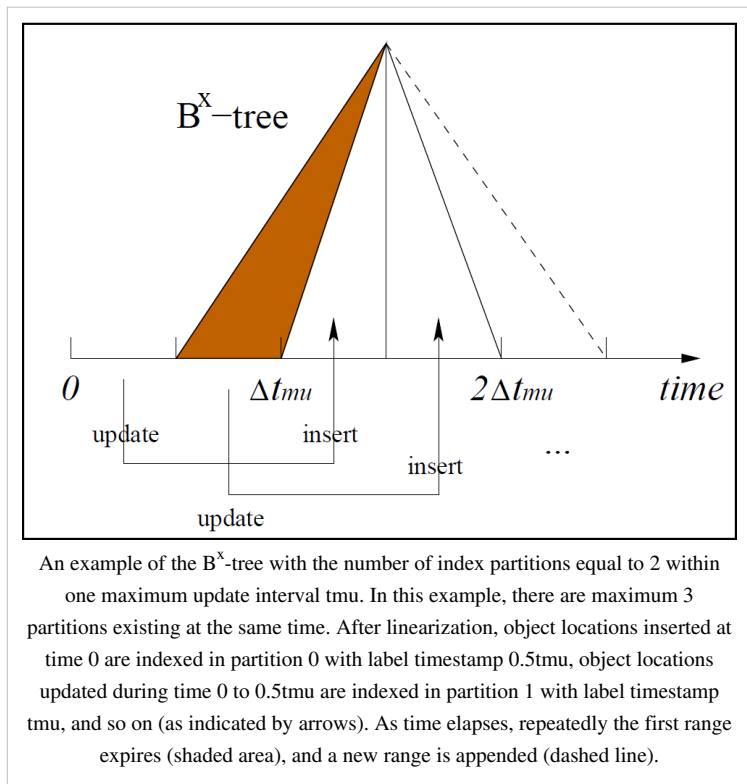
In computer science, the **B<sup>x</sup> tree** is a query and update efficient B+ tree-based index structure for moving objects.

## Index structure

The base structure of the B<sup>x</sup>-tree is a B+ tree in which the internal nodes serve as a directory, each containing a pointer to its right sibling. In the earlier version of the B<sup>x</sup>-tree,<sup>[1]</sup> the leaf nodes contained the moving-object locations being indexed and corresponding index time. In the optimized version,<sup>[2]</sup> each leaf node entry contains the id, velocity, single-dimensional mapping value and the latest update time of the object. The fanout is increased by not storing the locations of moving objects, as these can be derived from the mapping values.

## Utilize the B+ tree for moving objects

As for many other moving objects indexes, a 2-dimensional moving object is modeled as a linear function as  $O = ((x, y), (vx, vy), t)$ , where  $(x, y)$  and  $(vx, vy)$  are location and velocity of the object at a given time instance  $t$ , i.e., the time of last update. The B+ tree is a structure for indexing single dimensional data. In order to adopt the B+ tree as a moving object index, the B<sup>x</sup>-tree uses a linearization technique which helps to integrate objects' location at time  $t$  into single dimensional value. Specifically, objects are first partitioned according to their update time. For objects within the same partition, the B<sup>x</sup>-tree stores their locations at a given time which are estimated by linear interpolation. By doing so, the B<sup>x</sup>-tree keeps a consistent view of all objects within the same partition without storing the update time of an objects.



Secondly, the space is partitioned by a grid and the location of an object is linearized within the partitions according to a space-filling curve, e.g., the Peano or Hilbert curves.

Finally, with the combination of the partition number (time information) and the linear order (location information), an object is indexed in B<sup>x</sup>-tree with a one dimensional index key B<sup>x</sup>value:

$$B^x\text{value}(O, t) = [\text{indexpartition}]_2 + [\text{xrep}]_2$$

Here index-partition is an index partition determined by the update time and xrep is the space-filling curve value of the object position at the indexed time,  $[X]_2$  denotes the binary value of  $x$ , and "+" means concatenation.

Given an object  $O((7, 2), (-0.1, 0.05), 10)$ ,  $t_{mu} = 120$ , the B<sup>x</sup>value for  $O$  can be computed as follows.

1.  $O$  is indexed in partition 0 as mentioned. Therefore,  $\text{indexpartition} = (00)_2$ .
2.  $O$ 's position at the label timestamp of partition 0 is  $(1, 5)$ .
3. Using Z-curve with order = 3, the Z-value of  $O$ , i.e.,  $\text{xrep}$  is  $(010011)_2$ .
4. Concatenating  $\text{indexpartition}$  and  $\text{xrep}$ ,  $B^x\text{value} (00010011)_2 = 19$ .



## Insertion, Update and Deletion

Given a new object, its index key is computed and then the object is inserted into the  $B^x$ -tree as in the B+ tree. An update consists of a deletion followed by an insertion. An auxiliary structure is employed to keep the latest key of each index so that an object can be deleted by searching for the key. The indexing key is computed before affecting the tree. In this way, the  $B^x$ -tree directly inherits the good properties of the B+ tree, and achieves efficient update performance.

## Queries

### Range query

A range query retrieves all objects whose location falls within the rectangular range  $q = ([qx1, qy1]; [qx2, qy2])$  at time  $tq$  not prior to the current time.

The  $B^x$ -tree uses query-window enlargement technique to answer queries. Since the  $B^x$ -tree stores an object's location as of sometime after its update time, the enlargement involves two cases: a location must either be brought back to an earlier time or forward to a later time. The main idea is to enlarge the query window so that it encloses all objects whose positions are not within query window at its label timestamp but will enter the query window at the query timestamp.

After the enlargement, the partitions of the  $B^x$ -tree need to be traversed to find objects falling in the enlarged query window. In each partition, the use of a space-filling curve means that a range query in the native, two-dimensional space becomes a set of range queries in the transformed, one-dimensional space.<sup>[1]</sup>

To avoid excessively large query region after expansion in skewed datasets, an optimization of the query algorithm exists,<sup>[3]</sup> which improves the query efficiency by avoiding unnecessary query enlargement.

### K nearest neighbor query

K nearest neighbor query is computed by iteratively performing range queries with an incrementally enlarged search region until k answers are obtained. Another possibility is to employ similar querying ideas in The iDistance Technique.

### Other queries

The range query and K Nearest Neighbor query algorithms can be easily extended to support interval queries, continuous queries, etc.<sup>[2]</sup>

## Adapting relational database engines to accommodate moving objects

Since the  $B^x$ -tree is an index built on top of a B+ tree index, all operations in the  $B^x$ -tree, including the insertion, deletion and search, are the same as those in the B+ tree. There is no need to change the implementations of these operations. The only difference is to implement the procedure of deriving the indexing key as a stored procedure in an existing DBMS. Therefore the  $B^x$ -tree can be easily integrated into existing DBMS without touching the kernel.

SpADE<sup>[4]</sup> is moving object management system built on top of a popular relational database system MySQL, which uses the  $B^x$ -tree for indexing the objects. In the implementation, moving object data is transformed and stored directly on MySQL, and queries are transformed into standard SQL statements which are efficiently processed in the relational engine. Most importantly, all these are achieved neatly and independently without infiltrating into the MySQL core.

## Performance tuning

### Potential problem with data skew

The  $B^x$  tree uses a grid for space partitioning while mapping two-dimensional location into one-dimensional key. This may introduce performance degradation to both query and update operations while dealing with skewed data. If grid cell is oversize, many objects are contained in a cell. Since objects in a cell are indistinguishable to the index, there will be some overflow nodes in the underlying B+ tree. The existing of overflow pages not only destroys the balancing of the tree but also increases the update cost. As for the queries, for the given query region, large cell incurs more false positives and increases the processing time. On the other hand, if the space is partitioned with finer grid, i.e. smaller cells, each cell contains few objects. There is hardly overflow pages so that the update cost is minimized. Fewer false positives are retrieved in a query. However, more cells are needed to be searched. The increase in the number of cells searched also increases the workload of a query.

### Index tuning

The  $ST^2B$ -tree<sup>[5]</sup> introduces a self-tuning framework for tuning the performance of the  $B^x$ -tree while dealing with data skew in space and data change with time. In order to deal with data skew in space, the  $ST^2B$ -tree splits the entire space into regions of different object density using a set of reference points. Each region uses an individual grid whose cell size is determined by the object density inside of it.

The  $B^x$ -tree have multiple partitions regarding different time intervals. As time elapsed, each partition grows and shrinks alternately. The  $ST^2B$ -tree utilizes this feature to tune the index online in order to adjust the space partitioning to make itself accommodate to the data changes with time. In particular, as a partition shrinks to empty and starts growing, it chooses a new set of reference points and new grid for each reference point according to the latest data density. The tuning is based on the latest statistics collected during a given period of time, so that the way of space partitioning is supposed to fit the latest data distribution best. By this means, the  $ST^2B$ -tree is expected to minimize the effect caused by data skew in space and data changes with time.

## References

- [1] Christian S. Jensen, Dan Lin, and Beng Chin Ooi. Query and Update Efficient B+tree based Indexing of Moving Objects (<http://www.vldb.org/conf/2004/RS20P3.PDF>). In Proceedings of 30th International Conference on Very Large Data Bases (VLDB), pages 768-779, 2004.
- [2] Dan Lin. Indexing and Querying Moving Objects Databases ([http://web.mst.edu/~lindan/publication/thesis\\_lindan.pdf](http://web.mst.edu/~lindan/publication/thesis_lindan.pdf)), PhD thesis, National University of Singapore, 2006.
- [3] Jensen, C.S., D. Tiesyte, N. Tradisaukas, Robust B+-Tree-Based Indexing of Moving Objects, in Proceedings of the Seventh International Conference on Mobile Data Management ([http://www.cs.aau.dk/~csj/Papers/Files/2006\\_JensenMDM.pdf](http://www.cs.aau.dk/~csj/Papers/Files/2006_JensenMDM.pdf)), Nara, Japan, 9 pages, May 9–12, 2006.
- [4] SpADE (<http://www.comp.nus.edu.sg/~spade/>): A SPatio-temporal Autonomic Database Engine for location-aware services.
- [5] Su Chen, Beng Chin Ooi, Kan-Lee. Tan, and Mario A. Nacimento,  $ST^2B$ -tree: A Self-Tunable Spatio-Temporal B+-tree for Moving Objects (<http://www.comp.nus.edu.sg/~chensu/sigmod08.pdf>). In Proceedings of ACM SIGMOD International Conference on Management of Data (SIGMOD), page 29-42, 2008.

# Heaps

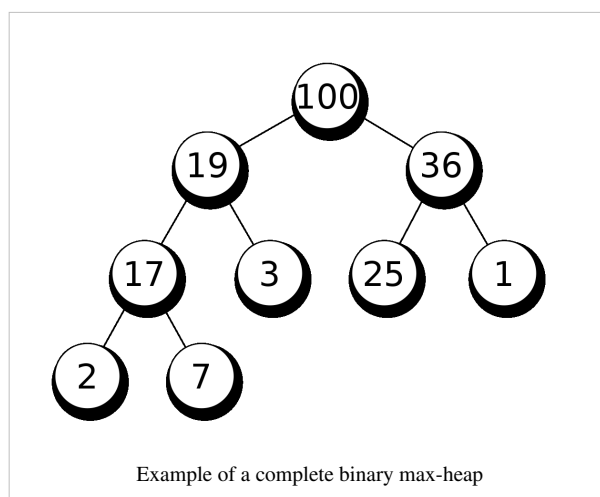
## Heap

In computer science, a **heap** is a specialized tree-based data structure that satisfies the *heap property*: If  $A$  is a parent node of  $B$  then  $\text{key}(A)$  is ordered with respect to  $\text{key}(B)$  with the same ordering applying across the heap. In a *max heap* the keys of parent nodes are always greater than or equal to those of the children and the highest key exists in the root node. In a *min heap* the keys of parent nodes are always less than or equal to those of the children and the root node has the lowest key.

Note that, as shown in the graphic, there is no implied ordering between siblings or cousins and no implied sequence for an in-order traversal (as there would be in, e.g., a binary search tree). The heap relation mentioned above applies only between nodes and their immediate parents.

The maximum number of children each node can have depends on the type of heap, but in many types it is at most two. The heap is one maximally efficient implementation of an abstract data type called a priority queue. Heaps are crucial in several efficient graph algorithms such as Dijkstra's algorithm, and in the sorting algorithm heapsort.

A *heap* data structure should not be confused with *the heap* which is a common name for dynamically allocated memory. The term was originally used only for the data structure.



## Implementation and operations

Heaps are usually implemented in an array, and do not require pointers between elements.

The operations commonly performed with a heap are:

- *create-heap*: create an empty heap
- *find-max* or *find-min*: find the maximum item of a max-heap or a minimum item of a min-heap, respectively
- *delete-max* or *delete-min*: removing the root node of a max- or min-heap, respectively
- *increase-key* or *decrease-key*: updating a key within a max- or min-heap, respectively
- *insert*: adding a new key to the heap
- *merge*: joining two heaps to form a valid new heap containing all the elements of both.

Different types of heaps implement the operations in different ways, but notably, insertion is often done by adding the new element at the end of the heap in the first available free space. This will tend to violate the heap property, and so the elements are then reordered until the heap property has been reestablished.

## Variants

- 2-3 heap
- Beap
- Binary heap
- Binomial heap
- Brodal queue
- D-ary heap
- Fibonacci heap
- Leftist heap
- Pairing heap
- Skew heap
- Soft heap
- Weak heap
- Leaf heap
- Radix heap
- Randomized meldable heap

## Comparison of theoretic bounds for variants

The following time complexities<sup>[1]</sup> are amortized (worst-time) time complexity for entries marked by an asterisk, and regular worst case time complexities for all other entries.  $O(f)$  gives asymptotic upper bound and  $\Theta(f)$  is asymptotically tight bound (see Big O notation). Function names assume a min-heap.

Operation	Binary <sup>[1]</sup>	Binomial <sup>[1]</sup>	Fibonacci <sup>[1]</sup>	Pairing <sup>[2]</sup>	Brodal*** <sup>[3]</sup>	RP <sup>[4]</sup>
find-min	$\Theta(1)$	$O(\log n)$	$\Theta(1)$	$O(1)^*$	$O(1)$	$\Theta(1)$
delete-min	$\Theta(\log n)$	$\Theta(\log n)$	$O(\log n)^*$	$O(\log n)^*$	$O(\log n)$	$O(\log n)^*$
insert	$\Theta(\log n)$	$O(\log n)$	$\Theta(1)$	$O(1)^*$	$O(1)$	$\Theta(1)$
decrease-key	$\Theta(\log n)$	$\Theta(\log n)$	$\Theta(1)^*$	$O(\log n)^*$	$O(1)$	$\Theta(1)^*$
merge	$\Theta(n)$	$O(\log n)^{**}$	$\Theta(1)$	$O(1)^*$	$O(1)$	$\Theta(1)$

(\*)Amortized time

(\*\*)Where  $n$  is the size of the larger heap

(\*\*\*)Brodal and Okasaki later describe a persistent variant with the same bounds except for decrease-key, which is not supported. Heaps with  $n$  elements can be constructed bottom-up in  $O(n)$ .<sup>[5]</sup>

## Applications

The heap data structure has many applications.

- Heapsort: One of the best sorting methods being in-place and with no quadratic worst-case scenarios.
- Selection algorithms: Finding the min, max, both the min and max, median, or even the  $k$ -th largest element can be done in linear time (often constant time) using heaps.<sup>[6]</sup>
- Graph algorithms: By using heaps as internal traversal data structures, run time will be reduced by polynomial order. Examples of such problems are Prim's minimal spanning tree algorithm and Dijkstra's shortest path problem.

Full and almost full binary heaps may be represented in a very space-efficient way using an array alone. The first (or last) element will contain the root. The next two elements of the array contain its children. The next four contain the four children of the two child nodes, etc. Thus the children of the node at position  $n$  would be at positions  $2n$  and

$2n+1$  in a one-based array, or  $2n+1$  and  $2n+2$  in a zero-based array. This allows moving up or down the tree by doing simple index computations. Balancing a heap is done by swapping elements which are out of order. As we can build a heap from an array without requiring extra memory (for the nodes, for example), heapsort can be used to sort an array in-place.

## Implementations

- The C++ Standard Template Library provides the `make_heap`, `push_heap` and `pop_heap` algorithms for heaps (usually implemented as binary heaps), which operate on arbitrary random access iterators. It treats the iterators as a reference to an array, and uses the array-to-heap conversion. Container adaptor `priority_queue` also exists. However, there is no standard support for the decrease/increase-key operation. See also `gheap`<sup>[7]</sup> - STL-like generalized heap implementation in C++ with D-heap and B-heap support.
- The Java 2 platform (since version 1.5) provides the binary heap implementation with class `java.util.PriorityQueue<E>`<sup>[8]</sup> in Java Collections Framework.
- Python has a `heapq`<sup>[6]</sup> module that implements a priority queue using a binary heap.
- PHP has both `maxheap` (`SplMaxHeap`) and `minheap` (`SplMinHeap`) as of version 5.3 in the Standard PHP Library.
- Perl has implementations of binary, binomial, and Fibonacci heaps in the `Heap`<sup>[9]</sup> distribution available on CPAN.
- The Go library contains a `heap`<sup>[7]</sup> package with heap algorithms that operate on an arbitrary type that satisfied a given interface.
- Apple's Core Foundation library contains a `CFBinaryHeap`<sup>[10]</sup> structure.

## References

- [1] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest (1990): Introduction to algorithms. MIT Press / McGraw-Hill.
- [2] Iacono, John (2000), "Improved upper bounds for pairing heaps", *Proc. 7th Scandinavian Workshop on Algorithm Theory*, Lecture Notes in Computer Science, **1851**, Springer-Verlag, pp. 63–77, doi:10.1007/3-540-44985-X\_5
- [3] <http://www.cs.au.dk/~gerth/papers/soda96.pdf>
- [4] Haeupler, Bernhard; Sen, Siddhartha; Tarjan, Robert E. (2009). "Rank-pairing heaps" (<http://www.cs.princeton.edu/~sssix/papers/rp-heaps.pdf>). *SIAM J. Computing*: 1463–1485. .
- [5] Goodrich, Michael T.; Tamassia, Roberto (2004). "7.3.6. Bottom-Up Heap Construction". *Data Structures and Algorithms in Java* (3rd ed.). pp. 338–341.
- [6] Frederickson, Greg N. (1993), "An Optimal Algorithm for Selection in a Min-Heap" ([http://ftp.cs.purdue.edu/research/technical\\_reports/1991/TR\\_91-027.pdf](http://ftp.cs.purdue.edu/research/technical_reports/1991/TR_91-027.pdf)), *Information and Computation*, **104**, Academic Press, pp. 197–214, doi:10.1006/inco.1993.1030,
- [7] <https://github.com/valyala/gheap>
- [8] <http://docs.oracle.com/javase/6/docs/api/java/util/PriorityQueue.html>
- [9] <http://search.cpan.org/perldoc?Heap>
- [10] <https://developer.apple.com/library/mac/#documentation/CoreFoundation/Reference/CFBinaryHeapRef/Reference/reference.html>

## External links

- Heap (<http://mathworld.wolfram.com/Heap.html>) at Wolfram MathWorld

# Binary heap

Binary Heap		
Type	Tree	
Time complexity in big O notation		
	Average	Worst case
Space	O(n)	O(n)
Search	N/A Operation	N/A Operation
Insert	O(log n)	O(log n)
Delete	O(log n)	O(log n)

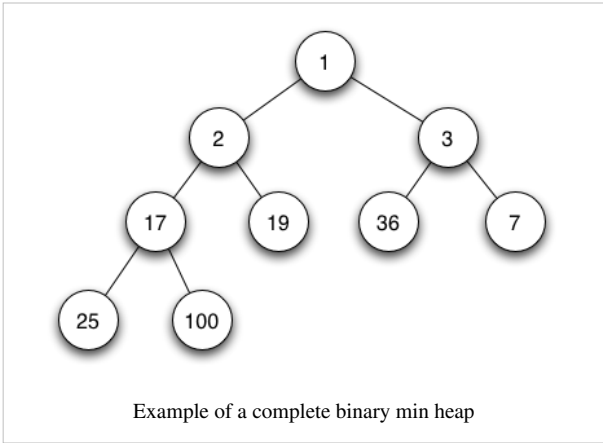
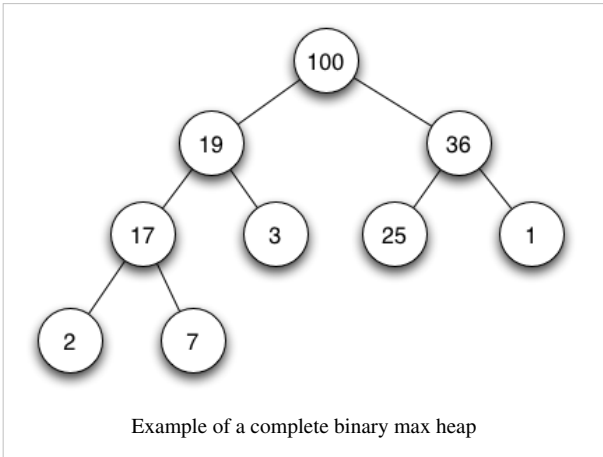
A **binary heap** is a heap data structure created using a binary tree. It can be seen as a binary tree with two additional constraints:

- The *shape property*: the tree is a *complete binary tree*; that is, all levels of the tree, except possibly the last one (deepest) are fully filled, and, if the last level of the tree is not complete, the nodes of that level are filled from left to right.
- The *heap property*: each node is greater than or equal to each of its children according to a comparison predicate defined for the data structure.

Heaps with a mathematical "greater than or equal to" comparison function are called *max-heaps*; those with a mathematical "less than or equal to" comparison function are called *min-heaps*. Min-heaps are often used to implement priority queues.<sup>[1][2]</sup>

Since the ordering of siblings in a heap is not specified by the heap property, a single node's two children can be freely interchanged unless doing so violates the shape property (compare with treap).

The binary heap is a special case of the d-ary heap in which  $d = 2$ .



## Heap operations

Both the insert and remove operations modify the heap to conform to the shape property first, by adding or removing from the end of the heap. Then the heap property is restored by traversing up or down the heap. Both operations take  $O(\log n)$  time.

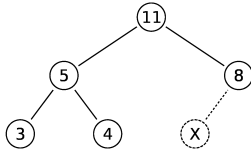
## Insert

To add an element to a heap we must perform an *up-heap* operation (also known as *bubble-up*, *percolate-up*, *sift-up*, *trickle up*, *heapify-up*, or *cascade-up*), by following this algorithm:

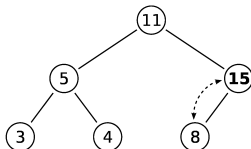
1. Add the element to the bottom level of the heap.
2. Compare the added element with its parent; if they are in the correct order, stop.
3. If not, swap the element with its parent and return to the previous step.

The number of operations required is dependent on the number of levels the new element must rise to satisfy the heap property, thus the insertion operation has a time complexity of  $O(\log n)$ .

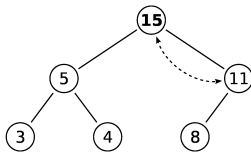
As an example, say we have a max-heap



and we want to add the number 15 to the heap. We first place the 15 in the position marked by the X. However, the heap property is violated since 15 is greater than 8, so we need to swap the 15 and the 8. So, we have the heap looking as follows after the first swap:



However the heap property is still violated since 15 is greater than 11, so we need to swap again:



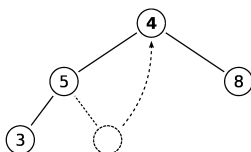
which is a valid max-heap. There is no need to check the children after this. Before we placed 15 on X, the heap was valid, meaning 11 is greater than 5. If 15 is greater than 11, and 11 is greater than 5, then 15 must be greater than 5, because of the transitive relation.

## Delete

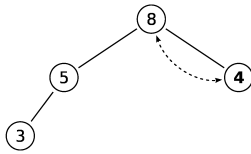
The procedure for deleting the root from the heap (effectively extracting the maximum element in a max-heap or the minimum element in a min-heap) and restoring the properties is called *down-heap* (also known as *bubble-down*, *percolate-down*, *sift-down*, *trickle down*, *heapify-down*, *cascade-down* and *extract-min/max*).

1. Replace the root of the heap with the last element on the last level.
2. Compare the new root with its children; if they are in the correct order, stop.
3. If not, swap the element with one of its children and return to the previous step. (Swap with its smaller child in a min-heap and its larger child in a max-heap.)

So, if we have the same max-heap as before, we remove the 11 and replace it with the 4.



Now the heap property is violated since 8 is greater than 4. In this case, swapping the two elements, 4 and 8, is enough to restore the heap property and we need not swap elements further:



The downward-moving node is swapped with the *larger* of its children in a max-heap (in a min-heap it would be swapped with its smaller child), until it satisfies the heap property in its new position. This functionality is achieved by the **Max-Heapify** function as defined below in pseudocode for an array-backed heap  $A$ . Note that " $A$ " is indexed starting at 1, not 0 as is common in many programming languages.

For the following algorithm to correctly re-heapify the array, the node at index  $i$  and its two direct children must violate the heap property. If they do not, the algorithm will fall through with no change to the array.

**Max-Heapify**<sup>[3]</sup> ( $A, i$ ):

$left \leftarrow 2i$

$right \leftarrow 2i + 1$

$largest \leftarrow i$

**if**  $left \leq \text{heap\_length}[A]$  **and**  $A[left] > A[largest]$  **then**:

$largest \leftarrow left$

**if**  $right \leq \text{heap\_length}[A]$  **and**  $A[right] > A[largest]$  **then**:

$largest \leftarrow right$

**if**  $largest \neq i$  **then**:

**swap**  $A[i] \leftrightarrow A[largest]$

Max-Heapify( $A, largest$ )

The down-heap operation (without the preceding swap) can also be used to modify the value of the root, even when an element is not being deleted.

In the worst case, the new root has to be swapped with its child on each level until it reaches the bottom level of the heap, meaning that the delete operation has a time complexity relative to the height of the tree, or  $O(\log n)$ .

An amortized analysis of a sequence of heap insertions and deletions can associate the insert operation with  $O(\log n)$  complexity, while the delete operation can be satisfied by  $O(1)$  complexity as long as deleting from an empty heap is forbidden. This is possible because any element's insert operation can be thought as "paying off in advance" the eventual complexity of its deletion from the tree for overall  $O(\log n)$  complexity. However, it's not possible to associate  $O(1)$  with the insert operation in such analysis. Amortized analysis by its very nature does not affect the actual complexity of individual operations.

## Building a heap

A heap could be built by successive insertions. This approach requires  $O(n \log n)$  time because each insertion takes  $O(\log n)$  time and there are  $n$  elements. However this is not the optimal method. The optimal method starts by arbitrarily putting the elements on a binary tree, respecting the shape property (the tree could be represented by an array, see below). Then starting from the lowest level and moving upwards, shift the root of each subtree downward as in the deletion algorithm until the heap property is restored. More specifically if all the subtrees starting at some height  $h$  (measured from the bottom) have already been "heapified", the trees at height  $h + 1$  can be heapified by sending their root down along the path of maximum valued children when building a max-heap, or minimum valued children when building a min-heap. This process takes  $O(h)$  operations (swaps) per node. In this method most of the heapification takes place in the lower levels. Since the height of the heap is  $\lfloor \lg(n) \rfloor$ , the number of nodes at

height  $h$  is  $\leq \left\lceil \frac{2^{\lg n - h} - 1}{2^{h+1}} \right\rceil = \left\lceil \frac{n}{2^{h+1}} \right\rceil$ . Therefore, the cost of heapifying all subtrees is:



$$\begin{aligned}
\sum_{h=0}^{\lceil \lg n \rceil} \frac{n}{2^{h+1}} O(h) &= O \left( n \sum_{h=0}^{\lceil \lg n \rceil} \frac{h}{2^{h+1}} \right) \\
&\leq O \left( n \sum_{h=0}^{\infty} \frac{h}{2^{h+1}} \right) \\
&= O(n)
\end{aligned}$$

This uses the fact that the given infinite series  $h / 2^h$  converges to 2.

The **Build-Max-Heap** function that follows, converts an array  $A$  which stores a complete binary tree with  $n$  nodes to a max-heap by repeatedly using **Max-Heapify** in a bottom up manner. It is based on the observation that the array elements indexed by  $\text{floor}(n/2) + 1, \text{floor}(n/2) + 2, \dots, n$  are all leaves for the tree, thus each is a one-element heap.

**Build-Max-Heap** runs **Max-Heapify** on each of the remaining tree nodes.

**Build-Max-Heap**<sup>[3]</sup> ( $A$ ):

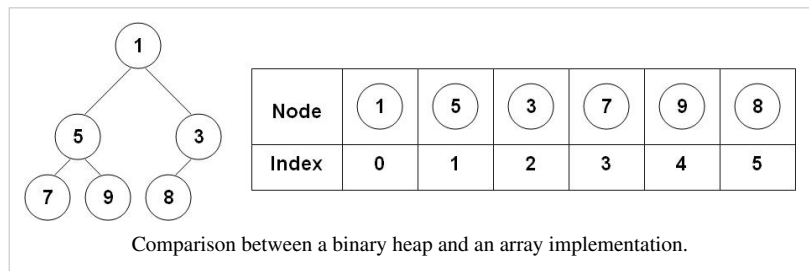
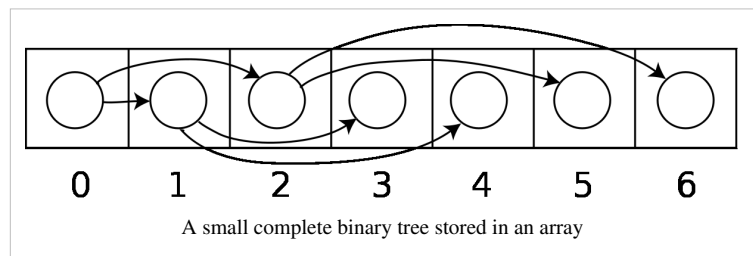
$\text{heap\_length}[A] \leftarrow \text{length}[A]$

**for**  $i \leftarrow \text{floor}(\text{length}[A]/2)$  **downto** 1 **do**

**Max-Heapify**( $A, i$ )

## Heap implementation

Heaps are commonly implemented with an array. Any binary tree can be stored in an array, but because a heap is always an almost complete binary tree, it can be stored compactly. No space is required for pointers; instead, the parent and children of each node can be found by arithmetic on array indices. These properties make this heap implementation a simple example of an implicit data structure or Ahnentafel list. Details depend on the root position, which in turn may depend on constraints of a programming language used for



implementation, or programmer preference. Specifically, sometimes the root is placed at index 1, wasting space in order to simplify arithmetic.

Let  $n$  be the number of elements in the heap and  $i$  be an arbitrary valid index of the array storing the heap. If the tree root is at index 0, with valid indices 0 through  $n-1$ , then each element  $a[i]$  has

- children  $a[2i+1]$  and  $a[2i+2]$
- parent  $a[\text{floor}((i-1)/2)]$

Alternatively, if the tree root is at index 1, with valid indices 1 through  $n$ , then each element  $a[i]$  has

- children  $a[2i]$  and  $a[2i+1]$
- parent  $a[\text{floor}(i/2)]$ .

This implementation is used in the heapsort algorithm, where it allows the space in the input array to be reused to store the heap (i.e. the algorithm is done in-place). The implementation is also useful for use as a Priority queue where use of a dynamic array allows insertion of an unbounded number of items.

The upheap/downheap operations can then be stated in terms of an array as follows: suppose that the heap property holds for the indices  $b, b+1, \dots, e$ . The sift-down function extends the heap property to  $b-1, b, b+1, \dots, e$ . Only index  $i = b-1$  can violate the heap property. Let  $j$  be the index of the largest child of  $a[i]$  (for a max-heap, or the smallest child for a min-heap) within the range  $b, \dots, e$ . (If no such index exists because  $2i > e$  then the heap property holds for the newly extended range and nothing needs to be done.) By swapping the values  $a[i]$  and  $a[j]$  the heap property for position  $i$  is established. At this point, the only problem is that the heap property might not hold for index  $j$ . The sift-down function is applied tail-recursively to index  $j$  until the heap property is established for all elements.

The sift-down function is fast. In each step it only needs two comparisons and one swap. The index value where it is working doubles in each iteration, so that at most  $\log_2 e$  steps are required.

For big heaps and using virtual memory, storing elements in an array according to the above scheme is inefficient: (almost) every level is in a different page. B-heaps are binary heaps that keep subtrees in a single page, reducing the number of pages accessed by up to a factor of ten.<sup>[4]</sup>

The operation of merging two binary heaps takes  $\Theta(n)$  for equal-sized heaps. The best you can do is (in case of array implementation) simply concatenating the two heap arrays and build a heap of the result.<sup>[5]</sup> When merging is a common task, a different heap implementation is recommended, such as binomial heaps, which can be merged in  $O(\log n)$ .

Additionally, a binary heap can be implemented with a traditional binary tree data structure, but there is an issue with finding the adjacent element on the last level on the binary heap when adding an element. This element can be determined algorithmically or by adding extra data to the nodes, called "threading" the tree—instead of merely storing references to the children, we store the inorder successor of the node as well.

It is possible to modify the heap structure to allow extraction of both the smallest and largest element in  $O(\log n)$  time.<sup>[6]</sup> To do this, the rows alternate between min heap and max heap. The algorithms are roughly the same, but, in each step, one must consider the alternating rows with alternating comparisons. The performance is roughly the same as a normal single direction heap. This idea can be generalised to a min-max-median heap.

## Derivation of children's index in an array implementation

This derivation will show how for any given node  $i$  (starts from zero), its children would be found at  $2i + 1$  and  $2i + 2$ .

### Mathematical proof

From the figure in "Heap Implementation" section, it can be seen that any node can store its children only after its right siblings and its left siblings' children have been stored. This fact will be used for derivation.

Total number of elements from root to any given level  $l = 2^{l+1} - 1$ , where  $l$  starts at zero.

Suppose the node  $i$  is at level  $l$ .

So, the total number of nodes from root to previous level would be  $= 2^{(l-1)+1} - 1 = 2^l - 1$

Total number of nodes stored in the array till the index  $i = i + 1$  (Counting  $i$  too)

So, total number of siblings on the left of  $i$  is

$$\begin{aligned}
 &= \text{Number of nodes including } i - \text{Number of nodes through the previous level} - \text{One node for } i \text{ itself} \\
 &= (i + 1) - (2^l - 1) - 1 \\
 &= i + 1 - 2^l + 1 - 1 \\
 &= i - 2^l + 1
 \end{aligned}$$

Hence, total number of children of these siblings  $= 2(i - 2^l + 1)$

Number of elements at any given level  $l = 2^l$

So, total siblings to right of  $i$  is:-

$$\begin{aligned}
 &= \text{Total nodes in level } l - (\text{Total siblings on left} + 1) \\
 &= (2^l) - (i - 2^l + 2) \\
 &= 2^l + 2^l - i - 2 \\
 &= 2^{l+1} - i - 2
 \end{aligned}$$

So, index of 1st child of node  $i$  would be:-

$$\begin{aligned}
 &= i + \text{Total siblings on right} + 2 * \text{Total siblings on left} + 1 \\
 &= i + (2^{l+1} - i - 2) + 2(i - 2^l + 1) + 1 \\
 &= i + 2^{l+1} - i - 2 + 2i - 2^{l+1} + 2 + 1 \\
 &= i - i + 2i + 2^{l+1} - 2^{l+1} - 2 + 2 + 1 \\
 &= 2i + 1 [\text{Proved}]
 \end{aligned}$$

## Intuitive proof

Although the mathematical approach proves this without doubt, the simplicity of the resulting equation suggests that there should be a simpler way to arrive at this conclusion.

For this two facts should be noted.

- Children for node  $i$  will be found at the very first empty slot.
- Second is that, all nodes previous to node  $i$ , right up to the root, will have exactly two children. This is necessary to maintain the shape of the heap.

Now since all nodes have two children (as per the second fact) so all memory slots taken by the children will be  $2((i + 1) - 1) = 2i$ . We add one since  $i$  starts at zero. Then we subtract one since node  $i$  doesn't yet have any children.

This means all filled memory slots have been accounted for except one – the root node. Root is child to none. So finally, the count of all filled memory slots are  $2i + 1$ .

So, by fact one and since our indexing starts at zero,  $2i + 1$  itself gives the index of the first child of  $i$ .

## References

- [1] "heapq – Heap queue algorithm" (<http://docs.python.org/library/heapq.html>). *Python Standard Library*. .
- [2] "Class PriorityQueue" (<http://download.oracle.com/javase/6/docs/api/java/util/PriorityQueue.html>). *Java™ Platform Standard Ed. 6*. .
- [3] Cormen, T. H. & al. (2001), *Introduction to Algorithms* (2nd ed.), Cambridge, Massachusetts: The MIT Press, ISBN 0-07-013151-1
- [4] Poul-Henning Kamp. "You're Doing It Wrong" (<http://queue.acm.org/detail.cfm?id=1814327>). ACM Queue. June 11, 2010.
- [5] Chris L. Kuszmaul. "binary heap" (<http://nist.gov/dads/HTML/binaryheap.html>). Dictionary of Algorithms and Data Structures, Paul E. Black, ed., U.S. National Institute of Standards and Technology. 16 November 2009.
- [6] Atkinson, M.D., J.-R. Sack, N. Santoro, and T. Strothotte (1 October 1986). "Min-max heaps and generalized priority queues." (<http://cg.scs.carleton.ca/~morin/teaching/5408/refs/minmax.pdf>). Programming techniques and Data structures. Comm. ACM, 29(10): 996–1000. .

## External links

- Binary Heap Applet (<http://people.ksp.sk/~kuko/bak/index.html>) by Kubo Kovac
- Using Binary Heaps in A\* Pathfinding (<http://www.policyalmanac.org/games/binaryHeaps.htm>)
- Java Implementation of Binary Heap ([http://sites.google.com/site/indy256/algo-en/binary\\_heap](http://sites.google.com/site/indy256/algo-en/binary_heap))
- C++ implementation of generalized heap with Binary Heap support (<https://github.com/valyala/gheap>)
- Open Data Structures - Section 10.1 - BinaryHeap: An Implicit Binary Tree ([http://opendatastructures.org/versions/edition-0.1e/ods-java/10\\_1\\_BinaryHeap\\_Implicit\\_Bi.html](http://opendatastructures.org/versions/edition-0.1e/ods-java/10_1_BinaryHeap_Implicit_Bi.html))
- Ighushev, Eduard. "Binary Heap C++ implementation" (<http://igushev.com/implementations/binary-heap-cpp/>).

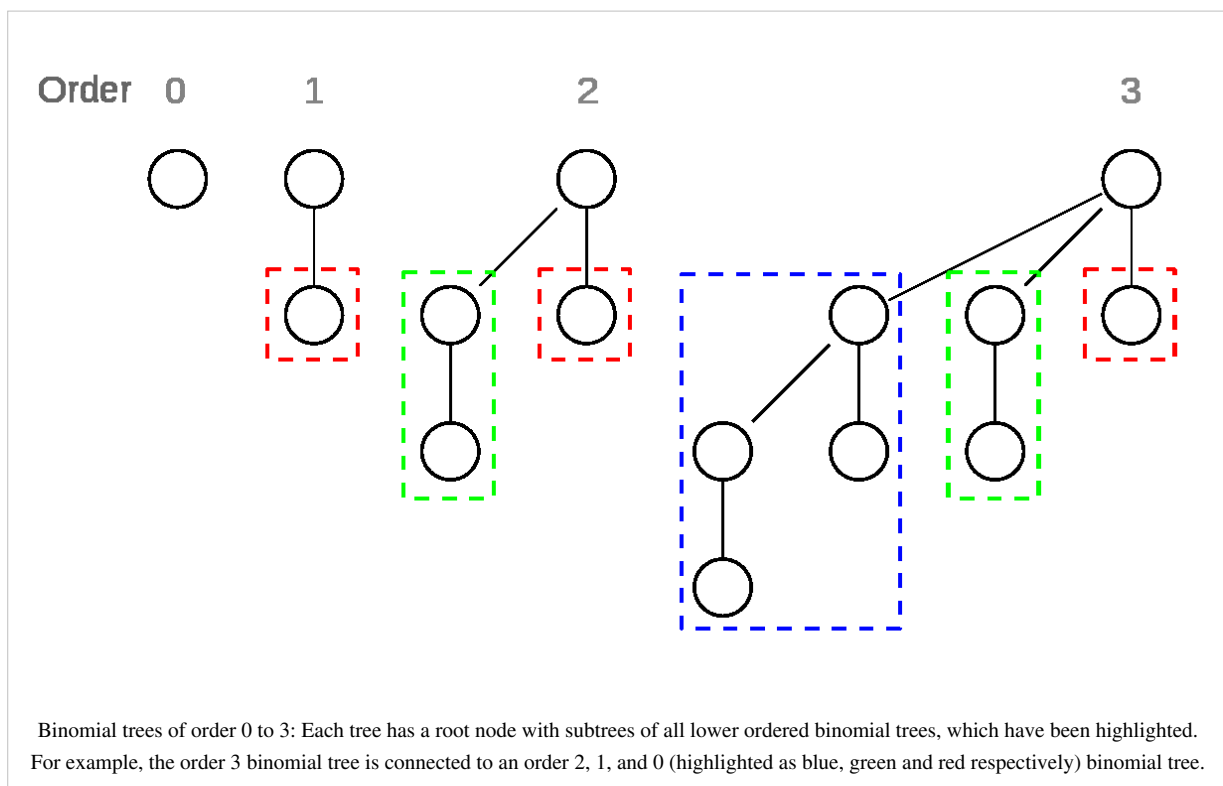
## Binomial heap

In computer science, a **binomial heap** is a heap similar to a binary heap but also supports quick merging of two heaps. This is achieved by using a special tree structure. It is important as an implementation of the mergeable heap abstract data type (also called meldable heap), which is a priority queue supporting merge operation.

### Binomial tree

A binomial heap is implemented as a collection of binomial trees (compare with a binary heap, which has a shape of a single binary tree). A **binomial tree** is defined recursively:

- A binomial tree of order 0 is a single node
- A binomial tree of order  $k$  has a root node whose children are roots of binomial trees of orders  $k-1, k-2, \dots, 2, 1, 0$  (in this order).



A binomial tree of order  $k$  has  $2^k$  nodes, height  $k$ .

Because of its unique structure, a binomial tree of order  $k$  can be constructed from two trees of order  $k-1$  trivially by attaching one of them as the leftmost child of root of the other one. This feature is central to the *merge* operation of a

binomial heap, which is its major advantage over other conventional heaps.

The name comes from the shape: a binomial tree of order  $n$  has  $\binom{n}{d}$  nodes at depth  $d$ . (See Binomial coefficient.)

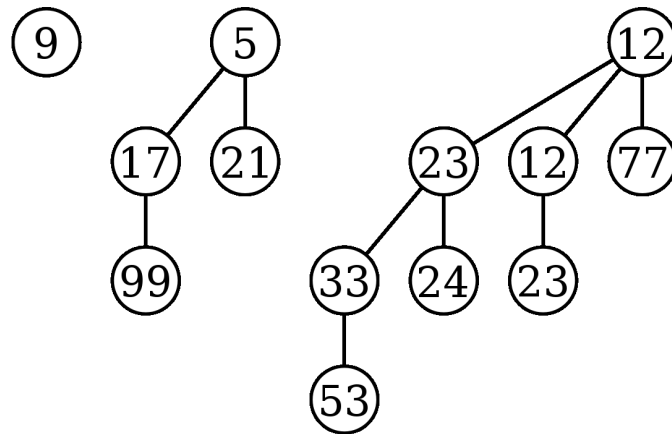
## Structure of a binomial heap

A binomial heap is implemented as a set of binomial trees that satisfy the *binomial heap properties*:

- Each binomial tree in a heap obeys the *minimum-heap property*: the key of a node is greater than or equal to the key of its parent.
- There can only be either *one* or *zero* binomial trees for each order, including zero order.

The first property ensures that the root of each binomial tree contains the smallest key in the tree, which applies to the entire heap.

The second property implies that a binomial heap with  $n$  nodes consists of at most  $\log n + 1$  binomial trees. In fact, the number and orders of these trees are uniquely determined by the number of nodes  $n$ : each binomial tree corresponds to one digit in the binary representation of number  $n$ : for example number 13 is 1101 in binary,  $2^3 + 2^2 + 2^0$ , and thus a binomial heap with 13 nodes will consist of three binomial trees of orders 3, 2, and 0 (see figure below).



*Example of a binomial heap containing 13 nodes with distinct keys.*

*The heap consists of three binomial trees with orders 0, 2, and 3.*

## Implementation

Because no operation requires random access to the root nodes of the binomial trees, the roots of the binomial trees can be stored in a linked list, ordered by increasing order of the tree.

### Merge

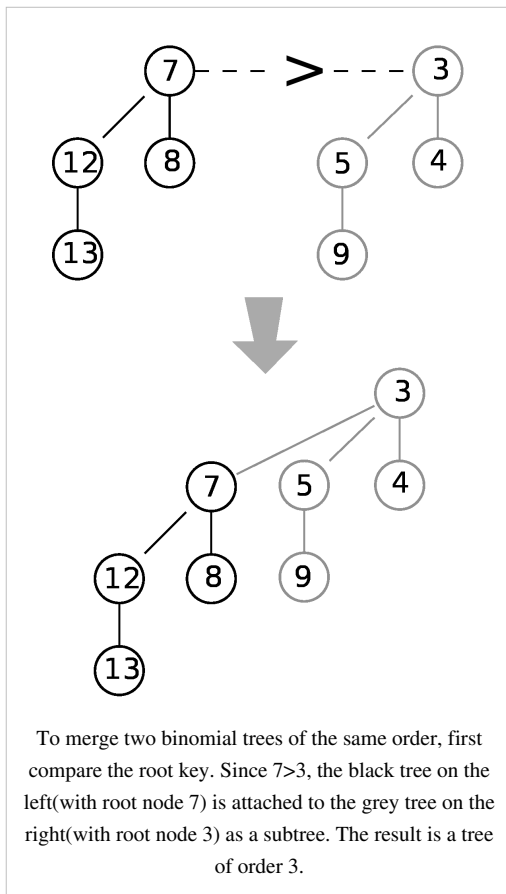
As mentioned above, the simplest and most important operation is the merging of two binomial trees of the same order within two binomial heaps. Due to the structure of binomial trees, they can be merged trivially. As their root node is the smallest element within the tree, by comparing the two keys, the smaller of them is the minimum key, and becomes the new root node. Then the other tree becomes a subtree of the combined tree. This operation is basic to the complete merging of two binomial heaps.

```

function mergeTree(p, q)
    if p.root.key <= q.root.key
        return p.addSubTree(q)
    else

```

```
return q.addSubTree(p)
```



The operation of **merging** two heaps is perhaps the most interesting and can be used as a subroutine in most other operations. The lists of roots of both heaps are traversed simultaneously, similarly as in the merge algorithm.

If only one of the heaps contains a tree of order  $j$ , this tree is moved to the merged heap. If both heaps contain a tree of order  $j$ , the two trees are merged to one tree of order  $j+1$  so that the minimum-heap property is satisfied. Note that it may later be necessary to merge this tree with some other tree of order  $j+1$  present in one of the heaps. In the course of the algorithm, we need to examine at most three trees of any order (two from the two heaps we merge and one composed of two smaller trees).

Because each binomial tree in a binomial heap corresponds to a bit in the binary representation of its size, there is an analogy between the merging of two heaps and the binary addition of the *sizes* of the two heaps, from right-to-left. Whenever a carry occurs during addition, this corresponds to a merging of two binomial trees during the merge.

Each tree has order at most  $\log n$  and therefore the running time is  $O(\log n)$ .

```
function merge(p, q)
  while not ( p.end() and q.end() )
    tree = mergeTree(p.currentTree(), q.currentTree())
    if not heap.currentTree().empty()
      tree = mergeTree(tree, heap.currentTree())
      heap.addTree(tree)
    else
      heap.addTree(tree)
  heap.next() p.next() q.next()
```

## Insert

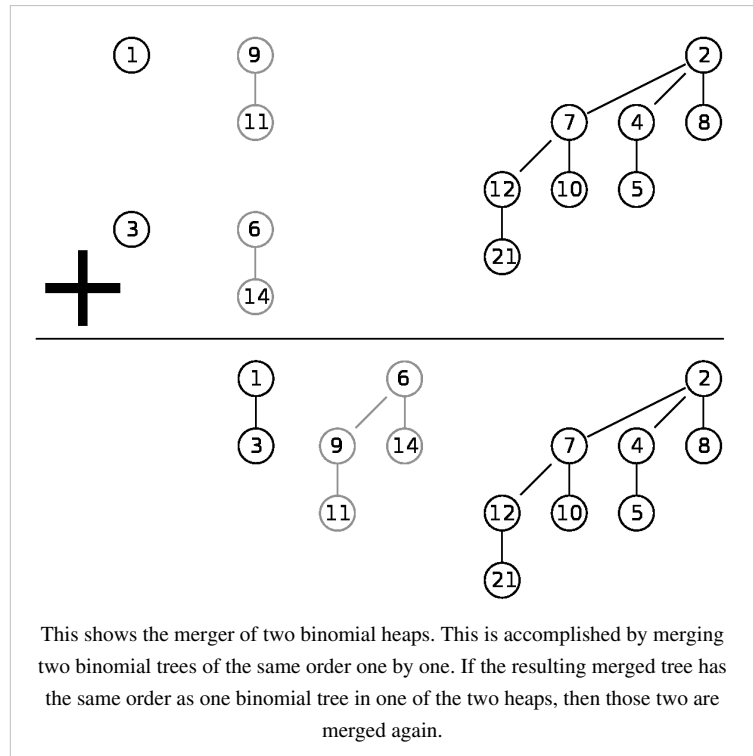
**Inserting** a new element to a heap can be done by simply creating a new heap containing only this element and then merging it with the original heap. Due to the merge, insert takes  $O(\log n)$  time, however it has an *amortized* time of  $O(1)$  (i.e. constant).

## Find minimum

To find the **minimum** element of the heap, find the minimum among the roots of the binomial trees. This can again be done easily in  $O(\log n)$  time, as there are just  $O(\log n)$  trees and hence roots to examine.

By using a pointer to the binomial tree that contains the minimum element, the time for this operation can be reduced to  $O(1)$ . The

pointer must be updated when performing any operation other than Find minimum. This can be done in  $O(\log n)$  without raising the running time of any operation.



## Delete minimum

To **delete the minimum element** from the heap, first find this element, remove it from its binomial tree, and obtain a list of its subtrees. Then transform this list of subtrees into a separate binomial heap by reordering them from smallest to largest order. Then merge this heap with the original heap. Since each tree has at most  $\log n$  children, creating this new heap is  $O(\log n)$ . Merging heaps is  $O(\log n)$ , so the entire delete minimum operation is  $O(\log n)$ .

```
function deleteMin(heap)
    min = heap.trees().first()
    for each current in heap.trees()
        if current.root < min then min = current
    for each tree in min.subTrees()
        tmp.addTree(tree)
    heap.removeTree(min)
    merge(heap, tmp)
```

## Decrease key

After **decreasing** the key of an element, it may become smaller than the key of its parent, violating the minimum-heap property. If this is the case, exchange the element with its parent, and possibly also with its grandparent, and so on, until the minimum-heap property is no longer violated. Each binomial tree has height at most  $\log n$ , so this takes  $O(\log n)$  time.

## Delete

To **delete** an element from the heap, decrease its key to negative infinity (that is, some value lower than any element in the heap) and then delete the minimum in the heap.

## Performance

All of the following operations work in  $O(\log n)$  time on a binomial heap with  $n$  elements:

- Insert a new element to the heap
- Find the element with minimum key
- Delete the element with minimum key from the heap
- Decrease key of a given element
- Delete given element from the heap
- Merge two given heaps to one heap

Finding the element with minimum key can also be done in  $O(1)$  by using an additional pointer to the minimum.

## Applications

- Discrete event simulation
- Priority queues

## References

- Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7. Chapter 19: Binomial Heaps, pp.455–475.
- Vuillemin, J. (1978). A data structure for manipulating priority queues.<sup>[1]</sup> *Communications of the ACM* **21**, 309–314.

## External links

- Java applet simulation of binomial heap<sup>[2]</sup>
  - Python implementation of binomial heap<sup>[3]</sup>
  - Two C implementations of binomial heap<sup>[4]</sup> (a generic one and one optimized for integer keys)
  - Haskell implementation of binomial heap<sup>[5]</sup>
  - Common Lisp implementation of binomial heap<sup>[6]</sup>
-



## References

- [1] <http://portal.acm.org/citation.cfm?id=359478>
- [2] <http://www.cs.yorku.ca/~aaw/Sotirios/BinomialHeap.html>
- [3] <http://aspn.activestate.com/ASPN/Cookbook/Python/Recipe/511508>
- [4] [http://www.cs.unc.edu/~bbb/#binomial\\_heaps](http://www.cs.unc.edu/~bbb/#binomial_heaps)
- [5] <http://hackage.haskell.org/packages/archive/TreeStructures/latest/doc/html/src/Data-Heap-Binomial.html>
- [6] <https://github.com/vy/binomial-heap>

## Fibonacci heap

In computer science, a **Fibonacci heap** is a heap data structure consisting of a collection of trees. It has a better amortized running time than a binomial heap. Fibonacci heaps were developed by Michael L. Fredman and Robert E. Tarjan in 1984 and first published in a scientific journal in 1987. The name of Fibonacci heap comes from Fibonacci numbers which are used in the running time analysis.

Find-minimum is  $O(1)$  amortized time.<sup>[1]</sup> Operations insert, decrease key, and merge (union) work in constant amortized time.<sup>[2]</sup> Operations delete and delete minimum work in  $O(\log n)$  amortized time<sup>[2]</sup>. This means that starting from an empty data structure, any sequence of  $a$  operations from the first group and  $b$  operations from the second group would take  $O(a + b \log n)$  time. In a binomial heap such a sequence of operations would take  $O((a + b) \log(n))$  time. A Fibonacci heap is thus better than a binomial heap when  $b$  is asymptotically smaller than  $a$ .

Using Fibonacci heaps for priority queues improves the asymptotic running time of important algorithms, such as Dijkstra's algorithm for computing the shortest path between two nodes in a graph.

## Structure

A Fibonacci heap is a collection of trees satisfying the minimum-heap property, that is, the key of a child is always greater than or equal to the key of the parent. This implies that the minimum key is always at the root of one of the trees. Compared with binomial heaps, the structure of a Fibonacci heap is more flexible. The trees do not have a prescribed shape and in the extreme case the heap can have every element in a separate tree. This flexibility allows some operations to be executed in a "lazy" manner, postponing the work for later operations. For example merging heaps is done simply by concatenating the two lists of trees, and operation *decrease key* sometimes cuts a node from its parent and forms a new tree.

However at some point some order needs to be introduced to the heap to achieve the desired running time. In particular, degrees of nodes (here degree means the number of children) are kept quite low: every node has degree at most  $O(\log n)$  and the size of a subtree rooted in a node of degree  $k$  is at least  $F_{k+2}$ , where  $F_k$  is the  $k$ th Fibonacci number. This is achieved by the rule that we can cut at most one child of each non-root node. When a second child is cut, the node itself needs to be cut from its parent and becomes the root of a new tree (see Proof of degree bounds, below). The number of trees is decreased in the operation *delete minimum*, where trees are linked together.

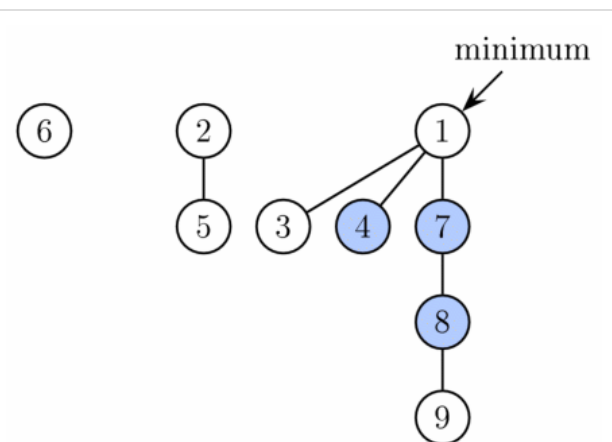


Figure 1. Example of a Fibonacci heap. It has three trees of degrees 0, 1 and 3. Three vertices are marked (shown in blue). Therefore the potential of the heap is 9.

As a result of a relaxed structure, some operations can take a long time while others are done very quickly. In the amortized running time analysis we pretend that very fast operations take a little bit longer than they actually do. This additional time is then later subtracted from the actual running time of slow operations. The amount of time saved for later use is measured at any given moment by a potential function. The potential of a Fibonacci heap is given by

$$\text{Potential} = t + 2m$$

where  $t$  is the number of trees in the Fibonacci heap, and  $m$  is the number of marked nodes. A node is marked if at least one of its children was cut since this node was made a child of another node (all roots are unmarked).

Thus, the root of each tree in a heap has one unit of time stored. This unit of time can be used later to link this tree with another tree at amortized time 0. Also, each marked node has two units of time stored. One can be used to cut the node from its parent. If this happens, the node becomes a root and the second unit of time will remain stored in it as in any other root.

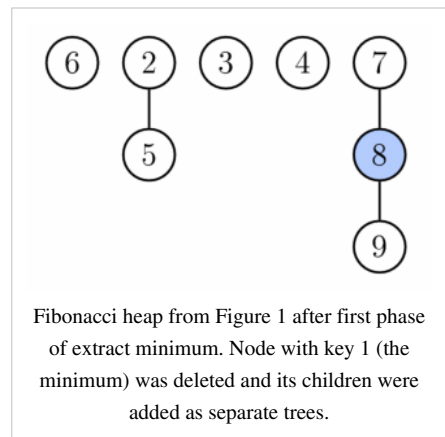
## Implementation of operations

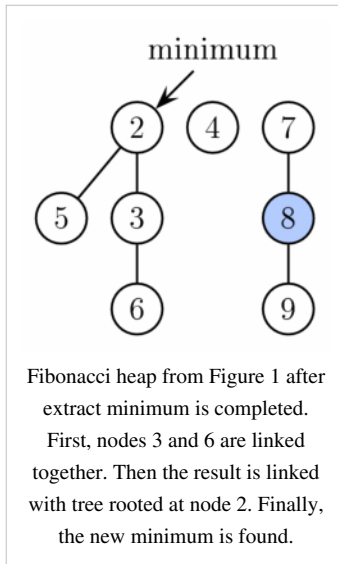
To allow fast deletion and concatenation, the roots of all trees are linked using a circular, doubly linked list. The children of each node are also linked using such a list. For each node, we maintain its number of children and whether the node is marked. Moreover we maintain a pointer to the root containing the minimum key.

Operation **find minimum** is now trivial because we keep the pointer to the node containing it. It does not change the potential of the heap, therefore both actual and amortized cost is constant. As mentioned above, **merge** is implemented simply by concatenating the lists of tree roots of the two heaps. This can be done in constant time and the potential does not change, leading again to constant amortized time.

Operation **insert** works by creating a new heap with one element and doing merge. This takes constant time, and the potential increases by one, because the number of trees increases. The amortized cost is thus still constant.

Operation **extract minimum** (same as *delete minimum*) operates in three phases. First we take the root containing the minimum element and remove it. Its children will become roots of new trees. If the number of children was  $d$ , it takes time  $O(d)$  to process all new roots and the potential increases by  $d-1$ . Therefore the amortized running time of this phase is  $O(d) = O(\log n)$ .



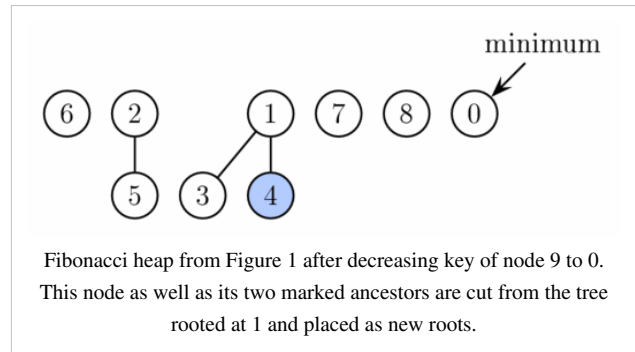


However to complete the extract minimum operation, we need to update the pointer to the root with minimum key. Unfortunately there may be up to  $n$  roots we need to check. In the second phase we therefore decrease the number of roots by successively linking together roots of the same degree. When two roots  $u$  and  $v$  have the same degree, we make one of them a child of the other so that the one with the smaller key remains the root. Its degree will increase by one. This is repeated until every root has a different degree. To find trees of the same degree efficiently we use an array of length  $O(\log n)$  in which we keep a pointer to one root of each degree. When a second root is found of the same degree, the two are linked and the array is updated. The actual running time is  $O(\log n + m)$  where  $m$  is the number of roots at the beginning of the second phase. At the end we will have at most  $O(\log n)$  roots (because each has a different degree). Therefore the difference in the potential function from before this phase to after it is:  $O(\log n) - m$ , and the amortized running time is then at most  $O(\log n + m) + O(\log n) - m = O(\log n)$ . Since we can scale up the units of potential stored at insertion in each

node by the constant factor in the  $O(m)$  part of the actual cost for this phase.

In the third phase we check each of the remaining roots and find the minimum. This takes  $O(\log n)$  time and the potential does not change. The overall amortized running time of extract minimum is therefore  $O(\log n)$ .

Operation **decrease key** will take the node, decrease the key and if the heap property becomes violated (the new key is smaller than the key of the parent), the node is cut from its parent. If the parent is not a root, it is marked. If it has been marked already, it is cut as well and its parent is marked. We continue upwards until we reach either the root or an unmarked node. In the process we create some number, say  $k$ , of new trees. Each of these new trees except possibly the first one was marked originally but as a root it will become unmarked. One node can become marked. Therefore the potential decreases by at least  $k - 2$ . The actual time to perform the cutting was  $O(k)$ , therefore the amortized running time is constant.



Finally, operation **delete** can be implemented simply by decreasing the key of the element to be deleted to minus infinity, thus turning it into the minimum of the whole heap. Then we call extract minimum to remove it. The amortized running time of this operation is  $O(\log n)$ .

## Proof of degree bounds

The amortized performance of a Fibonacci heap depends on the degree (number of children) of any tree root being  $O(\log n)$ , where  $n$  is the size of the heap. Here we show that the size of the (sub)tree rooted at any node  $x$  of degree  $d$  in the heap must have size at least  $F_{d+2}$ , where  $F_k$  is the  $k$ th Fibonacci number. The degree bound follows from this and the fact (easily proved by induction) that  $F_{d+2} \geq \varphi^d$  for all integers  $d \geq 0$ , where  $\varphi = (1 + \sqrt{5})/2 \doteq 1.618$ . (We then have  $n \geq F_{d+2} \geq \varphi^d$ , and taking the log to base  $\varphi$  of both sides gives  $d \leq \log_\varphi n$  as required.)

Consider any node  $x$  somewhere in the heap ( $x$  need not be the root of one of the main trees). Define **size**( $x$ ) to be the size of the tree rooted at  $x$  (the number of descendants of  $x$ , including  $x$  itself). We prove by induction on the height of  $x$  (the length of a longest simple path from  $x$  to a descendant leaf), that **size**( $x$ )  $\geq F_{d+2}$ , where  $d$  is the degree of  $x$ .

**Base case:** If  $x$  has height 0, then  $d = 0$ , and **size**( $x$ ) = 1 =  $F_2$ .

**Inductive case:** Suppose  $x$  has positive height and degree  $d > 0$ . Let  $y_1, y_2, \dots, y_d$  be the children of  $x$ , indexed in order of the times they were most recently made children of  $x$  ( $y_1$  being the earliest and  $y_d$  the latest), and let  $c_1, c_2, \dots, c_d$  be their respective degrees. We **claim** that  $c_i \geq i-2$  for each  $i$  with  $2 \leq i \leq d$ : Just before  $y_i$  was made a child of  $x$ ,  $y_1, \dots, y_{i-1}$  were already children of  $x$ , and so  $x$  had degree at least  $i-1$  at that time. Since trees are combined only when the degrees of their roots are equal, it must have been that  $y_i$  also had degree at least  $i-1$  at the time it became a child of  $x$ . From that time to the present,  $y_i$  can only have lost at most one child (as guaranteed by the marking process), and so its current degree  $c_i$  is at least  $i-2$ . This proves the **claim**.

Since the heights of all the  $y_i$  are strictly less than that of  $x$ , we can apply the inductive hypothesis to them to get  $\text{size}(y_i) \geq F_{c_i+2} \geq F_{(i-2)+2} = F_i$ . The nodes  $x$  and  $y_1$  each contribute at least 1 to  $\text{size}(x)$ , and so we have

$$\text{size}(x) \geq 2 + \sum_{i=2}^d \text{size}(y_i) \geq 2 + \sum_{i=2}^d F_i = 1 + \sum_{i=0}^d F_i.$$

A routine induction proves that  $1 + \sum_{i=0}^d F_i = F_{d+2}$  for any  $d \geq 0$ , which gives the desired lower bound on  $\text{size}(x)$ .

## Worst case

Although the total running time of a sequence of operations starting with an empty structure is bounded by the bounds given above, some (very few) operations in the sequence can take very long to complete (in particular delete and delete minimum have linear running time in the worst case). For this reason Fibonacci heaps and other amortized data structures may not be appropriate for real-time systems. It is possible to create a data structure which has the same worst case performance as the Fibonacci heap has amortized performance.<sup>[3]</sup> However the resulting structure is very complicated, so it is not useful in most practical cases.

## Summary of running times

Common Operations	Effect	Unsorted Linked List	Self-balancing binary search tree	Binary heap	Binomial heap	Fibonacci heap	Brodal queue <sup>[3]</sup>	Pairing heap
insert(data, key)	Adds <i>data</i> to the queue, tagged with <i>key</i>	O(1)	O(log n)	O(log n)	O(log n)	O(1)	O(1)	O(1) <sup>[2]</sup>
findMin() -> key, data	Returns <i>key, data</i> corresponding to min-value <i>key</i>	O(n)	O(log n) or O(1) (**)	O(1)	O(log n) <sup>[4]</sup>	O(1) <sup>[1]</sup>	O(1)	O(1) <sup>[2]</sup>
deleteMin()	Deletes <i>data</i> corresponding to min-value <i>key</i>	O(n)	O(log n)	O(log n)	O(log n)	O(log n)*	O(log n)	O(log n)* <sup>[2]</sup>
delete(node)	Deletes <i>data</i> corresponding to given <i>key</i> , given a pointer to the node being deleted	O(1)	O(log n)	O(log n)	O(log n)	O(log n)*	O(log n)	O(log n)* <sup>[2]</sup>

decreaseKey(node)	Decreases the <i>key</i> of a node, given a pointer to the node being modified	$O(1)$	$O(\log n)$	$O(\log n)$	$O(\log n)$	$O(1)^*$	$O(1)$	Unknown but bounded: $\Omega(\log \log n), 2^{O(\sqrt{\log \log n})}$ *[2]
merge(heap1, heap2) → heap3	Merges two heaps into a third	$O(1)$	$O(m \log(n+m))$	$O(m + n)$	$O(\log n)^{***}$	$O(1)$	$O(1)$	$O(1)^{[2]}$

(\*) Amortized time

(\*\*) With trivial modification to store an additional pointer to the minimum element

(\*\*\*) Where  $n$  is the size of the larger heap

## References

- [1] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7. Chapter 20: Fibonacci Heaps, pp.476–497. Third edition p518.
- [2] Fredman, Michael Lawrence; Tarjan, Robert E. (1987). "Fibonacci heaps and their uses in improved network optimization algorithms" ([http://www.cl.cam.ac.uk/~sos22/supervise/dsaa/fib\\_heaps.pdf](http://www.cl.cam.ac.uk/~sos22/supervise/dsaa/fib_heaps.pdf)) (PDF). *Journal of the Association for Computing Machinery* **34** (3): 596–615. doi:10.1145/28869.28874. .
- [3] Gerth Stølting Brodal (1996), "Worst-Case Efficient Priority Queues", *Proc. 7th ACM-SIAM Symposium on Discrete Algorithms* (Society for Industrial and Applied Mathematics): 52–58, doi:10.1145/313852.313883, ISBN 0-89871-366-8, CiteSeerX: 10.1.1.43.8133 (<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.43.8133>)
- [4] Cormen, Thomas H.; Leiserson, Charles E., Rivest, Ronald L. (1990). *Introduction to Algorithms* (1st ed.). MIT Press and McGraw-Hill. ISBN 0-262-03141-8.

## External links

- Java applet simulation of a Fibonacci heap (<http://www.cs.yorku.ca/~aaw/Jason/FibonacciHeapAnimation.html>)
- C implementation of Fibonacci heap ([http://resnet.uoregon.edu/~gurney\\_j/jmpc/fib.html](http://resnet.uoregon.edu/~gurney_j/jmpc/fib.html))
- MATLAB implementation of Fibonacci heap (<http://www.mathworks.com/matlabcentral/fileexchange/30072-fibonacci-heap>)
- De-recursive and memory efficient C implementation of Fibonacci heap (<http://www.labri.fr/perso/pelegri/code/#fibonacci>) (free/libre software, CeCILL-B license ([http://www.cecill.info/licences/Licence\\_CeCILL-B\\_V1-en.html](http://www.cecill.info/licences/Licence_CeCILL-B_V1-en.html)))
- C++ template Fibonacci heap, with demonstration (<http://ideone.com/9jYnv>)
- Ruby implementation of the Fibonacci heap (with tests) ([http://github.com/evansenter/f\\_heap](http://github.com/evansenter/f_heap))
- Pseudocode of the Fibonacci heap algorithm (<http://www.cs.princeton.edu/~wayne/cs423/fibonacci/FibonacciHeapAlgorithm.html>)
- Efficient C++ Fibonacci heap (<http://stackoverflow.com/q/504823/194609>)
- Various Java Implementations for Fibonacci heap (<http://stackoverflow.com/q/6273833/194609>)

## 2-3 heap

---

In computer science, a **2-3 heap** is a data structure, a variation on the heap, designed by Tadao Takaoka in 1999. The structure is similar to the Fibonacci heap, and borrows from the 2-3 tree.

Time costs for some common heap operations:

- *delete-min* takes  $O(\log(n))$  amortized time
- *decrease-key* takes constant amortized time
- *insertion* takes constant amortized time.

### References

- Tadao Takaoka. *Theory of 2-3 Heaps* <sup>[1]</sup>, Cocoon (1999).

### References

- [1] <http://www.cosc.canterbury.ac.nz/~tad/2-3heaps.pdf>

## Pairing heap

---

A **pairing heap** is a type of heap data structure with relatively simple implementation and excellent practical amortized performance. However, it has proven very difficult to determine the precise asymptotic running time of pairing heaps.

Pairing heaps are heap ordered multiway trees. Describing the various heap operations is relatively simple (in the following we assume a min-heap):

- *find-min*: simply return the top element of the heap.
- *merge*: compare the two root elements, the smaller remains the root of the result, the larger element and its subtree is appended as a child of this root.
- *insert*: create a new heap for the inserted element and *merge* into the original heap.
- *decrease-key* (optional): remove the subtree rooted at the key to be decreased then *merge* it with the heap.
- *delete-min*: remove the root and *merge* its subtrees. Various strategies are employed.

The amortized time per *delete-min* is  $O(\log n)$ .<sup>[1]</sup> The operations *find-min*, *merge*, and *insert* are  $O(1)$  <sup>[2]</sup> and *decrease-key* takes  $2^{O(\sqrt{\log \log n})}$  amortized time.<sup>[3]</sup> Fredman proved that the amortized time per *decrease-key* is at least  $\Omega(\log \log n)$ .<sup>[4]</sup>

Although this is worse than other priority queue algorithms such as Fibonacci heaps, which perform *decrease-key* in  $O(1)$  amortized time, the performance in practice is excellent. Stasko and Vitter<sup>[5]</sup> and Moret and Shapiro<sup>[6]</sup> conducted experiments on pairing heaps and other heap data structures. They concluded that the pairing heap is as fast as, and often faster than, other efficient data structures like the binary heaps.

## Implementation

A pairing heap is either an empty heap, or a pair consisting of a root element and a possibly empty list of pairing heaps. The heap ordering property requires that all the root elements of the subheaps in the list are not smaller than the root element of the heap. The following description assumes a purely functional heap that does not support the *decrease-key* operation.

```
type PairingHeap[Elem] = Empty Heap(elem: Elem, subheaps: List[PairingHeap[Elem])
```

## Operations

### *find-min*

The function *find-min* simply returns the root element of the heap:

```
function find-min(heap)
  if heap == Empty
    error
  else
    return heap.elem
```

### *merge*

Merging with an empty heap returns the other heap, otherwise a new heap is returned that has the minimum of the two root elements as its root element and just adds the heap with the larger root to the list of subheaps:

```
function merge(heap1, heap2)
  if heap1 == Empty
    return heap2
  elseif heap2 == Empty
    return heap1
  elseif heap1.elem < heap2.elem
    return Heap(heap1.elem, heap2 :: heap1.subheaps)
  else
    return Heap(heap2.elem, heap1 :: heap2.subheaps)
```

### *insert*

The easiest way to insert an element into a heap is to merge the heap with a new heap containing just this element and an empty list of subheaps:

```
function insert(elem, heap)
  return merge(Heap(elem, []), heap)
```

### *delete-min*

The only non-trivial fundamental operation is the deletion of the minimum element from the heap. The standard strategy first merges the subheaps in pairs (this is the step that gave this datastructure its name) from left to right and then merges the resulting list of heaps from right to left:

```
function delete-min(heap)
  if heap == Empty
    error
  else
```

```
return merge-pairs(heap.subheaps)
```

This uses the auxiliary function *merge-pairs*:

```
function merge-pairs(l)
  if length(l) == 0
    return Empty
  elseif length(l) == 1
    return l[0]
  else
    return merge(merge(l[0], l[1]), merge-pairs(l[2.. ]))
```

That this does indeed implement the described two-pass left-to-right then right-to-left merging strategy can be seen from this reduction:

```
merge-pairs([H1, H2, H3, H4, H5, H6, H7])
=> merge(merge(H1, H2), merge-pairs([H3, H4, H5, H6, H7]))
    # merge H1 and H2 to H12, then the rest of the list
=> merge(H12, merge(merge(H3, H4), merge-pairs([H5, H6, H7])))
    # merge H3 and H4 to H34, then the rest of the list
=> merge(H12, merge(H34, merge(merge(H5, H6), merge-pairs([H7]))))
    # merge H5 and H5 to H56, then the rest of the list
=> merge(H12, merge(H34, merge(H56, H7)))
    # switch direction, merge the last two resulting heaps, giving H567
=> merge(H12, merge(H34, H567))
    # merge the last two resulting heaps, giving H34567
=> merge(H12, H34567)
    # finally, merge the first merged pair with the result of merging the rest
=> H1234567
```

## References

- [1] Fredman, Michael L.; Sedgwick, Robert; Sleator, Daniel D.; Tarjan, Robert E. (1986), "The pairing heap: a new form of self-adjusting heap" (<http://www.lb.cs.cmu.edu/afs/cs.cmu.edu/user/sleator/www/papers/pairing-heaps.pdf>), *Algorithmica* **1** (1): 111–129, doi:10.1007/BF01840439, .
- [2] Iacono, John (2000), "Improved upper bounds for pairing heaps" (<http://john2.poly.edu/papers/swat00/paper.pdf>), *Proc. 7th Scandinavian Workshop on Algorithm Theory*, Lecture Notes in Computer Science, **1851**, Springer-Verlag, pp. 63–77, doi:10.1007/3-540-44985-X\_5, ISBN 978-3-540-67690-4, .
- [3] Pettie, Seth (2005), "Towards a final analysis of pairing heaps" (<http://www.eecs.umich.edu/~pettie/papers/focs05.pdf>), *Proc. 46th Annual IEEE Symposium on Foundations of Computer Science*, pp. 174–183, doi:10.1109/SFCS.2005.75, ISBN 0-7695-2468-0, .
- [4] Fredman, Michael L. (1999), "On the efficiency of pairing heaps and related data structures" (<http://www.wens.uqac.ca/azinflou/Fichiers840/EfficiencyPairingHeap.pdf>), *Journal of the ACM* **46** (4): 473–501, doi:10.1145/320211.320214, .
- [5] Stasko, John T.; Vitter, Jeffrey S. (1987), "Pairing heaps: experiments and analysis", *Communications of the ACM* **30** (3): 234–249, doi:10.1145/214748.214759, CiteSeerX: 10.1.1.106.2988 (<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.106.2988>).
- [6] Moret, Bernard M. E.; Shapiro, Henry D. (1991), "An empirical analysis of algorithms for constructing a minimum spanning tree", *Proc. 2nd Workshop on Algorithms and Data Structures*, Lecture Notes in Computer Science, **519**, Springer-Verlag, pp. 400–411, doi:10.1007/BFb0028279, ISBN 3-540-54343-0, CiteSeerX: 10.1.1.53.5960 (<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.53.5960>).



## External links

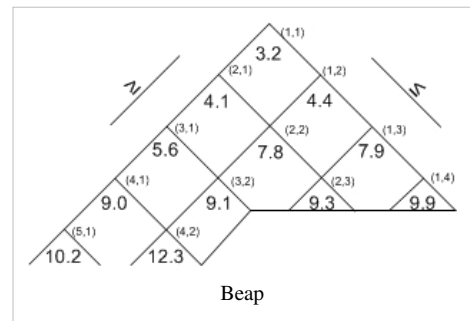
- Louis Wasserman discusses pairing heaps and their implementation in Haskell in The Monad Reader, Issue 16 (<http://themonadreader.files.wordpress.com/2010/05/issue16.pdf>) (pp. 37–52).
- pairing heaps (<http://www.cise.ufl.edu/~sahni/dsaa/enrich/c13/pairing.htm>), Sartaj Sahni
- Amr Elmasry (2009), "Pairing Heaps with  $O(\log \log n)$  decrease Cost" ([http://www.siam.org/proceedings/soda/2009/SODA09\\_052\\_elmasrya.pdf](http://www.siam.org/proceedings/soda/2009/SODA09_052_elmasrya.pdf)), *Proceedings of the twentieth Annual ACM-SIAM Symposium on Discrete Algorithms {SODA '09}* (New York): 471–476
- heaps library (<http://www.swi-prolog.org/pldoc/doc/swi/library/heaps.pl>) in SWI-Prolog, uses pairing heaps
- Open source implementation of pairing heaps in Erlang (<https://gist.github.com/1248317>)

## Beap

**Beap**, or **bi-parental heap**, is a data structure where a node usually has two parents (unless it is the first or last on a level) and two children (unless it is on the last level). Unlike a heap, a beap allows sublinear search. The beap was introduced by Ian Munro and Hendra Suwanda. A related data structure is the Young tableau.

## Performance

The height of the structure is approximately  $\sqrt{n}$ . Also, assuming the last level is full, the number of elements on that level is also  $\sqrt{n}$ . In fact, because of these properties all basic operations (insert, remove, find) run in  $O(\sqrt{n})$  time on average. Find operations in the heap can be  $O(n)$  in the worst case. Removal and insertion of new elements involves propagation of elements up or down (much like in a heap) in order to restore the beap invariant. An additional perk is that beap provides constant time access to the smallest element and  $O(\sqrt{n})$  time for the maximum element. Actually, a  $O(\sqrt{n})$  find operation can be implemented if parent pointers at each node are maintained. You would start at the absolute bottom-most element of the top node (similar to the left-most child in a heap) and move either up or right to find the element of interest.



## References

- J. Ian Munro and Hendra Suwanda. "Implicit data structures for fast search and update". *Journal of Computer and System Sciences*, 21(2):236250, 1980.
- J.W.J Williams in Algorithms 232, "Heapsort", *Comm. ACM* 7 (June 1964), 347-348

# Leftist tree

In computer science, a **leftist tree** or **leftist heap** is a priority queue implemented with a variant of a binary heap. Every node has an *s-value* which is the distance to the nearest leaf. In contrast to a *binary heap*, a leftist tree attempts to be very unbalanced. In addition to the heap property, leftist trees are maintained so the right descendant of each node has the lower s-value.

The leftist tree was invented by Clark Allan Crane. The name comes from the fact that the left subtree is usually taller than the right subtree.

When inserting a new node into a tree, a new one-node tree is created and merged into the existing tree. To delete a minimum item, we remove the root and the left and right sub-trees are then merged. Both these operations take  $O(\log n)$  time. For insertions, this is slower than binary heaps which support insertion in amortized constant time,  $O(1)$  and  $O(\log n)$  worst-case.

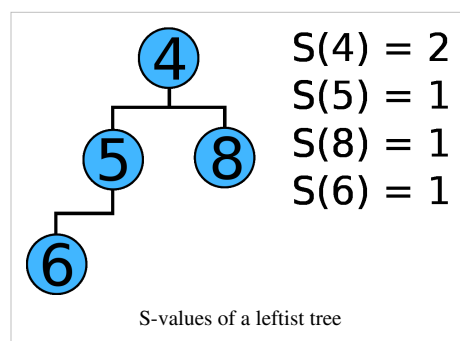
Leftist trees are advantageous because of their ability to merge quickly, compared to binary heaps which take  $\Theta(n)$ . In almost all cases, skew heaps have better performance.

## Bias

The usual leftist tree is a *height-biased* leftist tree. However, other biases can exist, such as in the *weight-biased* leftist tree.

## S-value

The s-value of a node is the distance from that node to the nearest leaf of the extended binary representation of the tree <sup>[1]</sup>. The extended representation (not shown) fills out the tree so that each node has 2 children (adding a total of 5 leaves here). The minimum distance to these leaves are marked in the diagram. Thus s-value of 4 is 2, since the closest leaf is that of 8 --if 8 were extended. The s-value of 5 is 1 since its extended representation would have one leaf itself.



## Merging height biased leftist trees

Merging two nodes together depends on whether the tree is a min or max height biased leftist tree. For a min height biased leftist tree, set the higher valued node as its right child of the lower valued node. If the lower valued node already has a right child, then merge the higher valued node with the sub-tree rooted by the right child of the lower valued node.

After merging, the s-value of the lower valued node must be updated (see above section, s-value). Now check if the lower valued node has a left child. If it does not, then move the right child to the left. If it does have a left child, then the child with the highest s-value should go on the left.

## Java code for merging a min height biased leftist tree

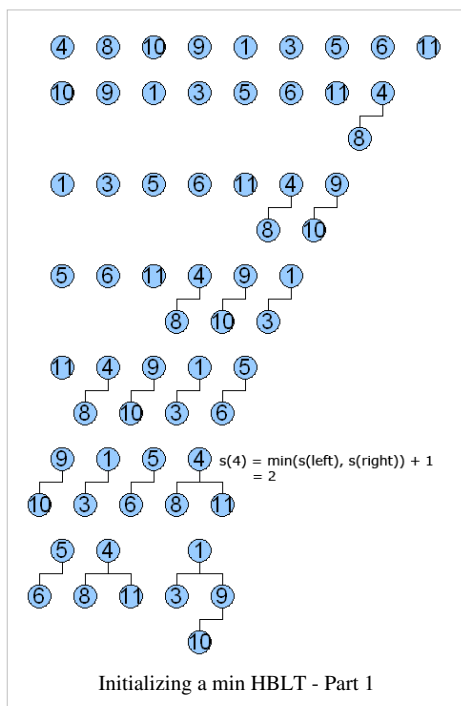
```
public Node merge(Node x, Node y) {
    if (x == null)
        return y;
    if (y == null)
        return x;

    // if this was a max height biased leftist tree, then the
    // next line would be: if(x.element < y.element)
    if (x.element.compareTo(y.element) > 0) {
        // x.element > y.element
        Node temp = x;
        x = y;
        y = temp;
    }

    x.rightChild = merge(x.rightChild, y);

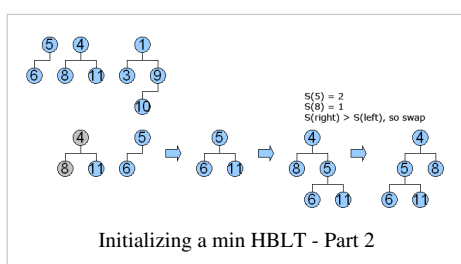
    if (x.leftChild == null) {
        // left child doesn't exist, so move right child to the left side
        x.leftChild = x.rightChild;
        x.rightChild = null;
        x.s = 1;
    } else {
        // left child does exist, so compare s-values
        if (x.leftChild.s < x.rightChild.s) {
            Node temp = x.leftChild;
            x.leftChild = x.rightChild;
            x.rightChild = temp;
        }
        // since we know the right child has the lower s-value, we can just
        // add one to its s-value
        x.s = x.rightChild.s + 1;
    }
    return x;
}
```

## Initializing a height biased leftist tree

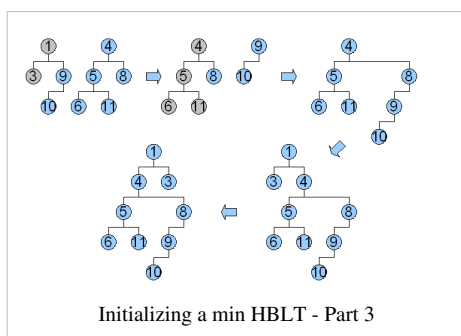


Initializing a height biased leftist tree is primarily done in one of two ways. The first is to merge each node one at a time into one HBLT. This process is inefficient and takes  $O(n \log n)$  time. The other approach is to use a queue to store each node and resulting tree. The first two items in the queue are removed, merged, and placed back into the queue. This can initialize a HBLT in  $O(n)$  time. This approach is detailed in the three diagrams supplied. A min height biased leftist tree is shown.

To initialize a min HBLT, place each element to be added to the tree into a queue. In the example (see Part 1 to the left), the set of numbers [4, 8, 10, 9, 1, 3, 5, 6, 11] are initialized. Each line of the diagram represents another cycle of the algorithm, depicting the contents of the queue. The first five steps are easy to follow. Notice that the freshly created HBLT is added to the end of the queue. In the fifth step, the first occurrence of an s-value greater than 1 occurs. The sixth step shows two trees merged with each other, with predictable results.



In part 2 a slightly more complex merge happens. The tree with the lower value (tree x) has a right child, so merge must be called again on the subtree rooted by tree x's right child and the other tree. After the merge with the subtree, the resulting tree is put back into tree x. The s-value of the right child (s=2) is now greater than the s-value of the left child (s=1), so they must be swapped. The s-value of the root node 4 is also now 2.



Part 3 is the most complex. Here, we recursively call merge twice (each time with the right child's subtree that is not grayed out). This uses the same process described for part 2.

## External links

- Leftist Trees <sup>[2]</sup>, Sartaj Sahni

## References

- [1] <http://mathworld.wolfram.com/ExtendedBinaryTree.html>
- [2] <http://www.cise.ufl.edu/~sahni/cop5536/slides/lec114.pdf>

# Skew heap

A **skew heap** (or **self-adjusting heap**) is a heap data structure implemented as a binary tree. Skew heaps are advantageous because of their ability to merge more quickly than binary heaps. In contrast with binary heaps, there are no structural constraints, so there is no guarantee that the height of the tree is logarithmic. Only two conditions must be satisfied:

- The general heap order must be enforced
- Every operation (add, remove\_min, merge) on two skew heaps must be done using a special *skew heap merge*.

A skew heap is a self-adjusting form of a leftist heap which attempts to maintain balance by unconditionally swapping all nodes in the merge path when merging two heaps. (The merge operation is also used when adding and removing values.)

With no structural constraints, it may seem that a skew heap would be horribly inefficient. However, amortized complexity analysis can be used to demonstrate that all operations on a skew heap can be done in  $O(\log n)$ .<sup>[1]</sup>

## Definition

Skew heaps may be described with the following recursive definition:

- A heap with only one element is a skew heap.
- The result of *skew merging* two skew heaps  $sh_1$  and  $sh_2$  is also a skew heap.

## Operations

### Merging two heaps

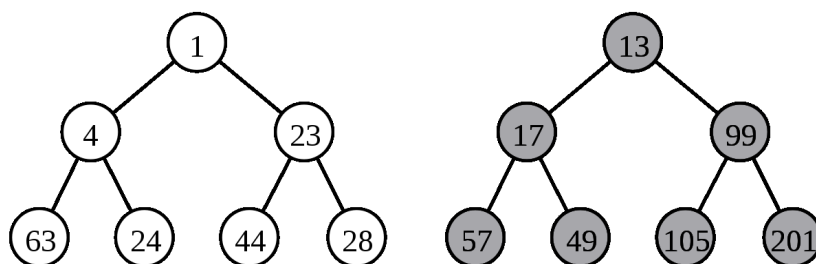
When two skew heaps are to be merged, we can use a similar process as the merge of two leftist heaps:

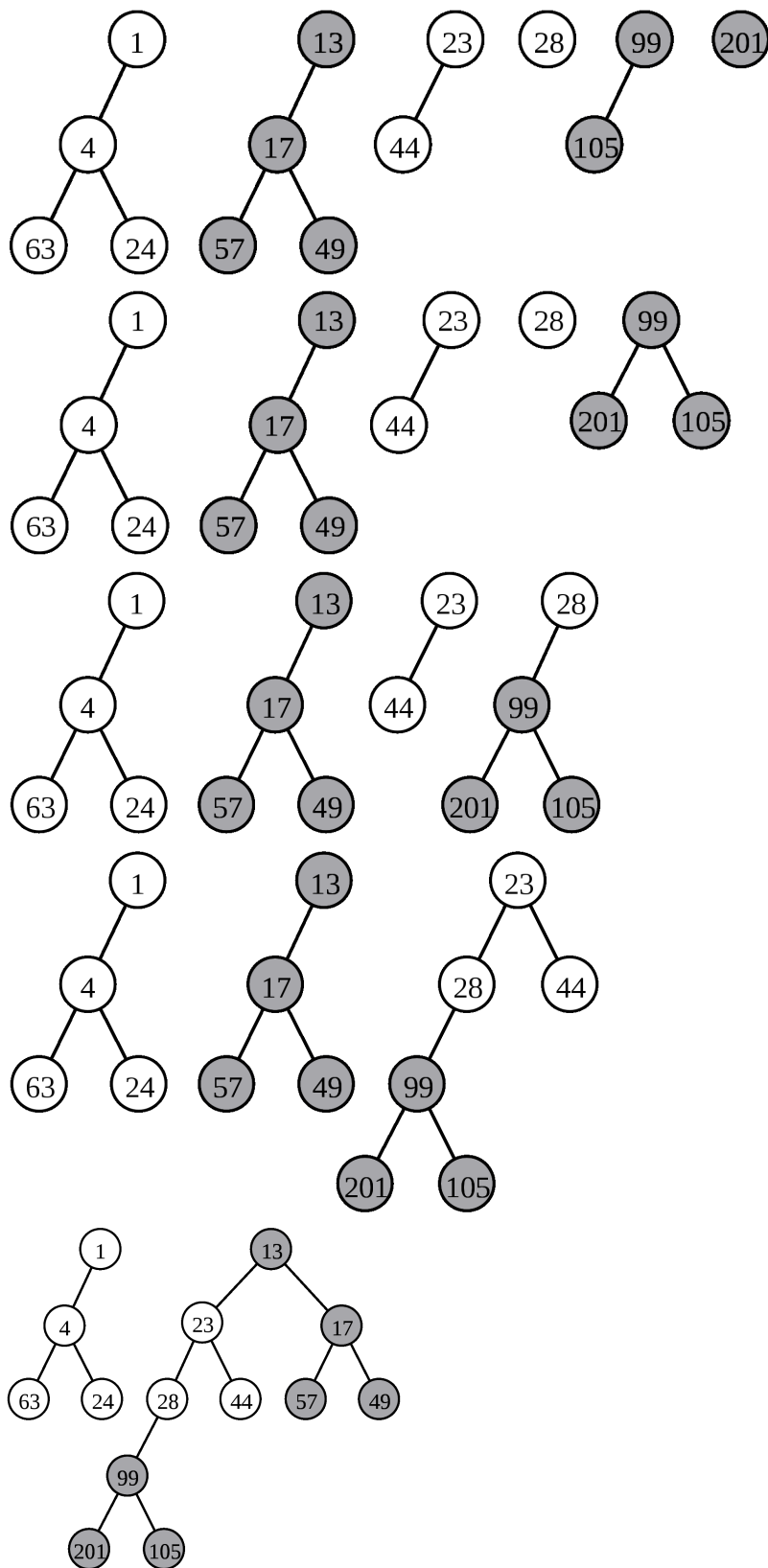
- Compare roots of two heaps; let  $p$  be the heap with the smaller root, and  $q$  be the other heap. Let  $r$  be the name of the resulting new heap.
- Let the root of  $r$  be the root of  $p$  (the smaller root), and let  $r$ 's right subtree be  $p$ 's left subtree.
- Now, compute  $r$ 's left subtree by recursively merging  $p$ 's right subtree with  $q$ .

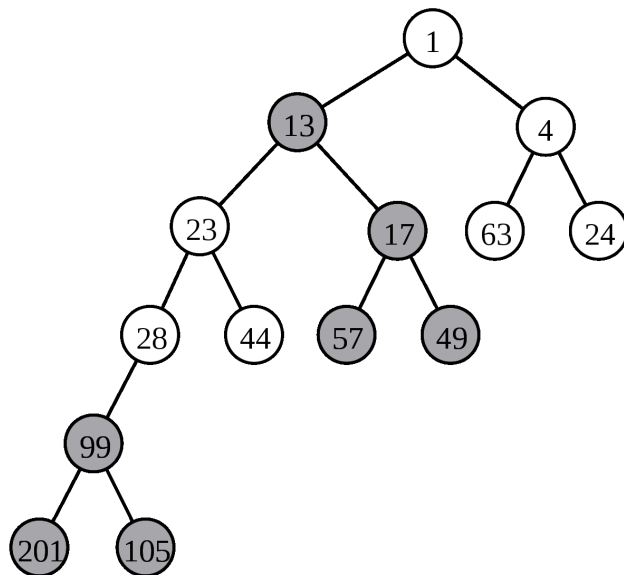
### Non-recursive merging

Alternatively, there is a non-recursive approach which is more wordy, and does require some sorting at the outset.

- Split each heap into subtrees by cutting every rightmost path. (From the root node, sever the right node and make the right child its own subtree.) This will result in a set of trees in which the root either only has a left child or no children at all.
- Sort the subtrees in ascending order based on the value of the root node of each subtree.
- While there are still multiple subtrees, iteratively recombine the last two (from right to left).
  - If the root of the second-to-last subtree has a left child, swap it to be the right child.
  - Link the root of the last subtree as the left child of the second-to-last subtree.







### Adding values

Adding a value to a skew heap is like merging a tree with one node together with the original tree.

### Removing values

Removing the first value in a heap can be accomplished by removing the root and merging the child subtrees. and child subtrees are the part of root.

### Implementation

In many functional languages, skew heaps become extremely simple to implement. Here is a complete sample implementation in Haskell.

```

data SkewHeap a = Empty
                | Node a (SkewHeap a) (SkewHeap a)

singleton :: Ord a => a -> SkewHeap a
singleton x = Node x Empty Empty

union :: Ord a => SkewHeap a -> SkewHeap a -> SkewHeap a
Empty      `union` t2      = t2
t1         `union` Empty   = t1
t1@(Node x1 l1 r1) `union` t2@(Node x2 l2 r2)
  | x1 <= x2      = Node x1 (t2 `union` r1) l1
  | otherwise     = Node x2 (t1 `union` r2)
l2

insert :: Ord a => a -> SkewHeap a -> SkewHeap a
insert x heap = singleton x `union` heap

extractMin :: Ord a => SkewHeap a -> Maybe (a, SkewHeap a)
extractMin Empty      = Nothing
extractMin (Node x l r) = Just (x, l `union` r)

```

## References

- Sleator, Daniel Dominic; Tarjan, Robert Endre (1986). "Self-Adjusting Heaps" <sup>[2]</sup>. *SIAM Journal on Computing* **15** (1): 52–69. doi:10.1137/0215004. ISSN 0097-5397.
- CSE 4101 lecture notes, York University <sup>[3]</sup>

[1] <http://www.cse.yorku.ca/~andy/courses/4101/lecture-notes/LN5.pdf>

[2] <http://www.cs.cmu.edu/~sleator/papers/Adjusting-Heaps.htm>

[3] <http://www.cse.yorku.ca/~andy/courses/4101/lecture-notes/LN5.pdf>

## External links

- Animations comparing leftist heaps and skew heaps, York University (<http://www.cse.yorku.ca/~aaw/Pourhashemi/>)
- Java applet for simulating heaps, Kansas State University (<http://people.cis.ksu.edu/~rhowell/viewer/heapviewer.html>)

# Soft heap

---

In computer science, a **soft heap** is a variant on the simple heap data structure that has constant amortized time for 5 types of operations. This is achieved by carefully "corrupting" (increasing) the keys of at most a certain fixed percentage of values in the heap. The constant time operations are:

- **create**( $S$ ): Create a new soft heap
- **insert**( $S, x$ ): Insert an element into a soft heap
- **meld**( $S, S'$ ): Combine the contents of two soft heaps into one, destroying both
- **delete**( $S, x$ ): Delete an element from a soft heap
- **findmin**( $S$ ): Get the element with minimum key in the soft heap

It was designed by Bernard Chazelle in 2000. The term "corruption" in the structure is the result of what Chazelle called "carpooling" in a soft heap. Each node in the soft heap contains a linked-list of keys and one common key. The common key is an upper bound on the values of the keys in the linked-list. Once a key is added to the linked-list, it is considered corrupted because its value is never again relevant in any of the soft heap operations: only the common keys are compared. It is unpredictable which keys will be corrupted in this manner; it is only known that at most a fixed percentage will be corrupted. This is what makes soft heaps "soft"; you can't be sure whether or not any particular value you put into it will be corrupted. The purpose of these corruptions is effectively to lower the information entropy of the data, enabling the data structure to break through information-theoretic barriers regarding heaps.

Other heaps such as Fibonacci heaps achieve most of these bounds without any corruption, but cannot provide a constant-time bound on the critical *delete* operation. The percentage of values which are corrupted can be chosen freely, but the lower this is set, the more time insertions require ( $O(\log 1/\epsilon)$  for an error rate of  $\epsilon$ ).

---



## Applications

Surprisingly, soft heaps are useful in the design of deterministic algorithms, despite their unpredictable nature. They were used to achieve the best complexity to date for finding a minimum spanning tree. They can also be used to easily build an optimal selection algorithm, as well as *near-sorting* algorithms, which are algorithms that place every element near its final position, a situation in which insertion sort is fast.

One of the simplest examples is the selection algorithm. Say we want to find the  $k$ th largest of a group of  $n$  numbers. First, we choose an error rate of  $1/3$ ; that is, at most 33% of the keys we insert will be corrupted. Now, we insert all  $n$  elements into the heap — at this point, at most  $n/3$  keys are corrupted. Next, we delete the minimum element from the heap about  $n/3$  times. Because this is decreasing the size of the heap, it cannot increase the number of corrupted elements. Thus there are still at most  $n/3$  keys that are corrupted.

Now at least  $2n/3 - n/3 = n/3$  of the remaining keys are not corrupted, so each must be larger than every element we removed. Let  $L$  be the element that we have removed with the largest (actual) value, which is not necessarily the last element that we removed (because the last element we removed could have had its key corrupted, or increased, to a value larger than another element that we have already removed).  $L$  is larger than all the other  $n/3$  elements that we removed and smaller than the remaining  $n/3$  uncorrupted elements in the soft heap. Therefore,  $L$  divides the elements somewhere between 33%/66% and 66%/33%. We then partition the set about  $L$  using the *partition* algorithm from quicksort and apply the same algorithm again to either the set of numbers less than  $L$  or the set of numbers greater than  $L$ , neither of which can exceed  $2n/3$  elements. Since each insertion and deletion requires  $O(1)$  amortized time, the total deterministic time is  $T(n) = T(2n/3) + O(n)$ . Using case 3 of the master theorem (with  $\epsilon=1$  and  $c=2/3$ ), we know that  $T(n) = \Theta(n)$ .

The final algorithm looks like this:

```
function softHeapSelect(a[1..n], k)
    if k = 1 then return minimum(a[1..n])
    create(S)
    for i from 1 to n
        insert(S, a[i])
    for i from 1 to n/3
        x := findmin(S)
        delete(S, x)
    xIndex := partition(a, x) // Returns new index of pivot x
    if k < xIndex
        softHeapSelect(a[1..xIndex-1], k)
    else
        softHeapSelect(a[xIndex..n], k-xIndex+1)
```

## References

- Chazelle, B. 2000. The soft heap: an approximate priority queue with optimal error rate. <sup>[1]</sup> *J. ACM* 47, 6 (Nov. 2000), 1012-1027.
- Kaplan, H. and Zwick, U. 2009. A simpler implementation and analysis of Chazelle's soft heaps. <sup>[2]</sup> In *Proceedings of the Nineteenth Annual ACM -SIAM Symposium on Discrete Algorithms* (New York, New York, January 4—6, 2009). Symposium on Discrete Algorithms. Society for Industrial and Applied Mathematics, Philadelphia, PA, 477-485.

## References

- [1] <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.5.9705>  
 [2] [http://www.siam.org/proceedings/soda/2009/SODA09\\_053\\_kaplanh.pdf](http://www.siam.org/proceedings/soda/2009/SODA09_053_kaplanh.pdf)

## $d$ -ary heap

The  **$d$ -ary heap** or  **$d$ -heap** is a priority queue data structure, a generalization of the binary heap in which the nodes have  $d$  children instead of 2.<sup>[1][2][3]</sup> Thus, a binary heap is a 2-heap. According to Tarjan<sup>[2]</sup> and Jensen et al.,<sup>[4]</sup>  $d$ -ary heaps were invented by Donald B. Johnson in 1975.<sup>[1]</sup>

This data structure allows decrease priority operations to be performed more quickly than binary heaps, at the expense of slower delete minimum operations. This tradeoff leads to better running times for algorithms such as Dijkstra's algorithm in which decrease priority operations are more common than delete min operations.<sup>[1][5]</sup> Additionally,  $d$ -ary heaps have better memory cache behavior than a binary heap, allowing them to run more quickly in practice despite having a theoretically larger worst-case running time.<sup>[6][7]</sup> Like binary heaps,  $d$ -ary heaps are an in-place data structure that uses no additional storage beyond that needed to store the array of items in the heap.<sup>[2][8]</sup>

## Data structure

The  $d$ -ary heap consists of an array of  $n$  items, each of which has a priority associated with it. These items may be viewed as the nodes in a complete  $d$ -ary tree, listed in breadth first traversal order: the item at position 0 of the array forms the root of the tree, the items at positions  $1-d$  are its children, the next  $d^2$  items are its grandchildren, etc. Thus, the parent of the item at position  $i$  (for any  $i > 0$ ) is the item at position  $\text{floor}((i-1)/d)$  and its children are the items at positions  $di+1$  through  $di+d$ . According to the heap property, in a min-heap, each item has a priority that is at least as large as its parent; in a max-heap, each item has a priority that is no larger than its parent.<sup>[2][3]</sup>

The minimum priority item in a min-heap (or the maximum priority item in a max-heap) may always be found at position 0 of the array. To remove this item from the priority queue, the last item  $x$  in the array is moved into its place, and the length of the array is decreased by one. Then, while item  $x$  and its children do not satisfy the heap property, item  $x$  is swapped with one of its children (the one with the smallest priority in a min-heap, or the one with the largest priority in a max-heap), moving it downward in the tree and later in the array, until eventually the heap property is satisfied. The same downward swapping procedure may be used to increase the priority of an item in a min-heap, or to decrease the priority of an item in a max-heap.<sup>[2][3]</sup>

To insert a new item into the heap, the item is appended to the end of the array, and then while the heap property is violated it is swapped with its parent, moving it upward in the tree and earlier in the array, until eventually the heap property is satisfied. The same upward-swapping procedure may be used to decrease the priority of an item in a min-heap, or to increase the priority of an item in a max-heap.<sup>[2][3]</sup>

To create a new heap from an array of  $n$  items, one may loop over the items in reverse order, starting from the item at position  $n-1$  and ending at the item at position 0, applying the downward-swapping procedure for each item.<sup>[2][3]</sup>

## Analysis

In a  $d$ -ary heap with  $n$  items in it, both the upward-swapping procedure and the downward-swapping procedure may perform as many as  $\log_d n = \log n / \log d$  swaps. In the upward-swapping procedure, each swap involves a single comparison of an item with its parent, and takes constant time. Therefore, the time to insert a new item into the heap, to decrease the priority of an item in a min-heap, or to increase the priority of an item in a max-heap, is  $O(\log n / \log d)$ . In the downward-swapping procedure, each swap involves  $d$  comparisons and takes  $O(d)$  time: it takes  $d-1$  comparisons to determine the minimum or maximum of the children and then one more comparison against the

parent to determine whether a swap is needed. Therefore, the time to delete the root item, to increase the priority of an item in a min-heap, or to decrease the priority of an item in a max-heap, is  $O(d \log n / \log d)$ .<sup>[2][3]</sup>

When creating a *d*-ary heap from a set of *n* items, most of the items are in positions that will eventually hold leaves of the *d*-ary tree, and no downward swapping is performed for those items. At most  $n/d + 1$  items are non-leaves, and may be swapped downwards at least once, at a cost of  $O(d)$  time to find the child to swap them with. At most  $n/d^2 + 1$  nodes may be swapped downward two times, incurring an additional  $O(d)$  cost for the second swap beyond the cost already counted in the first term, etc. Therefore, the total amount of time to create a heap in this way is

$$\sum_{i=1}^{\log_d n} \left( \frac{n}{d^i} + 1 \right) O(d) = O(n).^{[2][3]}$$

The space usage of the *d*-ary heap, with insert and delete-min operations, is linear, as it uses no extra storage other than an array containing a list of the items in the heap.<sup>[2][8]</sup> If changes to the priorities of existing items need to be supported, then one must also maintain pointers from the items to their positions in the heap, which again uses only linear storage.<sup>[2]</sup>

## Applications

Dijkstra's algorithm for shortest paths in graphs and Prim's algorithm for minimum spanning trees both use a min-heap in which there are *n* delete-min operations and as many as *m* decrease-priority operations, where *n* is the number of vertices in the graph and *m* is the number of edges. By using a *d*-ary heap with  $d = m/n$ , the total times for these two types of operations may be balanced against each other, leading to a total time of  $O(m \log_{m/n} n)$  for the algorithm, an improvement over the  $O(m \log n)$  running time of binary heap versions of these algorithms whenever the number of edges is significantly larger than the number of vertices.<sup>[1][5]</sup> An alternative priority queue data structure, the Fibonacci heap, gives an even better theoretical running time of  $O(m + n \log n)$ , but in practice *d*-ary heaps are generally at least as fast, and often faster, than Fibonacci heaps for this application.<sup>[9]</sup>

4-heaps may perform better than binary heaps in practice, even for delete-min operations.<sup>[2][3]</sup> Additionally, a *d*-ary heap typically runs much faster than a binary heap for heap sizes that exceed the size of the computer's cache memory: A binary heap typically requires more cache misses and virtual memory page faults than a *d*-ary heap, each one taking far more time than the extra work incurred by the additional comparisons a *d*-ary heap makes compared to a binary heap.<sup>[6][7]</sup>

## References

- [1] Johnson, D. B. (1975), "Priority queues with update and finding minimum spanning trees", *Information Processing Letters* **4** (3): 53–57, doi:10.1016/0020-0190(75)90001-0.
- [2] Tarjan, R. E. (1983), "3.2. *d*-heaps", *Data Structures and Network Algorithms*, CBMS-NSF Regional Conference Series in Applied Mathematics, **44**, Society for Industrial and Applied Mathematics, pp. 34–38.
- [3] Weiss, M. A. (2007), "*d*-heaps", *Data Structures and Algorithm Analysis* (2nd ed.), Addison-Wesley, p. 216, ISBN 0-321-37013-9.
- [4] Jensen, C.; Katajainen, J.; Vitale, F. (2004), *An extended truth about heaps* (<http://www.cphstl.dk/Report/In-place-multiway-heaps/experimental-study.pdf>), .
- [5] Tarjan (1983), pp. 77 and 91.
- [6] Naor, D.; Martel, C. U.; Matloff, N. S. (1991), "Performance of priority queue structures in a virtual memory environment", *Computer Journal* **34** (5): 428–437, doi:10.1093/comjnl/34.5.428.
- [7] Kamp, Poul-Henning (2010), "You're doing it wrong" (<http://queue.acm.org/detail.cfm?id=1814327>), *ACM Queue* **8** (6), .
- [8] Mortensen, C. W.; Pettie, S. (2005), "The complexity of implicit and space efficient priority queues", *Algorithms and Data Structures: 9th International Workshop, WADS 2005, Waterloo, Canada, August 15-17, 2005, Proceedings*, Lecture Notes in Computer Science, **3608**, Springer-Verlag, pp. 49–60, doi:10.1007/11534273\_6, ISBN 978-3-540-28101-6.
- [9] Cherkassky, B. V.; Goldberg, A. V.; Radzik, T. (1996), "Shortest paths algorithms: Theory and experimental evaluation", *Mathematical Programming* **73** (2): 129–174, doi:10.1007/BF02592101.

## External links

- C++ implementation of generalized heap with D-Heap support (<https://github.com/valyala/gheap>)

# Tries

## Trie

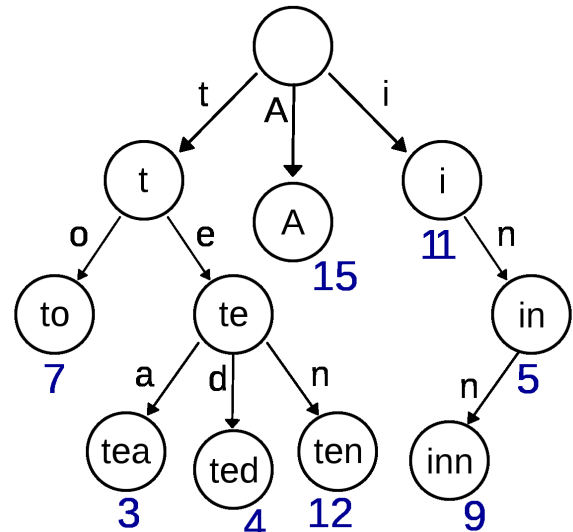
In computer science, a **trie**, or **prefix tree**, is an ordered tree data structure that is used to store a dynamic set or associative array where the keys are usually strings. Unlike a binary search tree, no node in the tree stores the key associated with that node; instead, its position in the tree defines the key with which it is associated. All the descendants of a node have a common prefix of the string associated with that node, and the root is associated with the empty string. Values are normally not associated with every node, only with leaves and some inner nodes that correspond to keys of interest. For the space-optimized presentation of prefix tree, see compact prefix tree.

The term trie comes from **retrieval**. Following the etymology, the inventor, Edward Fredkin, pronounces it /'tri:/ "tree".<sup>[1][2]</sup> However, it is pronounced /'traɪ/ "try" by other authors.<sup>[1][2][3]</sup>

In the example shown, keys are listed in the nodes and values below them. Each complete English word has an arbitrary integer value associated with it. A trie can be seen as a deterministic finite automaton, although the symbol on each edge is often implicit in the order of the branches.

It is not necessary for keys to be explicitly stored in nodes. (In the figure, words are shown only to illustrate how the trie works.)

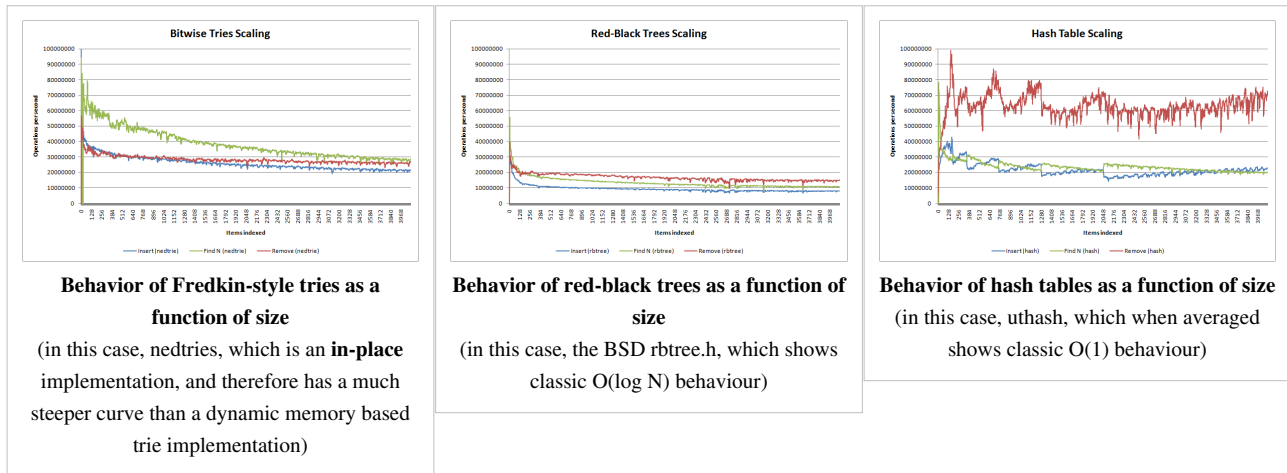
Though it is most common, tries need not be keyed by character strings. The same algorithms can easily be adapted to serve similar functions of ordered lists of any construct, e.g., permutations on a list of digits or shapes. In particular, a **bitwise trie** is keyed on the individual bits making up a short, fixed size of bits such as an integer number or pointer to memory.



A trie for keys "A", "to", "tea", "ted", "ten", "i", "in", and "inn".

## Advantages relative to other search algorithms

A series of graphs showing how different algorithms scale with number of items



Unlike most other algorithms, tries have the peculiar feature that the code path, and hence the time required, is almost identical for insert, delete, and find operations. As a result, for situations where code is inserting, deleting and finding in equal measure, tries can handily beat binary search trees, as well as provide a better basis for the CPU's instruction and branch caches.

The following are the main advantages of tries over binary search trees (BSTs):

- Looking up keys is faster. Looking up a key of length  $m$  takes worst case  $O(m)$  time. A BST performs  $O(\log(n))$  comparisons of keys, where  $n$  is the number of elements in the tree, because lookups depend on the depth of the tree, which is logarithmic in the number of keys if the tree is balanced. Hence in the worst case, a BST takes  $O(m \log n)$  time. Moreover, in the worst case  $\log(n)$  will approach  $m$ . Also, the simple operations tries use during lookup, such as array indexing using a character, are fast on real machines.
- Tries are more space-efficient when they contain a large number of short keys, since nodes are shared between keys with common initial subsequences.
- Tries facilitate longest-prefix matching.
- The number of internal nodes from root to leaf equals the length of the key. Balancing the tree is therefore of no concern.

The following are the main advantages of tries over hash tables:

- Tries support ordered iteration, whereas iteration over a hash table will result in a pseudorandom order given by the hash function (and further affected by the order of hash collisions, which is determined by the implementation).
- Tries facilitate longest-prefix matching, but hashing does not, as a consequence of the above. Performing such a "closest fit" find can, depending on implementation, be as quick as an exact find.
- Tries tend to be faster on average at insertion than hash tables because hash tables must rebuild their index when it becomes full - a very expensive operation. Tries therefore have much better bounded worst-case time costs, which is important for latency-sensitive programs.
- Since no hash function is used, tries are generally faster than hash tables for small keys.

## Applications

### As a replacement for other data structures

As mentioned, a trie has a number of advantages over binary search trees.<sup>[4]</sup> A trie can also be used to replace a hash table, over which it has the following advantages:

- Looking up data in a trie is faster in the worst case,  $O(m)$  time, compared to an imperfect hash table. An imperfect hash table can have key collisions. A key collision is the hash function mapping of different keys to the same position in a hash table. The worst-case lookup speed in an imperfect hash table is  $O(N)$  time, but far more typically is  $O(1)$ , with  $O(m)$  time spent evaluating the hash.
- There are no collisions of different keys in a trie.
- Buckets in a trie which are analogous to hash table buckets that store key collisions are necessary only if a single key is associated with more than one value.
- There is no need to provide a hash function or to change hash functions as more keys are added to a trie.
- A trie can provide an alphabetical ordering of the entries by key.

Tries do have some drawbacks as well:

- Tries can be slower in some cases than hash tables for looking up data, especially if the data is directly accessed on a hard disk drive or some other secondary storage device where the random-access time is high compared to main memory.<sup>[5]</sup>
- Some keys, such as floating point numbers, can lead to long chains and prefixes that are not particularly meaningful. Nevertheless a bitwise trie can handle standard IEEE single and double format floating point numbers.

### Dictionary representation

A common application of a trie is storing a dictionary, such as one found on a mobile telephone. Such applications take advantage of a trie's ability to quickly search for, insert, and delete entries; however, if storing dictionary words is all that is required (i.e. storage of information auxiliary to each word is not required), a minimal acyclic deterministic finite automaton would use less space than a trie. This is because an acyclic deterministic finite automaton can compress identical branches from the trie which correspond to the same suffixes (or parts) of different words being stored.

Tries are also well suited for implementing approximate matching algorithms, including those used in spell checking and hyphenation<sup>[2]</sup> software.

### Algorithms

We can describe trie lookup (and membership) easily. Given a recursive trie type, storing an optional value at each node, and a list of children tries, indexed by the next character, (here, represented as a Haskell data type):

```
data Trie a =
  Trie { value      :: Maybe a
        , children  :: [(Char, Trie a)] }
```

We can look up a value in the trie as follows:

```
find :: String -> Trie a -> Maybe a
find []      t = value t
find (k:ks) t = case lookup k (children t) of
                  Nothing  -> Nothing
                  Just t'  -> find ks t'
```

In an imperative style, and assuming an appropriate data type in place, we can describe the same algorithm in Python (here, specifically for testing membership). Note that `children` is map of a node's children; and we say that a "terminal" node is one which contains a valid word.

```
def find(node, key):
    for char in key:
        if char not in node.children:
            return None
        else:
            node = node.children[char]
    return node.value
```

A simple Ruby version:

```
class Trie
  def initialize
    @root = Hash.new
  end

  def build(str)
    node = @root
    str.each_char do |ch|
      node[ch] ||= Hash.new
      node = node[ch]
    end
    node[:end] = true
  end

  def find(str)
    node = @root
    str.each_char do |ch|
      return nil unless node = node[ch]
    end
    node[:end] == true
  end
end
```

A compiling Java version:

```
public class MinimalExample{
    private interface Node {
        public static final Node EMPTY_NODE = new Node() {
            @Override public String getValue() { return ""; }
            @Override public boolean containsChildValue(char c) {
return false; }
            @Override public Node getChild(char c) { return this; }
        };
        public String getValue();
        public boolean containsChildValue(char c);
        public Node getChild(char c);
    }
}
```



```
}

public Node findValue(Node startNode, String value) {
    Node current = startNode;
    for (char c : value.toCharArray()) {
        if (current.containsChildValue(c)) {
            current = current.getChild(c);
        } else {
            current = Node.EMPTY_NODE;
            break;
        }
    }
    return current;
}
```

## Sorting

Lexicographic sorting of a set of keys can be accomplished with a simple trie-based algorithm as follows:

- Insert all keys in a trie.
- Output all keys in the trie by means of pre-order traversal, which results in output that is in lexicographically increasing order. Pre-order traversal is a kind of depth-first traversal. In-order traversal is another kind of depth-first traversal that is more appropriate for outputting the values that are in a binary search tree rather than a trie.

This algorithm is a form of radix sort.

A trie forms the fundamental data structure of Burtsort, currently (2007) the fastest known, memory/cache-based, string sorting algorithm.<sup>[6]</sup>

A parallel algorithm for sorting  $N$  keys based on tries is  $O(1)$  if there are  $N$  processors and the lengths of the keys have a constant upper bound. There is the potential that the keys might collide by having common prefixes or by being identical to one another, reducing or eliminating the speed advantage of having multiple processors operating in parallel.

## Full text search

A special kind of trie, called a suffix tree, can be used to index all suffixes in a text in order to carry out fast full text searches.

## Bitwise tries

Bitwise tries are much the same as a normal character based trie except that individual bits are used to traverse what effectively becomes a form of binary tree. Generally, implementations use a special CPU instruction to very quickly find the first set bit in a fixed length key (e.g. GCC's `__builtin_clz()` intrinsic). This value is then used to index a 32 or 64 entry table which points to the first item in the bitwise trie with that number of leading zero bits. The search then proceeds by testing each subsequent bit in the key and choosing `child[0]` or `child[1]` appropriately until the item is found.

Although this process might sound slow, it is very cache-local and highly parallelizable due to the lack of register dependencies and therefore in fact has excellent performance on modern out-of-order execution CPUs. A red-black tree for example performs much better on paper, but is highly cache-unfriendly and causes multiple pipeline and

TLB stalls on modern CPUs which makes that algorithm bound by memory latency rather than CPU speed. In comparison, a bitwise trie rarely accesses memory and when it does it does so only to read, thus avoiding SMP cache coherency overhead, and hence is becoming increasingly the algorithm of choice for code which does a lot of insertions and deletions such as memory allocators (e.g. recent versions of the famous Doug Lea's allocator (dlmalloc) and its descendents).

A reference implementation of bitwise tries in C and C++ useful for further study can be found at <http://www.nedprod.com/programs/portable/nedtries/>.

## Compressing tries

When the trie is mostly static, i.e. all insertions or deletions of keys from a prefilled trie are disabled and only lookups are needed, and when the trie nodes are not keyed by node specific data (or if the node's data is common) it is possible to compress the trie representation by merging the common branches.<sup>[7]</sup> This application is typically used for compressing lookup tables when the total set of stored keys is very sparse within their representation space.

For example it may be used to represent sparse bitsets (i.e. subsets of a much fixed enumerable larger set) using a trie keyed by the bit element position within the full set, with the key created from the string of bits needed to encode the integral position of each element. The trie will then have a very degenerate form with many missing branches, and compression becomes possible by storing the leaf nodes (set segments with fixed length) and combining them after detecting the repetition of common patterns or by filling the unused gaps.

Such compression is also typically used in the implementation of the various fast lookup tables needed to retrieve Unicode character properties (for example to represent case mapping tables, or lookup tables containing the combination of base and combining characters needed to support Unicode normalization). For such application, the representation is similar to transforming a very large unidimensional sparse table into a multidimensional matrix, and then using the coordinates in the hyper-matrix as the string key of an uncompressed trie. The compression will then consist of detecting and merging the common columns within the hyper-matrix to compress the last dimension in the key; each dimension of the hypermatrix stores the start position within a storage vector of the next dimension for each coordinate value, and the resulting vector is itself compressible when it is also sparse, so each dimension (associated to a layer level in the trie) is compressed separately.

Some implementations do support such data compression within dynamic sparse tries and allow insertions and deletions in compressed tries, but generally this has a significant cost when compressed segments need to be split or merged, and some tradeoff has to be made between the smallest size of the compressed trie and the speed of updates, by limiting the range of global lookups for comparing the common branches in the sparse trie.

The result of such compression may look similar to trying to transform the trie into a directed acyclic graph (DAG), because the reverse transform from a DAG to a trie is obvious and always possible, however it is constrained by the form of the key chosen to index the nodes.

Another compression approach is to "unravel" the data structure into a single byte array.<sup>[8]</sup> This approach eliminates the need for node pointers which reduces the memory requirements substantially and makes memory mapping possible which allows the virtual memory manager to load the data into memory very efficiently.

Another compression approach is to "pack" the trie.<sup>[2]</sup> Liang describes a space-efficient implementation of a sparse packed trie applied to hyphenation, in which the descendants of each node may be interleaved in memory.

## External links

- NIST's Dictionary of Algorithms and Data Structures: Trie <sup>[9]</sup>
- Trie implementation and visualisation in flash <sup>[10]</sup>
- Tries <sup>[11]</sup> by Lloyd Allison
- Using Tries <sup>[12]</sup> Topcoder tutorial
- An Implementation of Double-Array Trie <sup>[13]</sup>
- de la Briandais Tree <sup>[14]</sup>
- Discussing a trie implementation in Lisp <sup>[15]</sup>
- ServerKit "parse trees" implement a form of Trie in C <sup>[16]</sup>
- A simple implementation of Trie in Python <sup>[17]</sup>
- A Trie implemented in Ruby <sup>[18]</sup>
- A reference implementation of bitwise tries in C and C++ <sup>[19]</sup>
- A reference implementation in Java <sup>[20]</sup>
- A quick tutorial on TRIE in Java and C++ <sup>[21]</sup>
- A compact C implementation of Judy Tries <sup>[22]</sup>
- Data::Trie <sup>[23]</sup> and Tree::Trie <sup>[24]</sup> Perl implementations.

## References

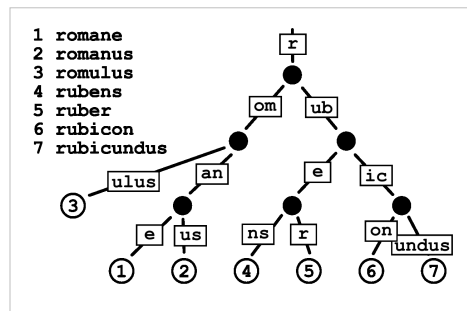
- [1] Black, Paul E. (2009-11-16). "trie" (<http://www.webcitation.org/5pqUULy24>). *Dictionary of Algorithms and Data Structures*. National Institute of Standards and Technology. Archived from the original (<http://www.nist.gov/dads/HTML/trie.html>) on 2010-05-19. .
- [2] Franklin Mark Liang (1983). *Word Hy-phen-a-tion By Com-put-er* (<http://www.webcitation.org/5pqOfzIIA>) (Doctor of Philosophy thesis). Stanford University. Archived from the original (<http://www.tug.org/docs/liang/liang-thesis.pdf>) on 2010-05-19. . Retrieved 2010-03-28.
- [3] Knuth, Donald (1997). "6.3: Digital Searching". *The Art of Computer Programming Volume 3: Sorting and Searching* (2nd ed.). Addison-Wesley. p. 492. ISBN 0-201-89685-0.
- [4] Bentley, Jon; Sedgewick, Robert (1998-04-01). "Ternary Search Trees" (<http://web.archive.org/web/20080623071352/http://www.ddj.com/windows/184410528>). *Dr. Dobb's Journal* (Dr Dobb's). Archived from the original (<http://www.ddj.com/windows/184410528>) on 2008-06-23. .
- [5] Edward Fredkin (1960). "Trie Memory". *Communications of the ACM* **3** (9): 490–499. doi:10.1145/367390.367400.
- [6] "Cache-Efficient String Sorting Using Copying" (<http://www.cs.mu.oz.au/~rsinha/papers/SinhaRingZobel-2006.pdf>) (PDF). . Retrieved 2008-11-15.
- [7] Jan Daciuk, Stoyan Mihov, Bruce W. Watson, Richard E. Watson (2000). "Incremental Construction of Minimal Acyclic Finite-State Automata" (<http://www.mitpressjournals.org/doi/abs/10.1162/089120100561601>). *Computational Linguistics* (Association for Computational Linguistics) **26**: 3. doi:10.1162/089120100561601. Archived from the original (<http://www.pg.gda.pl/~jandac/daciuk98.ps.gz>) on 2006-03-13. . Retrieved 2009-05-28. "This paper presents a method for direct building of minimal acyclic finite states automaton which recognizes a given finite list of words in lexicographical order. Our approach is to construct a minimal automaton in a single phase by adding new strings one by one and minimizing the resulting automaton on-the-fly"
- [8] Ulrich Germann, Eric Joanis, Samuel Larkin (2009). "Tightly packed tries: how to fit large models into memory, and make them load fast, too" (<http://www.aclweb.org/anthology/W/W09/W09-1505.pdf>) (PDF). *ACL Workshops: Proceedings of the Workshop on Software Engineering, Testing, and Quality Assurance for Natural Language Processing*. Association for Computational Linguistics. pp. 31–39. . "We present Tightly Packed Tries (TPTs), a compact implementation of read-only, compressed trie structures with fast on-demand paging and short load times. We demonstrate the benefits of TPTs for storing n-gram back-off language models and phrase tables for statistical machine translation. Encoded as TPTs, these databases require less space than flat text file representations of the same data compressed with the gzip utility. At the same time, they can be mapped into memory quickly and be searched directly in time linear in the length of the key, without the need to decompress the entire file. The overhead for local decompression during search is marginal."
- [9] <http://www.nist.gov/dads/HTML/trie.html>
- [10] <http://blog.ivank.net/trie-in-as3.html>
- [11] <http://www.csse.monash.edu.au/~lloyd/tildeAlgDS/Tree/Trie/>
- [12] <http://www.topcoder.com/tc?module=Static&d1=tutorials&d2=usingTries>
- [13] <http://linux.thai.net/~thep/datrie/datrie.html>
- [14] <http://tom.biodome.org/briandais.html>
- [15] [http://groups.google.com/group/comp.lang.lisp/browse\\_thread/thread/01e485291d150938/9aacb626fa26c516](http://groups.google.com/group/comp.lang.lisp/browse_thread/thread/01e485291d150938/9aacb626fa26c516)
- [16] <http://serverkit.org/apiref-wip/node59.html>
- [17] <http://kb.pyground.com/2011/05/trie-data-structure.html>

- [18] [http://scanty-evidence-1.herokuapp.com/past/2010/5/10/ruby\\_trie/](http://scanty-evidence-1.herokuapp.com/past/2010/5/10/ruby_trie/)
- [19] <http://www.nedprod.com/programs/portable/nedtries/>
- [20] <http://www.superliminal.com/sources/TrieMap.java.html>
- [21] <http://www.technicalypto.com/2010/04/trie-in-java.html>
- [22] <http://code.google.com/p/judyarray>
- [23] <http://search.cpan.org/~hammond/data-trie-0.01/Trie.pm>
- [24] <http://search.cpan.org/~avif/Tree-Trie-1.7/Trie.pm>

- de la Briandais, R. (1959). "File Searching Using Variable Length Keys". *Proceedings of the Western Joint Computer Conference*: 295–298.

## Radix tree

In computer science, a **radix tree** (also **patricia trie** or **radix trie** or **compact prefix tree**) is a space-optimized trie data structure where each node with only one child is merged with its child. The result is that every internal node has at least two children. Unlike in regular tries, edges can be labeled with sequences of characters as well as single characters. This makes them much more efficient for small sets (especially if the strings are long) and for sets of strings that share long prefixes.



As an optimization, edge labels can be stored in constant size by using two pointers to a string (for the first and last characters).<sup>[1]</sup>

## Applications

As mentioned, radix trees are useful for constructing associative arrays with keys that can be expressed as strings. They find particular application in the area of IP routing, where the ability to contain large ranges of values with a few exceptions is particularly suited to the hierarchical organization of IP addresses.<sup>[2]</sup> They are also used for inverted indexes of text documents in information retrieval.

## Operations

Radix tries support insertion, deletion, and searching operations. Insertion adds a new string to the trie while trying to minimize the amount of data stored. Deletion removes a string from the trie. Searching operations include exact lookup, find predecessor, find successor, and find all strings with a prefix. All of these operations are  $O(k)$  where  $k$  is the maximum length of all strings in the set. This list may not be exhaustive.

## Lookup

The lookup operation determines if a string exists in a trie. Most operations modify this approach in some way to handle their specific tasks. For instance, the node where a string terminates may be of importance. This operation is similar to tries except that some edges consume multiple characters.

The following pseudo code assumes that these classes exist.

### Edge

- *Node* targetNode
- *string* label

### Node

- *Array of Edges* edges
- *function* isLeaf()

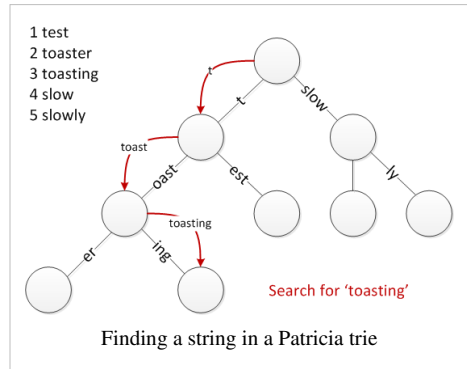
```
function lookup(string x)
{
    // Begin at the root with no characters found
    Node traverseNode := root;
    int charsFound := 0;

    // Traverse until a leaf is found or it is not possible to continue
    while (traverseNode != null && !traverseNode.isLeaf() && charsFound < x.length)
    {
        // Get the next edge to explore based on the characters not yet found in x
        Edge nextEdge := select edge from traverseNode.edges where edge.label is a prefix of x.suffix(charsFound)
        // x.suffix(charsFound) returns the last (x.length - charsFound) characters of x

        // Was an edge found?
        if (nextEdge != null)
        {
            // Set the next node to explore
            traverseNode := nextEdge.targetNode;

            // Increment characters found based on the label stored at the edge
            charsFound += nextEdge.label.length;
        }
        else
        {
            // Terminate loop
            traverseNode := null;
        }
    }

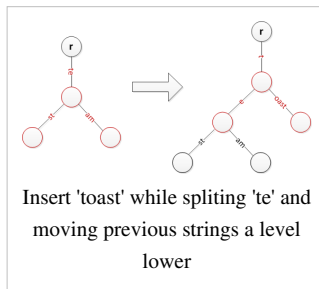
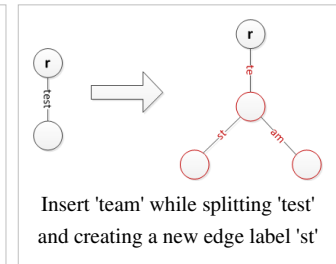
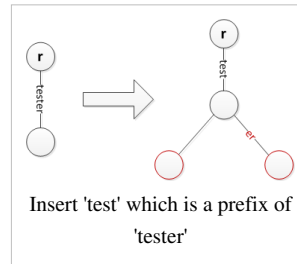
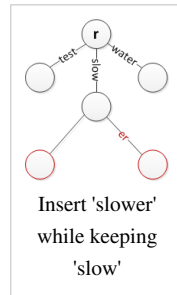
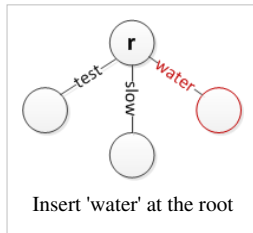
    // A match is found if we arrive at a leaf node and have used up exactly x.length characters
    return (traverseNode != null && traverseNode.isLeaf() && charsFound == x.length);
}
```



## Insertion

To insert a string, we search the trie until we can make no further progress. At this point we either add a new outgoing edge labeled with all remaining characters in the input string, or if there is already an outgoing edge sharing a prefix with the remaining input string, we split it into two edges (the first labeled with the common prefix) and proceed. This splitting step ensures that no node has more children than there are possible string characters.

Several cases of insertion are shown below, though more may exist. Note that *r* simply represents the root. It is assumed that edges can be labelled with empty strings to terminate strings where necessary and that the root has no incoming edge.



## Deletion

To delete a string *x* from a trie, we first locate the leaf representing *x*. Then, assuming *x* exists, we remove the corresponding leaf node. If the parent of our leaf node has only one other child, then that child's incoming label is appended to the parent's incoming label and the child is removed.

## Additional Operations

- Find all strings with common prefix: Returns an array of strings which begin with the same prefix.
- Find predecessor: Locates the largest string less than a given string, by lexicographic order.
- Find successor: Locates the smallest string greater than a given string, by lexicographic order.

## History

Donald R. Morrison first described what he called "Patricia trees" in 1968;<sup>[3]</sup> the name comes from the acronym **PATRICIA**, which stands for "*Practical Algorithm To Retrieve Information Coded In Alphanumeric*". Gernot Gwehenberger independently invented and described the data structure at about the same time.<sup>[4]</sup>

## Comparison to other data structures

(In the following comparisons, it is assumed that the keys are of length  $k$  and the data structure contains  $n$  elements.)

Unlike balanced trees, radix trees permit lookup, insertion, and deletion in  $O(k)$  time rather than  $O(\log n)$ . This doesn't seem like an advantage, since normally  $k \geq \log n$ , but in a balanced tree every comparison is a string comparison requiring  $O(k)$  worst-case time, many of which are slow in practice due to long common prefixes. In a trie, all comparisons require constant time, but it takes  $m$  comparisons to look up a string of length  $m$ . Radix trees can perform these operations with fewer comparisons and require many fewer nodes.

Radix trees also share the disadvantages of tries, however: as they can only be applied to strings of elements or elements with an efficiently reversible mapping (injection) to strings, they lack the full generality of balanced search trees, which apply to any data type with a total ordering. A reversible mapping to strings can be used to produce the required total ordering for balanced search trees, but not the other way around. This can also be problematic if a data type only provides a comparison operation, but not a (de)serialization operation.

Hash tables are commonly said to have expected  $O(1)$  insertion and deletion times, but this is only true when considering computation of the hash of the key to be a constant time operation. When hashing the key is taken into account, hash tables have expected  $O(k)$  insertion and deletion times, but may take longer in the worst-case depending on how collisions are handled. Radix trees have worst-case  $O(k)$  insertion and deletion. The successor/predecessor operations of radix trees are also not implemented by hash tables.

Since nodes in a Radix tree require arrays to be allocated that are the size of the alphabet being used, they start to consume very large amounts of space if two byte characters in various languages need to be stored. Another Trie data structure that doesn't have this problem is a Ternary search tree.

## Variants

A common extension of radix trees uses two colors of nodes, 'black' and 'white'. To check if a given string is stored in the tree, the search starts from the top and follows the edges of the input string until no further progress can be made. If the search-string is consumed and the final node is a black node, the search has failed; if it is white, the search has succeeded. This enables us to add a large range of strings with a common prefix to the tree, using white nodes, then remove a small set of "exceptions" in a space-efficient manner by *inserting* them using black nodes.

The **HAT-trie** is a radix tree based cache-conscious data structure that offers efficient string storage and retrieval, and ordered iterations. Performance, with respect to both time and space, is comparable to the cache-conscious hashtable.<sup>[5][6]</sup>

Nodes in a Radix Tree require allocating arrays whose size is equal to the size of the alphabet being used (typically  $\leq 256$  for one byte characters). If strings in multiple languages are to be stored, then a larger alphabet size is required which can cause the Radix Tree to consume too much memory. A small change to the Radix Tree where the edge array is replaced with a Hash Table that supports growth can be used to remove the effect of alphabet size on space consumption. Modification and access to the table runs in  $O(1)$  amortized time thus all theoretical run times for Radix Tree operations are preserved.

## References

- [1] Morin, Patrick. "Data Structures for Strings" (<http://cg.scs.carleton.ca/~morin/teaching/5408/notes/strings.pdf>). . Retrieved 15 April 2012.
- [2] Knizhnik, Konstantin. "Patricia Tries: A Better Index For Prefix Searches" (<http://www.ddj.com/architect/208800854>), *Dr. Dobb's Journal*, June, 2008.
- [3] Morrison, Donald R. Practical Algorithm to Retrieve Information Coded in Alphanumeric (<http://portal.acm.org/citation.cfm?id=321481>)
- [4] G. Gwehenberger, Anwendung einer binären Verweiskettenmethode beim Aufbau von Listen. (<http://cr.yp.to/bib/1968/gwehenberger.html>) Elektronische Rechenanlagen 10 (1968), pp. 223–226
- [5] Askitis, Nikolas; Sinha, Ranjan (2007). *HAT-trie: A Cache-conscious Trie-based Data Structure for Strings* (<http://portal.acm.org/citation.cfm?id=1273749.1273761&coll=GUIDE&dl=>). **62**. pp. 97–105. ISBN 1-920682-43-0.
- [6] Askitis, Nikolas; Sinha, Ranjan (2010). *Engineering scalable, cache and space efficient tries for strings* (<http://www.springerlink.com/content/86574173183j6565/>). doi:10.1007/s00778-010-0183-9. ISBN 1066-8888 (Print) 0949-877X (Online). .

## External links

- Algorithms and Data Structures Research & Reference Material: PATRICIA (<http://www.csse.monash.edu.au/~lloyd/tildeAlgDS/Tree/PATRICIA/>), by Lloyd Allison, Monash University
- Patricia Tree (<http://www.nist.gov/dads/HTML/patriciatree.html>), NIST Dictionary of Algorithms and Data Structures
- Crit-bit trees (<http://cr.yp.to/critbit.html>), by Daniel J. Bernstein
- Radix Tree API in the Linux Kernel (<http://lwn.net/Articles/175432/>), by Jonathan Corbet
- Kart (key alteration radix tree) (<http://code.dogmap.org/kart/>), by Paul Jarc

## Implementations

- GNU C++ Standard library has a trie implementation ([http://gcc.gnu.org/onlinedocs/libstdc++/ext/pb\\_ds/trie\\_based\\_containers.html](http://gcc.gnu.org/onlinedocs/libstdc++/ext/pb_ds/trie_based_containers.html))
- Java implementation of Radix Tree (<http://badgenow.com/p/radixtree/>), by Tahseen Ur Rehman
- Java implementation of Concurrent Radix Tree (<http://code.google.com/p/concurrent-trees/>), by Niall Gallagher
- C# implementation of a Radix Tree (<http://paratechnical.blogspot.com/2011/03/radix-tree-implementation-in-c.html>)
- Practical Algorithm Template Library (<http://code.google.com/p/patl/>), a C++ library on PATRICIA tries (VC++ >=2003, GCC G++ 3.x), by Roman S. Klyujkov
- Patricia Trie C++ template class implementation (<http://www.codeproject.com/KB/string/PatriciaTrieTemplateClass.aspx>), by Radu Gruian
- Haskell standard library implementation (<http://hackage.haskell.org/packages/archive/containers/latest/doc/html/Data-IntMap.html>) "based on big-endian patricia trees". Web-browsable source code (<http://hackage.haskell.org/packages/archive/containers/latest/doc/html/src/Data-IntMap.html>).
- Patricia Trie implementation in Java (<http://code.google.com/p/patricia-trie/>), by Roger Kapsi and Sam Berlin
- Crit-bit trees (<http://github.com/agl/critbit>) forked from C code by Daniel J. Bernstein
- Patricia Trie implementation in C ([http://cprops.sourceforge.net/gen/docs/trie\\_8c-source.html](http://cprops.sourceforge.net/gen/docs/trie_8c-source.html)), in libcprops (<http://cprops.sourceforge.net>)
- Patricia Trees : efficient sets and maps over integers in (<http://www.lri.fr/~filliatr/ftp/ocaml/ds>) OCaml, by Jean-Christophe Filliâtre

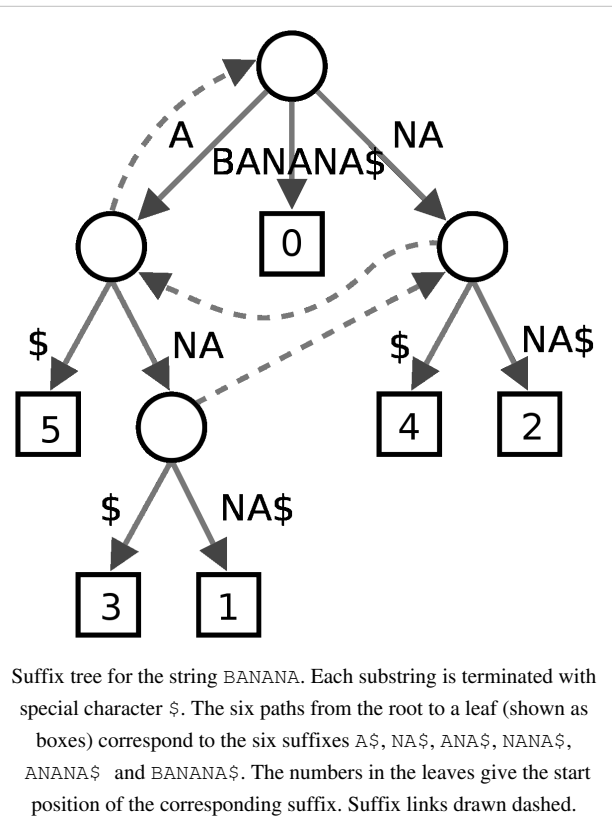


# Suffix tree

In computer science, a **suffix tree** (also called **PAT tree** or, in an earlier form, **position tree**) is a data structure that presents the suffixes of a given string in a way that allows for a particularly fast implementation of many important string operations.

The suffix tree for a string  $S$  is a tree whose edges are labeled with strings, such that each suffix of  $S$  corresponds to exactly one path from the tree's root to a leaf. It is thus a radix tree (more specifically, a Patricia tree) for the suffixes of  $S$ .

Constructing such a tree for the string  $S$  takes time and space linear in the length of  $S$ . Once constructed, several operations can be performed quickly, for instance locating a substring in  $S$ , locating a substring if a certain number of mistakes are allowed, locating matches for a regular expression pattern etc. Suffix trees also provided one of the first linear-time solutions for the longest common substring problem. These speedups come at a cost: storing a string's suffix tree typically requires significantly more space than storing the string itself.



## History

The concept was first introduced as a *position tree* by Weiner (1973), which Donald Knuth subsequently characterized as "Algorithm of the Year 1973". The construction was greatly simplified by McCreight (1976), and also by Ukkonen (1995).<sup>[1]</sup> Ukkonen provided the first online-construction of suffix trees, now known as Ukkonen's algorithm, with running time that matched the then fastest algorithms. These algorithms are all linear-time for constant-size alphabet, and have worst-case running time of  $O(n \log n)$  in general.

Farach (1997) gave the first suffix tree construction algorithm that is optimal for all alphabets. In particular, this is the first linear-time algorithm for strings drawn from an alphabet of integers in a polynomial range. Farach's algorithm has become the basis for new algorithms for constructing both suffix trees and suffix arrays, for example, in external memory, compressed, succinct, etc.

## Definition

The suffix tree for the string  $S$  of length  $n$  is defined as a tree such that:<sup>[2]</sup>

- the paths from the root to the leaves have a one-to-one relationship with the suffixes of  $S$ ,
- edges spell non-empty strings,
- and all internal nodes (except perhaps the root) have at least two children.

Since such a tree does not exist for all strings,  $S$  is padded with a terminal symbol not seen in the string (usually denoted  $\$$ ). This ensures that no suffix is a prefix of another, and that there will be  $n$  leaf nodes, one for each of the  $n$  suffixes of  $S$ . Since all internal non-root nodes are branching, there can be at most  $n - 1$  such nodes, and  $n + (n - 1) + 1 = 2n$  nodes in total ( $n$  leaves,  $n - 1$  internal nodes, 1 root).

**Suffix links** are a key feature for older linear-time construction algorithms, although most newer algorithms, which are based on Farach's algorithm, dispense with suffix links. In a complete suffix tree, all internal non-root nodes have a suffix link to another internal node. If the path from the root to a node spells the string  $\chi\alpha$ , where  $\chi$  is a single character and  $\alpha$  is a string (possibly empty), it has a suffix link to the internal node representing  $\alpha$ . See for example the suffix link from the node for ANA to the node for NA in the figure above. Suffix links are also used in some algorithms running on the tree.

## Generalized suffix tree

A generalized suffix tree is a suffix tree made for a set of words instead only for a single word. It represents all suffixes from this set of words. Each word must be terminated by a different termination symbol or word.

## Functionality

A suffix tree for a string  $S$  of length  $n$  can be built in  $\Theta(n)$  time, if the letters come from an alphabet of integers in a polynomial range (in particular, this is true for constant-sized alphabets).<sup>[3]</sup> For larger alphabets, the running time is dominated by first sorting the letters to bring them into a range of size  $O(n)$ ; in general, this takes  $O(n \log n)$  time. The costs below are given under the assumption that the alphabet is constant.

Assume that a suffix tree has been built for the string  $S$  of length  $n$ , or that a generalised suffix tree has been built for the set of strings  $D = \{S_1, S_2, \dots, S_K\}$  of total length  $n = |n_1| + |n_2| + \dots + |n_K|$ . You can:

- Search for strings:
  - Check if a string  $P$  of length  $m$  is a substring in  $O(m)$  time.<sup>[4]</sup>
  - Find the first occurrence of the patterns  $P_1, \dots, P_q$  of total length  $m$  as substrings in  $O(m)$  time.
  - Find all  $z$  occurrences of the patterns  $P_1, \dots, P_q$  of total length  $m$  as substrings in  $O(m + z)$  time.<sup>[5]</sup>
  - Search for a regular expression  $P$  in time expected sublinear in  $n$ .<sup>[6]</sup>
  - Find for each suffix of a pattern  $P$ , the length of the longest match between a prefix of  $P[i \dots m]$  and a substring in  $D$  in  $\Theta(m)$  time.<sup>[7]</sup> This is termed the **matching statistics** for  $P$ .
- Find properties of the strings:
  - Find the longest common substrings of the string  $S_i$  and  $S_j$  in  $\Theta(n_i + n_j)$  time.<sup>[8]</sup>
  - Find all maximal pairs, maximal repeats or supermaximal repeats in  $\Theta(n + z)$  time.<sup>[9]</sup>
  - Find the Lempel–Ziv decomposition in  $\Theta(n)$  time.<sup>[10]</sup>
  - Find the longest repeated substrings in  $\Theta(n)$  time.
  - Find the most frequently occurring substrings of a minimum length in  $\Theta(n)$  time.
  - Find the shortest strings from  $\Sigma$  that do not occur in  $D$ , in  $O(n + z)$  time, if there are  $z$  such strings.
  - Find the shortest substrings occurring only once in  $\Theta(n)$  time.
  - Find, for each  $i$ , the shortest substrings of  $S_i$  not occurring elsewhere in  $D$  in  $\Theta(n)$  time.

The suffix tree can be prepared for constant time lowest common ancestor retrieval between nodes in  $\Theta(n)$  time.<sup>[11]</sup>

One can then also:

- Find the longest common prefix between the suffixes  $S_i[p..n_i]$  and  $S_j[q..n_j]$  in  $\Theta(1)$ .<sup>[12]</sup>
- Search for a pattern  $P$  of length  $m$  with at most  $k$  mismatches in  $O(kn + z)$  time, where  $z$  is the number of hits.<sup>[13]</sup>
- Find all  $z$  maximal palindromes in  $\Theta(n)$ ,<sup>[14]</sup> or  $\Theta(gn)$  time if gaps of length  $g$  are allowed, or  $\Theta(kn)$  if  $k$  mismatches are allowed.<sup>[15]</sup>
- Find all  $z$  tandem repeats in  $O(n \log n + z)$ , and  $k$ -mismatch tandem repeats in  $O(kn \log(n/k) + z)$ .<sup>[16]</sup>
- Find the longest substrings common to at least  $k$  strings in  $D$  for  $k = 2, \dots, K$  in  $\Theta(n)$  time.<sup>[17]</sup>
- Find the longest palindromic substring of a given string (using the suffix trees of both the string and its reverse) in linear time.<sup>[18]</sup>

## Applications

Suffix trees can be used to solve a large number of string problems that occur in text-editing, free-text search, computational biology and other application areas.<sup>[19]</sup> Primary applications include:<sup>[19]</sup>

- String search, in  $O(m)$  complexity, where  $m$  is the length of the sub-string (but with initial  $O(n)$  time required to build the suffix tree for the string)
- Finding the longest repeated substring
- Finding the longest common substring
- Finding the longest palindrome in a string

Suffix trees are often used in bioinformatics applications, searching for patterns in DNA or protein sequences (which can be viewed as long strings of characters). The ability to search efficiently with mismatches might be considered their greatest strength. Suffix trees are also used in data compression; they can be used to find repeated data, and can be used for the sorting stage of the Burrows–Wheeler transform. Variants of the LZW compression schemes use suffix trees (LZSS). A suffix tree is also used in suffix tree clustering, a data clustering algorithm used in some search engines.<sup>[20]</sup>

## Implementation

If each node and edge can be represented in  $\Theta(1)$  space, the entire tree can be represented in  $\Theta(n)$  space. The total length of all the strings on all of the edges in the tree is  $O(n^2)$ , but each edge can be stored as the position and length of a substring of  $S$ , giving a total space usage of  $\Theta(n)$  computer words. The worst-case space usage of a suffix tree is seen with a fibonacci word, giving the full  $2n$  nodes.

An important choice when making a suffix tree implementation is the parent-child relationships between nodes. The most common is using linked lists called **sibling lists**. Each node has a pointer to its first child, and to the next node in the child list it is a part of. Hash maps, sorted/unordered arrays (with array doubling), and balanced search trees may also be used, giving different running time properties. We are interested in:

- The cost of finding the child on a given character.
- The cost of inserting a child.
- The cost of enlisting all children of a node (divided by the number of children in the table below).

Let  $\sigma$  be the size of the alphabet. Then you have the following costs:

	Lookup	Insertion	Traversal
Sibling lists / unsorted arrays	$O(\sigma)$	$\Theta(1)$	$\Theta(1)$
Hash maps	$\Theta(1)$	$\Theta(1)$	$O(\sigma)$
Balanced search tree	$O(\log \sigma)$	$O(\log \sigma)$	$O(1)$
Sorted arrays	$O(\log \sigma)$	$O(\sigma)$	$O(1)$
Hash maps + sibling lists	$O(1)$	$O(1)$	$O(1)$

Note that the insertion cost is amortised, and that the costs for hashing are given perfect hashing.

The large amount of information in each edge and node makes the suffix tree very expensive, consuming about ten to twenty times the memory size of the source text in good implementations. The suffix array reduces this requirement to a factor of 8 (For array including LCP values built within 32-bit address space and 8-bit characters. This factor depends on this properties and may reach 2 with usage of 4-byte wide characters (needed to contain any symbol in some UNIX-like systems, see `wchar_t` on 32-bit systems), and researchers have continued to find smaller indexing structures.

## External construction

Suffix trees quickly outgrow the main memory on standard machines for sequence collections in the order of gigabytes. As such, their construction calls for external memory approaches.

There are theoretical results for constructing suffix trees in external memory. The algorithm by Farach-Colton, Ferragina & Muthukrishnan (2000) is theoretically optimal, with an I/O complexity equal to that of sorting. However the overall intricacy of this algorithm has prevented, so far, its practical implementation.<sup>[21]</sup>

On the other hand, there have been practical works for constructing disk-based suffix trees which scale to (few) GB/hours. The state of the art methods are TDD,<sup>[22]</sup> TRELLIS,<sup>[23]</sup> DiGeST,<sup>[24]</sup> and B<sup>2</sup>ST.<sup>[25]</sup>

TDD and TRELLIS scale up to the entire human genome – approximately 3GB – resulting in a disk-based suffix tree of a size in the tens of gigabytes.<sup>[22][23]</sup> However, these methods cannot handle efficiently collections of sequences exceeding 3GB.<sup>[24]</sup> DiGeST performs significantly better and is able to handle collections of sequences in the order of 6GB in about 6 hours.<sup>[24]</sup> All these methods can efficiently build suffix trees for the case when the tree does not fit in main memory, but the input does. The most recent method, B<sup>2</sup>ST,<sup>[25]</sup> scales to handle inputs that do not fit in main memory. ERA is a recent parallel suffix tree construction method that is significantly faster. ERA can index the entire human genome in 19 minutes on an 8-core desktop computer with 16GB RAM. On a simple Linux cluster with 16 nodes (4GB RAM per node), ERA can index the entire human genome in less than 9 minutes.<sup>[26]</sup>

## Notes

- [1] Giegerich & Kurtz (1997)
- [2] Gusfield (1999), p.90.
- [3] Farach (1997)
- [4] Gusfield (1999), p.92.
- [5] Gusfield (1999), p.123.
- [6] Baeza-Yates & Gonnet (1996)
- [7] Gusfield (1999), p.132.
- [8] Gusfield (1999), p.125.
- [9] Gusfield (1999), p.144.
- [10] Gusfield (1999), p.166.
- [11] Gusfield (1999), Chapter 8.
- [12] Gusfield (1999), p.196.
- [13] Gusfield (1999), p.200.
- [14] Gusfield (1999), p.198.

- [15] Gusfield (1999), p.201.
- [16] Gusfield (1999), p.204.
- [17] Gusfield (1999), p.205.
- [18] Gusfield (1999), pp.197–199.
- [19] Allison, L.. "Suffix Trees" (<http://www.allisons.org/ll/AlgDS/Tree/Suffix/>). . Retrieved 2008-10-14.
- [20] First introduced by Zamir & Etzioni (1998).
- [21] Smyth (2003)
- [22] Tata, Hankins & Patel (2003).
- [23] Phoophakdee & Zaki (2007).
- [24] Barsky et al. (2008).
- [25] Barsky et al. (2009).
- [26] Mansour et al. (2011)

## References

- Baeza-Yates, Ricardo A.; Gonnet, Gaston H. (1996), "Fast text searching for regular expressions or automaton searching on tries", *Journal of the ACM* **43** (6): 915–936, doi:10.1145/235809.235810.
- Barsky, Marina; Stege, Ulrike; Thomo, Alex; Upton, Chris (2008), "A new method for indexing genomes using on-disk suffix trees", *CIKM '08: Proceedings of the 17th ACM Conference on Information and Knowledge Management*, New York, NY, USA: ACM, pp. 649–658.
- Barsky, Marina; Stege, Ulrike; Thomo, Alex; Upton, Chris (2009), "Suffix trees for very large genomic sequences", *CIKM '09: Proceedings of the 18th ACM Conference on Information and Knowledge Management*, New York, NY, USA: ACM.
- Farach, Martin (1997), "Optimal Suffix Tree Construction with Large Alphabets" (<http://www.cs.rutgers.edu/~farach/pubs/Suffix.pdf>), *38th IEEE Symposium on Foundations of Computer Science (FOCS '97)*, pp. 137–143.
- Farach-Colton, Martin; Ferragina, Paolo; Muthukrishnan, S. (2000), "On the sorting-complexity of suffix tree construction.", *Journal of the ACM* **47** (6): 987–1011, doi:10.1145/355541.355547.
- Giegerich, R.; Kurtz, S. (1997), "From Ukkonen to McCreight and Weiner: A Unifying View of Linear-Time Suffix Tree Construction" (<http://europa.zbh.uni-hamburg.de/pubs/pdf/GieKur1997.pdf>), *Algorithmica* **19** (3): 331–353, doi:10.1007/PL00009177.
- Gusfield, Dan (1999), *Algorithms on Strings, Trees and Sequences: Computer Science and Computational Biology*, Cambridge University Press, ISBN 0-521-58519-8.
- Mansour, Essam; Allam, Amin; Skiadopoulos, Spiros; Kalnis, Panos (2011), "ERA: Efficient Serial and Parallel Suffix Tree Construction for Very Long Strings" ([http://www.vldb.org/pvldb/vol5/p049\\_essammansour\\_vldb2012.pdf](http://www.vldb.org/pvldb/vol5/p049_essammansour_vldb2012.pdf)), *PVLDB* **5** (1): 49–60.
- McCreight, Edward M. (1976), "A Space-Economical Suffix Tree Construction Algorithm", *Journal of the ACM* **23** (2): 262–272, doi:10.1145/321941.321946.
- Phoophakdee, Benjarath; Zaki, Mohammed J. (2007), "Genome-scale disk-based suffix tree indexing", *SIGMOD '07: Proceedings of the ACM SIGMOD International Conference on Management of Data*, New York, NY, USA: ACM, pp. 833–844.
- Smyth, William (2003), *Computing Patterns in Strings*, Addison-Wesley.
- Tata, Sandeep; Hankins, Richard A.; Patel, Jignesh M. (2003), "Practical Suffix Tree Construction", *VLDB '03: Proceedings of the 30th International Conference on Very Large Data Bases*, Morgan Kaufmann, pp. 36–47.
- Ukkonen, E. (1995), "On-line construction of suffix trees" (<http://www.cs.helsinki.fi/u/ukkonen/SuffixT1withFigs.pdf>), *Algorithmica* **14** (3): 249–260, doi:10.1007/BF01206331.
- Weiner, P. (1973), "Linear pattern matching algorithm", *14th Annual IEEE Symposium on Switching and Automata Theory*, pp. 1–11, doi:10.1109/SWAT.1973.13.
- Zamir, Oren; Etzioni, Oren (1998), "Web document clustering: a feasibility demonstration", *SIGIR '98: Proceedings of the 21st annual international ACM SIGIR conference on Research and development in*

*information retrieval*, New York, NY, USA: ACM, pp. 46–54.

## External links

- Suffix Trees (<http://www.cise.ufl.edu/~sahni/dsaaj/enrich/c16/suffix.htm>) by Sartaj Sahni
- Suffix Trees (<http://www.allisons.org/ll/AlgDS/Tree/Suffix/>) by Lloyd Allison
- NIST's Dictionary of Algorithms and Data Structures: Suffix Tree (<http://www.nist.gov/dads/HTML/suffixtree.html>)
- suffix\_tree ([http://mila.cs.technion.ac.il/~yona/suffix\\_tree/](http://mila.cs.technion.ac.il/~yona/suffix_tree/)) ANSI C implementation of a Suffix Tree
- libstree (<http://www.cl.cam.ac.uk/~cpk25/libstree/>), a generic suffix tree library written in C
- Tree::Suffix (<http://search.cpan.org/dist/Tree-Suffix/>), a Perl binding to libstree
- Strmat (<http://www.cs.ucdavis.edu/~gusfield/strmat.html>) a faster generic suffix tree library written in C (uses arrays instead of linked lists)
- SuffixTree ([http://hkn.eecs.berkeley.edu/~dyoo/python/suffix\\_trees/](http://hkn.eecs.berkeley.edu/~dyoo/python/suffix_trees/)) a Python binding to Strmat
- Universal Data Compression Based on the Burrows-Wheeler Transformation: Theory and Practice (<http://www.balkenhol.net/papers/t1043.pdf.gz>), application of suffix trees in the BWT
- Theory and Practice of Succinct Data Structures (<http://www.cs.helsinki.fi/group/suds/>), C++ implementation of a compressed suffix tree]
- Practical Algorithm Template Library (<http://code.google.com/p/patl/>), a C++ library with suffix tree implementation on PATRICIA trie, by Roman S. Klyujkov
- A Java implementation ([http://en.literateprograms.org/Suffix\\_tree\\_\(Java\)](http://en.literateprograms.org/Suffix_tree_(Java)))
- A Java implementation of Concurrent Suffix Tree (<http://code.google.com/p/concurrent-trees/>)

## Suffix array

In computer science, a **suffix array** is an array of integers giving the starting positions of suffixes of a string in lexicographical order.

### Details

Consider the string

1	2	3	4	5	6	7	8	9	10	11	12
a	b	r	a	c	a	d	a	b	r	a	\$

of length 12, that ends with a sentinel letter \$, appearing only once, and that is less (in lexicographical order) than any other letter in the string.

It has twelve suffixes: "abracadabra\$", "bracadabra\$", "racadabra\$", and so on down to "a\$" and "\$" that can be sorted into lexicographical order to obtain:

index	sorted suffix	lcp
12	\$	0
11	a\$	0
8	abra\$	1
1	abracadabra\$	4
4	acadabra\$	1
6	adabra\$	1
9	bra\$	0
2	bracadabra\$	3
5	cadabra\$	0
7	dabra\$	0
10	ra\$	0
3	racadabra\$	2

If the original string is available, each suffix can be completely specified by the index of its first character. The suffix array is the array of the indices of suffixes sorted in lexicographical order. For the string "abracadabra\$", using one-based indexing, the suffix array is  $\{12, 11, 8, 1, 4, 6, 9, 2, 5, 7, 10, 3\}$ , because the suffix "\$" begins at position 12, "a\$" begins at position 11, "abra\$" begins at position 8, and so forth.

The longest common prefix is also shown above as lcp. This value, stored alongside the list of prefix indices, indicates how many characters a particular suffix has in common with the suffix directly above it, starting at the beginning of both suffixes. The lcp is useful in making some string operations more efficient. For example, it can be used to avoid comparing characters that are already known to be the same when searching through the list of suffixes. The fact that the minimum lcp value belonging to a consecutive set of sorted suffixes gives the longest common prefix among all of those suffixes can also be useful.

## Algorithms

The easiest way to construct a suffix array is to use an efficient comparison sort algorithm. This requires  $O(n \log n)$  suffix comparisons, but a suffix comparison requires  $O(n)$  time, so the overall runtime of this approach is  $O(n^2 \log n)$ . More sophisticated algorithms improve this to  $O(n \log n)$  by exploiting the results of partial sorts to avoid redundant comparisons. Several  $\Theta(n)$  algorithms (of Pang Ko and Srinivas Aluru, Juha Kärkkäinen and Peter Sanders, etc.) have also been developed which provide faster construction and have space usage of  $O(n)$  with low constants. These latter are derived from the suffix tree construction algorithm of Farach. Recent work by Salson *et al.* proposes an algorithm for updating the suffix array of a text that has been edited instead of rebuilding a new suffix array from scratch. Even if the theoretical worst-case time complexity is  $O(n \log n)$ , it appears to perform well in practice: experimental results from the authors showed that their implementation of dynamic suffix arrays is generally more efficient than rebuilding when considering the insertion of a reasonable number of letters in the original text (Léonard, Mouchard and Salson).

## Applications

The suffix array of a string can be used as an index to quickly locate every occurrence of a substring within the string. Finding every occurrence of the substring is equivalent to finding every suffix that begins with the substring. Thanks to the lexicographical ordering, these suffixes will be grouped together in the suffix array, and can be found efficiently with a binary search. If implemented straightforwardly, this binary search takes  $O(m \log n)$  time, where  $m$  is the length of the substring  $W$ . The following pseudo-code from Manber and Myers shows how to find  $W$  (or the suffix lexicographically immediately before  $W$  if  $W$  is not present) in a suffix array with indices stored in  $\text{pos}$ , starting with 1 as the first index.

```

if W <= suffixAt(pos[1]) then
    ans = 1
else if W > suffixAt(pos[n]) then
    ans = n
else
{
    L = 1, R = n
    while R-L > 1 do
    {
        M = (L + R) / 2
        if W <= suffixAt(pos[M]) then
            R = M
        else
            L = M
    }
    ans = R
}

```

To avoid redoing comparisons, extra data structures giving information about the longest common prefixes (LCPs) of suffixes are constructed, giving  $O(m + \log n)$  search time.

Suffix sorting algorithms can be used to perform the Burrows–Wheeler transform (BWT). Technically the BWT requires sorting cyclic permutations of a string, not suffixes. We can fix this by appending to the string a special end-of-string character which sorts lexicographically before every other character. Sorting cyclic permutations is then equivalent to sorting suffixes.

Suffix arrays are used to look up substrings in Example-Based Machine Translation, demanding much less storage than a full phrase table as used in Statistical machine translation.

## History

Suffix arrays were originally developed by Gene Myers and Udi Manber to reduce memory consumption compared to a suffix tree. This began the trend towards compressed suffix arrays and BWT-based compressed full-text indices.

## References

- Udi Manber and Gene Myers (1991). "Suffix arrays: a new method for on-line string searches". *SIAM Journal on Computing*, Volume 22, Issue 5 (October 1993), pp. 935–948.
- Pang Ko and Srinivas Aluru (2003). "Space efficient linear time construction of suffix arrays." In *Combinatorial Pattern Matching (CPM 03)*. LNCS 2676, Springer, 2003, pp 203–210.



- Juha Kärkkäinen and Peter Sanders (2003). "Simple linear work suffix array construction." In *Proc. 30th International Colloquium on Automata, Languages and Programming (ICALP '03)*. LNCS 2719, Springer, 2003, pp. 943–955.
- M. Farach (1997). "Optimal Suffix Tree Construction with Large Alphabets". FOCS: 137–143.
- Klaus-Bernd Schürmann and Jens Stoye (2005). "An incomplex algorithm for fast suffix array construction". In *Proceedings of the 7th Workshop on Algorithm Engineering and Experiments and the 2nd Workshop on Analytic Algorithmics and Combinatorics (ALENEX/ANALCO 2005)*, 2005, pp. 77–85.
- Mikaël Salson, Martine Léonard, Thierry Lecroq and Laurent Mouchard (2009) "Dynamic Extended Suffix Arrays", *Journal of Discrete Algorithms* 2009 <sup>[1]</sup>.
- Martine Léonard, Laurent Mouchard and Mikaël Salson (2011) "On the number of elements to reorder when updating a suffix array", *Journal of Discrete Algorithms*, 2011 <sup>[2]</sup>.

## External links

- Suffix sorting module for BWT in C code <sup>[3]</sup>
- Suffix Array Implementation in Ruby <sup>[4]</sup>
- Suffix array library and tools <sup>[5]</sup>
- Project containing various Suffix Array c/c++ Implementations with a unified interface <sup>[6]</sup>
- A fast, lightweight, and robust C API library to construct the suffix array <sup>[7]</sup>
- Suffix Array implementation in Python <sup>[8]</sup>

## References

- [1] <http://dx.doi.org/10.1016/j.jda.2009.02.007>
  - [2] <http://dx.doi.org/10.1016/j.jda.2011.01.002>
  - [3] <http://code.google.com/p/compression-code/downloads/list>
  - [4] <http://www.codeodor.com/index.cfm/2007/12/24/The-Suffix-Array/1845>
  - [5] <http://sary.sourceforge.net/index.html.en>
  - [6] <http://pizzachili.dcc.uchile.cl/>
  - [7] <http://code.google.com/p/libdivsufsort/>
  - [8] <http://code.google.com/p/pysuffix/>
-

# Compressed suffix array

---

In computer science, a **compressed suffix array**<sup>[1][2]</sup> is a compressed data structure for pattern matching. Given a text  $T$  of  $n$  characters from an alphabet  $\Sigma$ , the compressed suffix array support searching for arbitrary patterns in  $T$ . For an input pattern  $P$  of  $m$  characters, the search time is equal to  $n$  times the higher-order entropy of the text  $T$ , plus some extra bits to store the empirical statistical model plus  $o(n)$ .

The original instantiation of the compressed suffix array<sup>[1]</sup> solved a long-standing open problem by showing that fast pattern matching was possible using only a linear-space data structure, namely, one proportional to the size of the text  $T$ , which takes  $O(n \log |\Sigma|)$  bits. The conventional suffix array and suffix tree use  $\Omega(n \log n)$  bits, which is substantially larger. The basis for the data structure is a recursive decomposition using the "neighbor function," which allows a suffix array to be represented by one of half its length. The construction is repeated multiple times until the resulting suffix array uses a linear number of bits. Following work showed that the actual storage space was related to the zeroth-order entropy and that the index supports self-indexing.<sup>[3]</sup> The space bound was further improved achieving the ultimate goal of higher-order entropy; the compression is obtained by partitioning the neighbor function by high-order contexts, and compressing each partition with a wavelet tree.<sup>[2]</sup> The space usage is extremely competitive in practice with other state-of-the-art compressors,<sup>[4]</sup> and it also supports fast pattern matching.

The memory accesses made by compressed suffix arrays and other compressed data structures for pattern matching are typically not localized, and thus these data structures have been notoriously hard to design efficiently for use in external memory. Recent progress using geometric duality takes advantage of the block access provided by disks to speed up the I/O time significantly<sup>[5]</sup>

## References

- [1] R. Grossi and J. S. Vitter, Compressed Suffix Arrays and Suffix Trees, with Applications to Text Indexing and String Matching, *SIAM Journal on Computing*, 35(2), 2005, 378-407. An earlier version appeared in *Proceedings of the 32nd ACM Symposium on Theory of Computing*, May 2000, 397-406.
  - [2] R. Grossi, A. Gupta, and J. S. Vitter, High-Order Entropy-Compressed Text Indexes, *Proceedings of the 14th Annual SIAM/ACM Symposium on Discrete Algorithms*, January 2003, 841-850.
  - [3] K. Sadakane, Compressed Text Databases with Efficient Query Algorithms Based on the Compressed Suffix Arrays, *Proceedings of the International Symposium on Algorithms and Computation*, Lecture Notes in Computer Science, vol. 1969, Springer, December 2000, 410-421.
  - [4] L. Foschini, R. Grossi, A. Gupta, and J. S. Vitter, Indexing Equals Compression: Experiments on Suffix Arrays and Trees, *ACM Transactions on Algorithms*, 2(4), 2006, 611-639.
  - [5] W.-K. Hon, R. Shah, S. V. Thankachan, and J. S. Vitter, On Entropy-Compressed Text Indexing in External Memory, *Proceedings of the Conference on String Processing and Information Retrieval*, August 2009.
-

# FM-index

In computer science, an **FM-index** is a compressed full-text substring index based on the Burrows-Wheeler transform, with some similarities to the suffix array. It was created by Paolo Ferragina and Giovanni Manzini<sup>[1]</sup>, who describe it as an opportunistic data structure as it allows compression of the input text while still permitting fast substring queries. The name stands for Full-text index in Minute space<sup>[2]</sup>.

It can be used to efficiently find the number of occurrences of a pattern within the compressed text, as well as locate the position of each occurrence. Both the query time and storage space requirements are sublinear with respect to the size of the input data.

The original authors have devised improvements to their original approach and dubbed it "FM-Index version 2".<sup>[3]</sup> A further improvement, the alphabet-friendly FM-index, combines the use of compression boosting and wavelet trees<sup>[4]</sup> to significantly reduce the space usage for large alphabets.

The FM-index has found use in, among other places, bioinformatics<sup>[5]</sup>

## Background

Using an index is a common strategy to efficiently search a large body of text. When the text is larger than what reasonably fits within a computer's main memory, there is a need to compress not only the text but also the index. When the FM-index was introduced, there were several suggested solutions that were based on traditional compression methods and tried to solve the compressed matching problem. In contrast, the FM-index is a compressed self-index, which means that compresses the data and indexes it at the same time.

## FM-index data structure

An FM-index is created by first taking the Burrows-Wheeler transform (BWT) of the input text. For example, the BWT of the string  $T = \text{"abracadabra"}$  is  $\text{"ard\$rcaaaabb"}$ , and here it is represented by the matrix  $M$  where each row is a rotation of the text that has been sorted. The transform corresponds to the last column labeled  $L$ .

<b>F</b>	<b>L</b>
\$ abracadabr	a
a \$abracadab	r
a bra\$abraca	d
a bracadabra	\$
a cadabra\$ab	r
a dabra\$abra	c
b ra\$abracad	a
b racadabra\$	a
c adabra\$abr	a
d abra\$abrac	a
r a\$abracada	b
r acadabra\$a	b

The BWT in itself allows for some compression with, for instance, move-to-front and Huffman encoding, but the transform has even more uses. The rows in the matrix are essentially the sorted suffixes of the text and the first column  $F$  of the matrix shares similarities with suffix arrays. How the suffix array relates to the BWT lies at the heart

of the FM-index.

### C[c] of "ard\$rcaaaabb"

c	\$	a	b	c	d	r
C[c]	0	1	6	8	9	10

It is possible to make a last-to-first column mapping  $LF(i)$  from a character  $L[i]$  to  $F[j]$ , with the help of a table  $C[c]$  and a function  $Occ(c, k)$ .  $C[c]$  is a table that, for each character  $c$  in the alphabet, contains the number of occurrences of lexically smaller characters in the text. The function  $Occ(c, k)$  is the number of occurrences of character  $c$  in the prefix  $L[1..k]$ . Ferragina and Manzini showed<sup>[1]</sup> that it is possible to compute  $Occ(c, k)$  in constant time.

### Occ(c, k) of "ard\$rcaaaabb"

	a	r	d	\$	r	c	a	a	a	a	b	b
	1	2	3	4	5	6	7	8	9	10	11	12
\$	0	0	0	1	1	1	1	1	1	1	1	1
a	1	1	1	1	1	1	2	3	4	5	5	5
b	0	0	0	0	0	0	0	0	0	0	1	2
c	0	0	0	0	0	1	1	1	1	1	1	1
d	0	0	1	1	1	1	1	1	1	1	1	1
r	0	1	1	1	2	2	2	2	2	2	2	2

The last-to-first mapping can now be defined as  $LF(i) = C[L[i]] + Occ(L[i], i)$ . For instance, on row 9,  $L$  is a and the same a can be found on row 5 in the first column  $F$ , so  $LF(9)$  should be 5 and  $LF(9) = C[a] + Occ(a, 9) = 5$ . For any row  $i$  of the matrix, the character in the last column  $L[i]$  precedes the character in the first column  $F[i]$  also in  $T$ . Finally, if  $L[i] = T[k]$ , then  $L[LF(i)] = T[k - 1]$ , and using the equality it is possible to extract a string of  $T$  from  $L$ .

The FM-index itself a compression of the string  $L$  together with  $C$  and  $Occ$  in some form, as well as information that maps a selection of indices in  $L$  to positions in the original string  $T$ .

## Count

The operation *count* takes a pattern  $P[1..p]$  and returns the number of occurrences of that pattern in the original text  $T$ . Since the rows of matrix  $M$  is sorted, and it contains every suffix of  $T$ , the occurrences of pattern  $P$  will be next to each other in a single continuous range. The operation iterates backwards over the pattern. For every character in the pattern, the range that has the character as a suffix is found. For example, the count of the pattern "bra" in "abracadabra" follows these steps:

1. The first character we look for is a, the last character in the pattern. The initial range is set to  $[C[a] + 1..C[a+1]] = [2..6]$ . This range over  $L$  represents every character of  $T$  that has a suffix beginning with  $a$ .
2. The next character to look for is r. The new range is  $[C[r] + Occ(r, start-1) + 1..C[r] + Occ(r, end)] = [10 + 0 + 1..10 + 2] = [11..12]$ , if start is the index of the beginning of the range and end is the end. This range over  $L$  is all the characters of  $T$  that have suffixes beginning with  $ra$ .
3. The last character to look at is b. The new range is  $[C[b] + Occ(b, start-1) + 1..C[b] + Occ(b, end)] = [6 + 0 + 1..6 + 2] = [7..8]$ . This range over  $L$  is all the characters that have a suffix that begins with  $bra$ . Now that the whole pattern has been processed, the count is the same as the size of the range:  $8 - 7 + 1 = 2$ .

If the range at becomes empty or the range boundaries cross each other before the whole pattern has been looked up, the pattern does not occur in  $T$ . Because  $Occ(c, k)$  can be performed in constant time, count can complete in  $O(p)$

time.

## Locate

The operation *locate* takes as input an index of a character in  $L$  and returns its position  $i$  in  $T$ . For instance  $\text{locate}(7) = 8$ . To locate every occurrence of a pattern, first the range of character is found whose suffix is the pattern in the same way the *count* operation found the range. Then the position of every character in the range can be located.

To map an index in  $L$  to one in  $T$ , a subset of the indices in  $L$  are associated with a position in  $T$ . If  $L[j]$  has a position associated with it,  $\text{locate}(j)$  is trivial. If it's not associated, the string is followed with  $\text{LF}(i)$  until an associated index is found. By associating a suitable number of indices, an upper bound can be found. *Locate* can be implemented to find  $occ$  occurrences of a pattern  $P[1..p]$  in a text  $T[1..u]$  in  $O(p + occ \log^\epsilon u)$  time with  $O(H_k(T) + \frac{\log \log u}{\log^\epsilon u})$  bits per input symbol for any  $k \geq 0$  <sup>[1]</sup>.

## References

- [1] Paolo Ferragina and Giovanni Manzini (2000). "Opportunistic Data Structures with Applications". Proceedings of the 41st Annual Symposium on Foundations of Computer Science. p.390.
- [2] Paolo Ferragina and Giovanni Manzini (2005). "Indexing Compressed Text". Journal of the ACM (JACM), 52, 4 (Jul. 2005). p. 553
- [3] Paolo Ferragina and Rossano Venturini "FM-Index version 2" (<http://www.di.unipi.it/~ferragin/Libraries/fmindexV2/index.html>)
- [4] P. Ferragina, G. Manzini, V. Mäkinen and G. Navarro. An Alphabet-Friendly FM-index. *In Proc. SPIRE'04*, pages 150-160. LNCS 3246.
- [5] Simpson, Jared T. and Durbin, Richard (2010). "Efficient construction of an assembly string graph using the FM-index". *Bioinformatics*, 26, 12 (Jun. 17). p. i367

# Generalised suffix tree

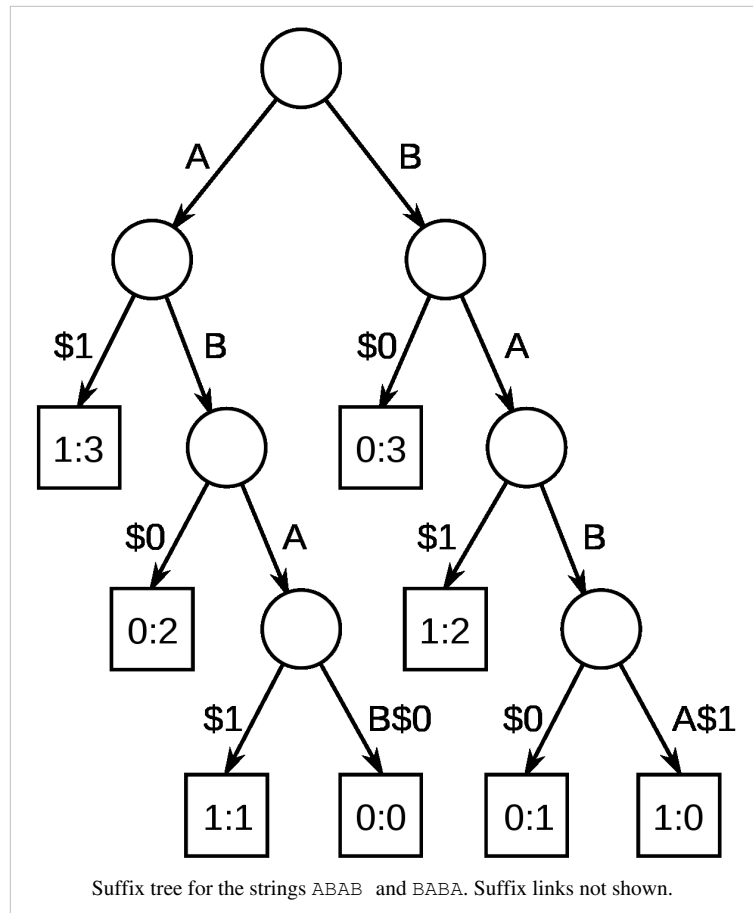
In computer science, a **generalised suffix tree** is a suffix tree for a set of strings. Given the set of strings  $D = S_1, S_2, \dots, S_d$  of total length  $n$ , it is a Patricia tree containing all  $n$  suffixes of the strings. It is mostly used in bioinformatics.<sup>BRCR</sup>

## Functionality

It can be built in  $\Theta(n)$  time and space, and can be used to find all  $z$  occurrences of a string  $P$  of length  $m$  in  $O(m + z)$  time, which is asymptotically optimal (assuming the size of the alphabet is constant, see <sup>Gus97</sup> page 119).

When constructing such a tree, each string should be padded with a unique out-of-alphabet marker symbol (or string) to ensure no suffix is a substring of another, guaranteeing each suffix is represented by a unique leaf node.

Algorithms for constructing a GST include Ukkonen's algorithm and McCreight's algorithm.



## Example

A suffix tree for the strings ABAB and BABA is shown in a figure above. They are padded with the unique terminator strings \$0 and \$1. The numbers in the leaf nodes are string number and starting position. Notice how a left to right traversal of the leaf nodes corresponds to the sorted order of the suffixes. The terminators might be strings or unique single symbols. Edges on \$ from the root are left out in this example.

## Alternatives

An alternative to building a generalised suffix tree is to concatenate the strings, and build a regular suffix tree or suffix array for the resulting string. When hits are evaluated after a search, global positions are mapped into documents and local positions with some algorithm and/or data structure, such as a binary search in the starting/ending positions of the documents.

## References

- Lucas Chi Kwong Hui (1992). "Color Set Size Problem with Applications to String Matching"<sup>[1]</sup>. *Combinatorial Pattern Matching, Lecture Notes in Computer Science*, 644.. pp. 230–243.
- Paul Bieganski, John Riedl, John Carlis, and Ernest F. Retzel (1994). "Generalized Suffix Trees for Biological Sequence Data"<sup>[2]</sup>. *Biotechnology Computing, Proceedings of the Twenty-Seventh Hawaii International*

*Conference on..* pp. 35–44.

- Gusfield, Dan (1999) [1997]. *Algorithms on Strings, Trees and Sequences: Computer Science and Computational Biology*. USA: Cambridge University Press. ISBN 0-521-58519-8.

## References

[1] <http://www.springerlink.com/content/y565487707522555/>

[2] [http://ieeexplore.ieee.org/xpl/freeabs\\_all.jsp?arnumber=323593](http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=323593)

## B-trie

The **B-trie** is a trie-based data structure that can store and retrieve variable-length strings efficiently on disk.<sup>[1]</sup>

The B-trie was compared against several high-performance variants of B-tree that were designed for string keys. It was shown to offer superior performance, particularly under skew access (i.e., many repeated searches). It is currently a leading choice for maintaining a string dictionary on disk, along with other disk-based tasks, such as maintaining an index to a string database or for accumulating the vocabulary of a large text collection.

## References

- [1] Askitis, Nikolas; Zobel, Justin (2008), "B-tries for Disk-based String Management" (<http://www.springerlink.com/content/x7545u2g85675u17/>), *VLDB Journal*: 1–26, ISBN 1066-8888 (Print) 0949-877X (Online),

## Judy array

In computer science and software engineering, a **Judy array** is a data structure that has high performance, low memory usage and implements an associative array. Unlike normal arrays, Judy arrays may be sparse, that is, they may have large ranges of unassigned indices. They can be used for storing and looking up values using integer or string keys. The key benefits of using Judy is its scalability, high performance, memory efficiency and ease of use.<sup>[1]</sup>

Judy arrays are both speed- and memory-efficient, with no tuning or configuration required and therefore they can replace common data structures (skiplists, linked lists, binary, ternary, b-trees, hashing) and work better with very large data sets.

Roughly speaking, it is similar to a highly-optimised 256-ary trie data structure.<sup>[2]</sup> To make memory consumption small, Judy arrays use over 20 different compression techniques to compress trie nodes.

The Judy array was invented by Douglas Baskins and named after his sister.<sup>[3]</sup>

## Terminology

Expanse, population and density are commonly used when it comes to Judy. As they are not commonly used in tree search literature, it is important to define them:

1. *Expanse* is a range of possible keys, e.g. 200, 300, etc.
2. *Population* is the count of keys contained in an expanse, e.g. 200, 360, 400, 512, 720 = 5
3. *Density* is used to describe the sparseness of an expanse of keys: Density = Population/Expanse

## Benefits

### Memory allocation

Judy arrays are designed to be unbounded arrays and therefore their sizes are not pre-allocated. They can dynamically choose to grow or shrink the memory used according to the population of the array and can scale to a large number of elements. Since it allocates memory dynamically as it grows, it is only bounded by machine memory.<sup>[4]</sup> The memory used by Judy is nearly proportional to the number of elements (population) in the Judy array.

### Speed

Judy arrays are designed to keep the number of processor cache-line fills as low as possible, and the algorithm is internally complex in an attempt to satisfy this goal as often as possible. Due to these cache optimizations, Judy arrays are fast, sometimes even faster than a hash table, especially for very big datasets. Despite Judy arrays being a type of trie, they consume much less memory than hash tables. Also because a Judy array is a trie, it is possible to do an ordered sequential traversal of keys, which is not possible in hash tables.

## Drawbacks

Judy arrays appear to be the subject of US patent 6735595.<sup>[5]</sup>

## References

- [1] <http://packages.debian.org/lenny/libjudy-dev>
- [2] Alan Silverstein, " Judy IV Shop Manual ([http://judy.sourceforge.net/application/shop\\_interim.pdf](http://judy.sourceforge.net/application/shop_interim.pdf))", 2002
- [3] <http://judy.sourceforge.net/>
- [4] Advances in databases: concepts, systems and applications : By Kotagiri Ramamohanarao
- [5] <http://www.google.com/patents/US6735595>

## External links

- Main Judy arrays site (<http://judy.sourceforge.net/>)
- How Judy arrays work and why they are so fast (<http://judy.sourceforge.net/downloads/10minutes.htm>)
- A complete technical description of Judy arrays ([http://judy.sourceforge.net/application/shop\\_interim.pdf](http://judy.sourceforge.net/application/shop_interim.pdf))
- An independent performance comparison of Judy to Hash Tables (<http://www.nothings.org/computer/judy/>)
- A compact implementation of Judy arrays in 1K lines of C code (<http://code.google.com/p/judyarray>)



# Directed acyclic word graph

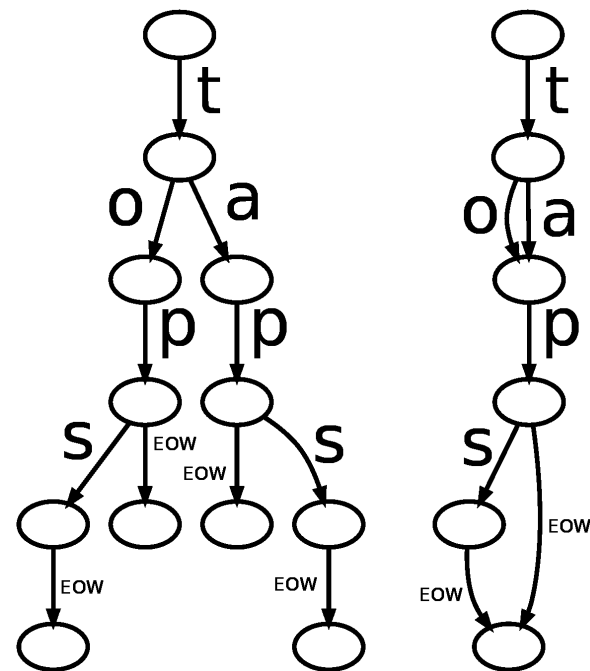
In computer science, a **directed acyclic word graph** (sometimes abbreviated as **DAWG**) is a data structure that represents a set of strings, and allows for a query operation that tests whether a given string belongs to the set in time proportional to its length. In these respects, a DAWG is very similar to a trie, but it is much more space efficient.

A DAWG is represented as a directed acyclic graph with a single source vertex (a vertex with no incoming edges), in which each edge of the graph is labeled by a letter, symbol, or special end-of-string marker, and in which each vertex has at most one outgoing edge for each possible letter or symbol. The strings represented by the DAWG are formed by the symbols on paths in the DAWG from the source vertex to any sink vertex (a vertex with no outgoing edges). A DAWG can also be interpreted as an acyclic finite automaton that accepts the words that are stored in the DAWG.

Thus, a trie (a rooted tree with the same properties of having edges labeled by symbols and strings formed by root-to-leaf paths) is a special kind of DAWG. However, by allowing the same vertices to be reached by multiple paths, a DAWG may use significantly fewer vertices than a trie. Consider, for example, the four English words "tap", "taps", "top", and "tops". A trie for those four words would have 11 vertices, one for each of the strings formed as a prefix of one of these words, or for one of the words followed by the end-of-string marker. However, a DAWG can represent these same four words using only six vertices  $v_i$  for  $0 \leq i \leq 5$ , and the following edges: an edge from  $v_0$  to  $v_1$  labeled "t", two edges from  $v_1$  to  $v_2$  labeled "a" and "o", an edge from  $v_2$  to  $v_3$  labeled "p", an edge  $v_3$  to  $v_4$  labeled "s", and edges from  $v_3$  and  $v_4$  to  $v_5$  labeled with the end-of-string marker.

The primary difference between DAWG and trie is the elimination of suffix redundancy in storing strings. The trie eliminates prefix redundancy since all common prefixes are shared between strings, such as between *doctors* and *doctorate* the *doctor* prefix is shared. In a DAWG common suffixes are also shared, such as between *desertion* and *destruction* both the prefix *des-* and suffix *-tion* are shared. For dictionary sets of common English words, this translates into major memory usage reduction.

Because the terminal nodes of a DAWG can be reached by multiple paths, a DAWG cannot directly store auxiliary information relating to each path, e.g. a word's frequency in the English language. However, if at each node we store a count of the number of unique paths through the structure from that point, we can use it to retrieve the index of a word, or a word given its index.<sup>[1]</sup> The auxiliary information can then be stored in an array.



The strings "tap", "taps", "top", and "tops" stored in a Trie (left) and a DAWG (right), EOW stands for End-of-word.

## References

- Blumer, A. and Blumer, J. and Haussler, D. and Ehrenfeucht, A. and Chen, M.T. and Seiferas, J. (1985), "The smallest automation recognizing the subwords of a text", *Theoretical computer science* **40**: 31--55
- Appel, Andrew; Jacobsen, Guy (1988), *one of the early mentions of the data structure* (<http://www.cs.cmu.edu/afs/cs/academic/class/15451-s06/www/lectures/scrabble.pdf>), "The World's Fastest Scrabble Program" (PDF), *Communications of the ACM*.
- Crochemore, Maxime; V  rin, Renaud (1997), "Direct construction of compact directed acyclic word graphs", *Combinatorial Pattern Matching*, Lecture Notes in Computer Science, Springer-Verlag, pp. 116–129, doi:10.1007/3-540-63220-4\_55.
- Inenaga, S.; Hoshino, H.; Shinohara, A.; Takeda, M.; Arikawa, S. (2001), "On-line construction of symmetric compact directed acyclic word graphs" ([http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=989743](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=989743)), *Proc. 8th Int. Symp. String Processing and Information Retrieval, 2001. SPIRE 2001*, pp. 96–110, doi:10.1109/SPIRE.2001.989743, ISBN 0-7695-1192-9.
- Jansen, Cees J. A.; Boekee, Dick E. (1990), "On the significance of the directed acyclic word graph in cryptology", *Advances in Cryptology — AUSCRYPT '90*, Lecture Notes in Computer Science, **453**, Springer-Verlag, pp. 318–326, doi:10.1007/BFb0030372, ISBN 3-540-53000-2.
- Do, H.H.; Sung, W.K. (2011), "Compressed Directed Acyclic Word Graph with Application in Local Alignment", *Computing and Combinatorics*, Lecture Notes in Computer Science, **6842**, Springer-Verlag, pp. 503-518, doi:10.1007/978-3-642-22685-4\_44, ISBN 978-3-642-22684-7

## External links

- National Institute of Standards and Technology (<http://www.nist.gov/dads/HTML/directedAcyclicWordGraph.html>)
  - DAWG implementation in C# by Samuel Allen (<http://dotnetperls.com/directed-acyclic-word-graph>)
  - Optimal DAWG Creation Step By Step Treatment (<http://www.pathcom.com/~vadco/dawg.html>)
  - Documentation for The World's Most Powerful DAWG Encoding: Caroline Word Graph (<http://www.pathcom.com/~vadco/cwg.html>)
-

---

# Multiway trees

---

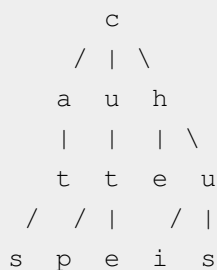
## Ternary search tree

---

In computer science, a **ternary search tree** is a special trie data structure where the child nodes of a standard trie are ordered as a binary search tree. Ternary search trees are effective for many applications involving mapping strings to other values as well as completing near-neighbor lookups and other string-related queries. In addition, ternary search trees are a more space efficient (at the cost of speed) equivalent of typical tries. Further, **ternary search trees** can be used any time a hashtable would be used to store strings.

### Description

A ternary search tree is similar to other tries.<sup>[1]</sup> Each node of a ternary search tree stores a single character, an object (or a pointer to an object depending on implementation), and pointers to its three children conventionally named "equal kid" "lo kid" and "hi kid."<sup>[1][2]</sup> A node may also have a pointer to its parent node as well as an indicator as to whether or not the node marks the end of a word<sup>[1]</sup>. The lo kid pointer must point to a node whose character value is less than the current node. Conversely, the hi kid pointer must point to a node whose character is greater than the current node.<sup>[2]</sup> The figure below shows a ternary search tree with the strings "as", "at", "cup", "cute", "he", "i" and "us":



As with other trie data structures, each node in a ternary search tree represents a prefix of the stored strings. All strings in the middle subtree of a node start with that prefix.

### Comparison to other data structures

#### Tries

When compared to tries, ternary search trees run slower. However, tries are often unacceptable solutions for large sets of data because of their large memory footprint.<sup>[2]</sup> For this reason, ternary search trees can be used as a slower, but more space-efficient alternative.<sup>[2]</sup>

#### Hash maps

Hashtables can also be used in place of ternary search trees for mapping strings to values. However, hash maps also frequently use more memory than ternary search trees (but not as much as tries). Additionally, hash maps are typically slower at reporting a string that is not in the same data structure because it must compare the entire string rather than just the first few characters. There is some evidence that shows ternary search trees running faster than hash maps.<sup>[2]</sup> Additionally, hash maps do not allow for many of the uses of ternary search trees such as near-neighbor lookups.

---

## Ternary search tree operations

### Look up

To look up a particular node or the data associated with a node, a string key is needed. A lookup procedure begins by checking the root node of the tree and determining which of the following conditions has occurred. If the first character of the string is less than the character in the root node, a recursive lookup can be called on the tree whose root is the lo kid of the current root. Similarly, if the first character is greater than the current node in the tree, then a recursive call can be made to the tree whose root is the hi kid of the current node.<sup>[2]</sup> As a final case, if the first character of the string is equal to the character of the current node then the function returns the node if there are no more characters in the key. If there are more characters in the key then the first character of the key must be removed and a recursive call is made given the equal kid node and the modified key.<sup>[2]</sup> This can also be written in a non-recursive way by using a pointer to the current node and a pointer to the current character of the key.<sup>[2]</sup>

### Insertion

Inserting a value into a ternary search can be defined recursively much as lookups are defined. This recursive method is continually called on nodes of the tree given a key which gets progressively shorter by pruning characters off the front of the key. If this method reaches a node that has not been created, it creates the node and assigns it the character value of the first character in the key. Whether a new node is created or not, the method checks to see if the first character in the string is greater than or less than the character value in the node and makes a recursive call on the appropriate node as in the lookup operation. If, however, the key's first character is equal to the node's value then the insertion procedure is called on the equal kid and the key's first character is pruned away.<sup>[2]</sup> Like binary search trees and other data structures, ternary search trees can become degenerate depending on the order of the keys.<sup>[3]</sup> Inserting keys in order is one way to attain the worst possible degenerate tree.<sup>[2]</sup> Inserting the keys in random order often produces a well-balanced tree.<sup>[2]</sup>

### Running Time

The running time of ternary search trees varies significantly with the input. For this reason it is important to use them only when the situation lends itself to their use. Ternary search trees run best when given several similar strings, especially when those strings share a common prefix. Alternatively, ternary search trees are effective when storing a large number of relatively short strings (such as words in a dictionary).<sup>[2]</sup> Running times for binary search trees are similar to binary search trees in that they typically run in logarithmic time but, in the worst case, when degenerate trees are formed, can run in linear time.

Time complexities for ternary search tree operations:<sup>[2]</sup>

	Average-Case Running Time	Worst-Case Running Time
Lookup	$O(\log n)$	$O(n)$
Insertion	$O(\log n)$	$O(n)$
Delete	$O(\log n)$	$O(n)$

## Uses

Ternary search trees can be used to solve many problems in which a large number of strings must be stored and retrieved in an arbitrary order. Some of the most common or most useful of these are below:

- Anytime a trie could be used but a less memory-consuming structure is preferred. <sup>[2]</sup>
- A quick and space-saving data structure for mapping strings to other data. <sup>[3]</sup>
- To implement an auto-complete feature. <sup>[1]</sup>
- As a spell check <sup>[4]</sup>
- Near-neighbor searching (Of which a spell-check is a special case) <sup>[2]</sup>
- As a database especially when indexing by several non-key fields is desirable <sup>[4]</sup>
- In place of a hash table. <sup>[4]</sup>

## References

- [1] Ostrovsky, Igor. "Efficient auto-complete with a ternary search tree" (<http://igoro.com/archive/efficient-auto-complete-with-a-ternary-search-tree/>). .
- [2] Dobbs. "Ternary Search Trees" (<http://www.drdobbs.com/database/184410528>). .
- [3] Wrobel, Lukasz. "Ternary Search Tree" (<http://lukaszwrobel.pl/blog/ternary-search-tree>). .
- [4] Flint, Wally. "Plant your data in a ternary search tree" (<http://www.javaworld.com/javaworld/jw-02-2001/jw-0216-ternary.html>). .

## External links

- Ternary Search Trees (<http://www.cs.princeton.edu/~rs/strings/>)
- Tree::Ternary (Perl module) (<http://search.cpan.org/~mrogaski/Tree-Ternary-0.03/Ternary.pm>)
- Ternary Search Tree code (<http://dasnar.sdf-eu.org/res/ctst-README.html>)
- A key-value store implementation based on Ternary Search Tree (<http://code.google.com/p/tstddb/>)
- STL-compliant Ternary Search Tree in C++ (<http://abc.se/~re/code/tst/>)
- Ternary Search Tree in C++ (<http://ternary.sourceforge.net>)
- Ternary Search Tree in Ruby (<https://github.com/kanwei/algorithms/tree/>)
- pytst - C++ Ternary Search Tree implementation with Python bindings (<https://github.com/nlehuen/pytst/>)
- Algorithm for generating search strings given a Ternary Search Tree (<http://stackoverflow.com/questions/8143527/is-it-possible-to-generate-all-possible-terms-findable-in-a-ternary-search-tree>)
- Python Implementation Ternary Search Tree (<http://hacktalks.blogspot.in/2012/03/implementing-auto-complete-with-ternary.html>)
- Java Concurrent Ternary Search Tree (<http://trasahin.blogspot.co.uk/2012/06/concurrent-ternary-search-tree.html>)

# And-or tree

An **and-or tree** is a graphical representation of the reduction of problems (or goals) to conjunctions and disjunctions of subproblems (or subgoals).

## Example

The and-or tree:

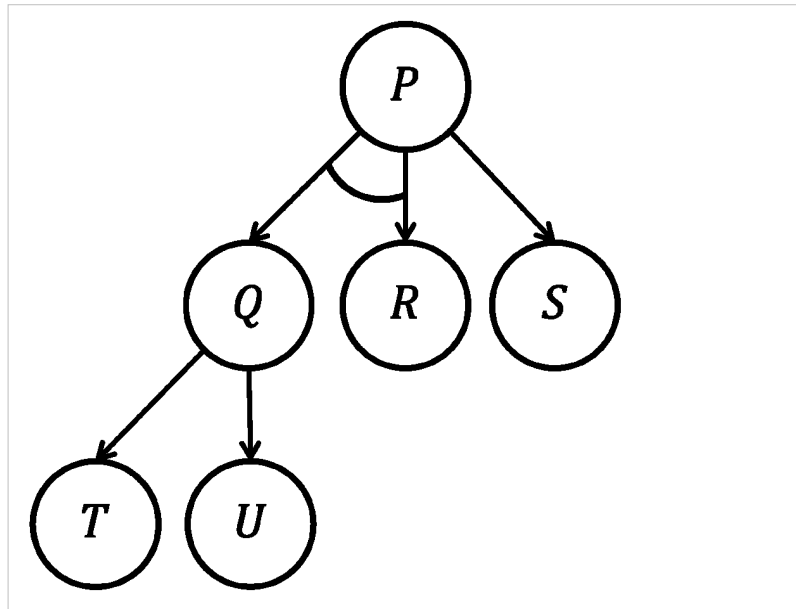
represents the search space for solving the problem  $P$ , using the goal-reduction methods:

$P$  if  $Q$  and  $R$

$P$  if  $S$

$Q$  if  $T$

$Q$  if  $U$



## Definitions

Given an initial problem  $P_0$  and set of problem solving methods of the form:

$P$  if  $P_1$  and ... and  $P_n$

the associated and-or tree is a set of labelled nodes such that:

1. The root of the tree is a node labelled by  $P_0$ .
1. For every node  $N$  labelled by a problem or sub-problem  $P$  and for every method of the form  $P$  if  $P_1$  and ... and  $P_n$ , there exists a set of children nodes  $N_1, \dots, N_n$  of the node  $N$ , such that each node  $N_i$  is labelled by  $P_i$ . The nodes are conjoined by an arc, to distinguish them from children of  $N$  that might be associated with other methods.

A node  $N$ , labelled by a problem  $P$ , is a success node if there is a method of the form  $P$  if nothing (i.e.,  $P$  is a "fact"). The node is a failure node if there is no method for solving  $P$ .

If all of the children of a node  $N$ , conjoined by the same arc, are success nodes, then the node  $N$  is also a success node. Otherwise the node is a failure node.

## Search strategies

An and-or tree specifies only the search space for solving a problem. Different search strategies for searching the space are possible. These include searching the tree depth-first, breadth-first, or best-first using some measure of desirability of solutions. The search strategy can be sequential, searching or generating one node at a time, or parallel, searching or generating several nodes in parallel.

## Relationship with logic programming

The methods used for generating and-or trees are propositional logic programs (without variables). In the case of logic programs containing variables, the solutions of conjoint sub-problems must be compatible. Subject to this complication, sequential and parallel search strategies for and-or trees provide a computational model for executing logic programs.

## Relationship with two-player games

And–or trees can also be used to represent the search spaces for two-person games. The root node of such a tree represents the problem of one of the players winning the game, starting from the initial state of the game. Given a node  $N$ , labelled by the problem  $P$  of the player winning the game from a particular state of play, there exists a single set of conjoint children nodes, corresponding to all of the opponents responding moves. For each of these children nodes, there exists a set of non-conjoint children nodes, corresponding to all of the player's defending moves.

For solving game trees with proof-number search family of algorithms, game trees are to be mapped to And/Or trees. MAX-nodes (i.e. maximizing player to move) are represented as OR nodes, MIN-nodes map to AND nodes. The mapping is possible, when the search is done with only a binary goal, which usually is "player to move wins the game".

## Bibliography

- Luger, G.F. and Stubblefield, W.A. *Artificial Intelligence: Structures and Strategies for Complex Problem Solving* (2nd Edition), The Begjamins/Cummings Publishing Company, Inc., 1993.
- Nilsson, N.J. *Artificial Intelligence, A New Synthesis*, Morgan Kaufmann Publishers, Inc., 1998.

## (a,b)-tree

---

In computer science, an **(a,b) tree** is a specific kind of search tree.

An (a,b) tree has all of its leaves at the same depth, and all internal nodes have between  $a$  and  $b$  children, where  $a$  and  $b$  are integers such that  $2 \leq a \leq (b + 1)/2$ . The root may have as few as zero children.

### Definition

Let  $a, b \in \mathbb{N}$  such that  $a \leq b$ . Then a tree  $T$  is an **(a,b) tree** when:

- Every inner node except the root has at least  $a$  and maximally  $b$  child nodes.
- Root has maximally  $b$  child nodes.
- All paths from the root to the leaves are of the same length.

### Inner node representation

Every inner node  $v$  has the following representation:

- Let  $\rho_v$  be the number of child nodes of node  $v$ .
- Let  $S_v[1 \dots \rho_v]$  be pointers to child nodes.
- Let  $H_v[1 \dots \rho_v - 1]$  be an array of keys such that  $H_v[i]$  equals the largest key in the subtree pointed to by  $S_v[i]$ .

### References

- Paul E. Black, (a,b)-tree<sup>[1]</sup> at the NIST Dictionary of Algorithms and Data Structures.

### References

[1] <http://www.nist.gov/dads/HTML/abtree.html>

# Link/cut tree

A **link/cut tree** is a type of data structure that can merge (link) and split (cut) data sets in  $O(\log(n))$  amortized time, and can find which tree an element belongs to in  $O(\log(n))$  amortized time. In the original publication, Sleator and Tarjan referred to link/cut trees as “dynamic trees”.

## Further reading

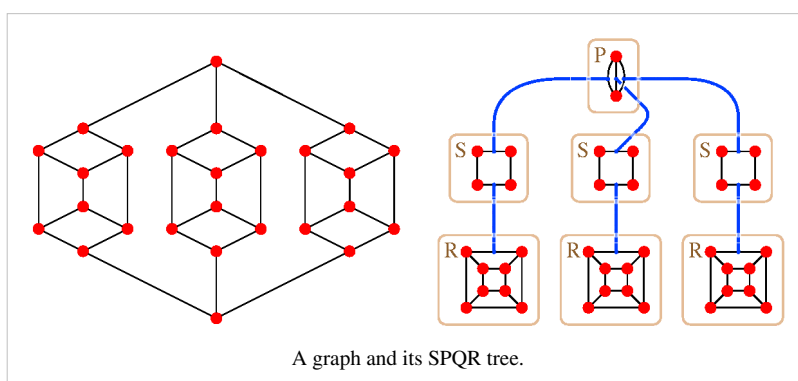
- Sleator, D. D.; Tarjan, R. E. (1983). "A Data Structure for Dynamic Trees" <sup>[1]</sup>. *Proceedings of the thirteenth annual ACM symposium on Theory of computing - STOC '81*. pp. 114. doi:10.1145/800076.802464.
- Sleator, D. D.; Tarjan, R. E. (1985). "Self-Adjusting Binary Search Trees" <sup>[2]</sup>. *Journal of the ACM* **32** (3): 652. doi:10.1145/3828.3835.
- Goldberg, A. V.; Tarjan, R. E. (1989). "Finding minimum-cost circulations by canceling negative cycles". *Journal of the ACM* **36** (4): 873. doi:10.1145/76359.76368. — Application to min-cost circulation
- <http://compgeom.cs.uiuc.edu/~jeffe/teaching/datastructures/2006/notes/07-linkcut.pdf>

## References

- [1] <http://www.cs.cmu.edu/~sleator/papers/dynamic-trees.pdf>  
 [2] <http://www.cs.cmu.edu/~sleator/papers/self-adjusting.pdf>

# SPQR tree

In graph theory, a branch of mathematics, the **triconnected components** of a biconnected graph are a system of smaller graphs that describe all of the 2-vertex cuts in the graph. An **SPQR tree** is a tree data structure used in computer science, and more specifically graph algorithms, to represent the triconnected components of a graph. The SPQR tree of a graph may be constructed in linear time<sup>[1]</sup> and has several applications in dynamic graph algorithms and graph drawing.



The basic structures underlying the SPQR tree, the triconnected components of a graph, and the connection between this decomposition and the planar embeddings of a planar graph, were first investigated by Saunders Mac Lane (1937); these structures were used in efficient algorithms by several other researchers<sup>[2]</sup> prior to their formalization as the SPQR tree by Di Battista and Tamassia (1989, 1990, 1996).



## Structure

An SPQR tree takes the form of an unrooted tree in which for each node  $x$  there is associated an undirected graph or multigraph  $G_x$ . The node, and the graph associated with it, may have one of four types, given the initials SPQR:

- In an S node, the associated graph is a cycle graph with three or more vertices and edges. This case is analogous to series composition in series-parallel graphs; the S stands for "series".<sup>[3]</sup>
- In a P node, the associated graph is a dipole graph, a multigraph with two vertices and three or more edges, the planar dual to a cycle graph. This case is analogous to parallel composition in series-parallel graphs; the P stands for "parallel".<sup>[3]</sup>
- In a Q node, the associated graph has a single edge. This trivial case is necessary to handle the graph that has only one edge, but does not appear in the SPQR trees of more complex graphs.
- In an R node, the associated graph is a 3-connected graph that is not a cycle or dipole. The R stands for "rigid": in the application of SPQR trees in planar graph embedding, the associated graph of an R node has a unique planar embedding.<sup>[3]</sup>

Each edge  $xy$  between two nodes of the SPQR tree is associated with two directed *virtual edges*, one of which is an edge in  $G_x$  and the other of which is an edge in  $G_y$ . Each edge in a graph  $G_x$  may be a virtual edge for at most one SPQR tree edge.

An SPQR tree  $T$  represents a 2-connected graph  $G_T$  formed as follows. Whenever SPQR tree edge  $xy$  associates the virtual edge  $ab$  of  $G_x$  with the virtual edge  $cd$  of  $G_y$ , form a single larger graph by merging  $a$  and  $c$  into a single supervertex, merging  $b$  and  $d$  into another single supervertex, and deleting the two virtual edges. That is, the larger graph is the 2-clique-sum of  $G_x$  and  $G_y$ . Performing this gluing step on each edge of the SPQR tree produces the graph  $G_T$ ; the order of performing the gluing steps does not affect the result. Each vertex in one of the graphs  $G_x$  may be associated in this way with a unique vertex in  $G_T$ , the supervertex into which it was merged.

Typically, it is not allowed within an SPQR tree for two S nodes to be adjacent, nor for two P nodes to be adjacent, because if such an adjacency occurred the two nodes could be merged into a single larger node. With this assumption, the SPQR tree is uniquely determined from its graph. When a graph  $G$  is represented by an SPQR tree with no adjacent P nodes and no adjacent S nodes, then the graphs  $G_x$  associated with the nodes of the SPQR tree are known as the triconnected components of  $G$ .

## Finding 2-vertex cuts

With the SPQR tree of a graph  $G$  (without Q nodes) it is straightforward to find every pair of vertices  $u$  and  $v$  in  $G$  such that removing  $u$  and  $v$  from  $G$  leaves a disconnected graph, and the connected components of the remaining graphs:

- The two vertices  $u$  and  $v$  may be the two endpoints of a virtual edge in the graph associated with an R node, in which case the two components are represented by the two subtrees of the SPQR tree formed by removing the corresponding SPQR tree edge.
- The two vertices  $u$  and  $v$  may be the two vertices in the graph associated with a P node that has two or more virtual edges. In this case the components formed by the removal of  $u$  and  $v$  are the represented by subtrees of the SPQR tree, one for each virtual edge in the node.
- The two vertices  $u$  and  $v$  may be two vertices in the graph associated with an S node such that either  $u$  and  $v$  are not adjacent, or the edge  $uv$  is virtual. If the edge is virtual, then the pair  $(u,v)$  also belongs to a node of type P and R and the components are as described above. If the two vertices are not adjacent then the two components are represented by two paths of the cycle graph associated with the S node and with the SPQR tree nodes attached to those two paths.

## Embeddings of planar graphs

If a planar graph is 3-connected, it has a unique planar embedding up to the choice of which face is the outer face and of orientation of the embedding: the faces of the embedding are exactly the nonseparating cycles of the graph. However, for a planar graph (with labeled vertices and edges) that is 2-connected but not 3-connected, there may be greater freedom in finding a planar embedding. Specifically, whenever two nodes in the SPQR tree of the graph are connected by a pair of virtual edges, it is possible to flip the orientation of one of the nodes relative to the other one. Additionally, in a P node of the SPQR tree, the different parts of the graph connected to virtual edges of the P node may be arbitrarily permuted. All planar representations may be described in this way.

## Notes

- [1] Hopcroft & Tarjan (1973); Gutwenger & Mutzel (2001).
- [2] E.g., Hopcroft & Tarjan (1973) and Bienstock & Monma (1988), both of which are cited as precedents by Di Battista and Tamassia.
- [3] Di Battista & Tamassia (1989).

## References

- Bienstock, Daniel; Monma, Clyde L. (1988), "On the complexity of covering vertices by faces in a planar graph", *SIAM Journal on Computing* **17** (1): 53–76, doi:10.1137/0217004.
- Di Battista, Giuseppe; Tamassia, Roberto (1989), "Incremental planarity testing", *Proc. 30th Annual Symposium on Foundations of Computer Science*, pp. 436–441, doi:10.1109/SFCS.1989.63515.
- Di Battista, Giuseppe; Tamassia, Roberto (1990), "On-line graph algorithms with SPQR-trees", *Proc. 17th International Colloquium on Automata, Languages and Programming*, Lecture Notes in Computer Science, **443**, Springer-Verlag, pp. 598–611, doi:10.1007/BFb0032061.
- Di Battista, Giuseppe; Tamassia, Roberto (1996), "On-line planarity testing" (<http://cs.brown.edu/research/pubs/pdfs/1996/DiBattista-1996-OPT.pdf>), *SIAM Journal on Computing* **25** (5): 956–997, doi:10.1137/S0097539794280736.
- Gutwenger, Carsten; Mutzel, Petra (2001), "A linear time implementation of SPQR-trees", *Proc. 8th International Symposium on Graph Drawing (GD 2000)*, Lecture Notes in Computer Science, **1984**, Springer-Verlag, pp. 77–90, doi:10.1007/3-540-44541-2\_8.
- Hopcroft, John; Tarjan, Robert (1973), "Dividing a graph into triconnected components", *SIAM Journal on Computing* **2** (3): 135–158, doi:10.1137/0202012.
- Mac Lane, Saunders (1937), "A structural characterization of planar combinatorial graphs", *Duke Mathematical Journal* **3** (3): 460–472, doi:10.1215/S0012-7094-37-00336-3.

## External links

- SQPR tree implementation ([http://www.ogdf.net/doc-ogdf/classogdf\\_1\\_1\\_s\\_p\\_q\\_r\\_tree.html](http://www.ogdf.net/doc-ogdf/classogdf_1_1_s_p_q_r_tree.html)) in the Open Graph Drawing Framework.
- The tree of the triconnected components Java implementation (<http://code.google.com/p/jbpt/>) in the jBPT library (see TCTree class).

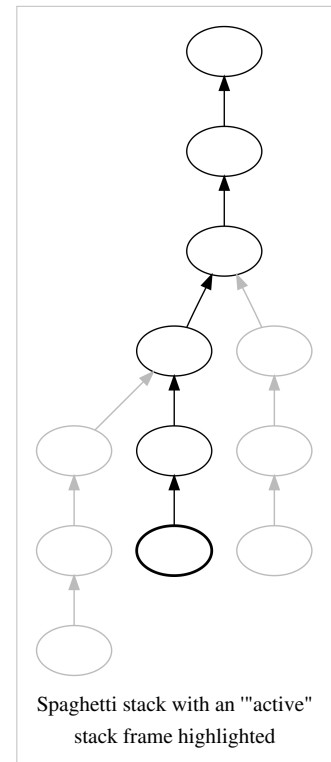
# Spaghetti stack

A **spaghetti stack** (also called a **cactus stack**, **saguaro stack** or **in-tree**) in computer science is an N-ary tree data structure in which child nodes have pointers to the parent nodes (but not vice-versa). When a list of nodes is traversed from a leaf node to the root node by chasing these parent pointers, the structure looks like a linked list stack.<sup>[1]</sup> It can be analogized to a linked list having one and only parent pointer called "next" or "link", and ignoring that each parent may have other children (which are not accessible anyway since there are no downward pointers).

Spaghetti stack structures arise in situations when records are dynamically pushed and popped onto a stack as execution progresses, but references to the popped records remain in use.

For example, a compiler for a language such as C creates a spaghetti stack as it opens and closes symbol tables representing block scopes. When a new block scope is opened, a symbol table is pushed onto a stack. When the closing curly brace is encountered, the scope is closed and the symbol table is popped. But that symbol table is remembered, rather than destroyed. And of course it remembers its higher level "parent" symbol table and so on. Thus when the compiler is later performing translations over the abstract syntax tree, for any given expression, it can fetch the symbol table representing that expression's environment and can resolve references to identifiers. If the expression refers to a variable X, it is first sought after in the leaf symbol table representing the inner-most lexical scope, then in the parent and so on.

A similar data structure appears in disjoint-set forests, a type of disjoint-set data structure.



## Use in programming language runtimes

The term *spaghetti stack* is closely associated with implementations of programming languages that support continuations. Spaghetti stacks are used to implement the actual run-time stack containing variable bindings and other environmental features. When continuations must be supported, a function's local variables cannot be destroyed when that function returns: a saved continuation may later re-enter into that function, and will expect not only the variables there to be intact, but it will also expect the entire stack to be there, so it can return again! To resolve this problem, stack frames can be dynamically allocated in a spaghetti stack structure, and simply left behind to be garbage collected when no continuations refer to them any longer. This type of structure also solves both the upward and downward funarg problems, so first-class lexical closures are readily implemented in that substrate also.

Examples of languages that use spaghetti stacks are:

- Languages having first-class continuations such as Scheme and Standard ML of New Jersey
- Languages where the execution stack can be inspected and modified at runtime such as Smalltalk
- Felix
- Rust

## References

- [1] Machinery, Sponsored (1988). *Proceedings of the 1988 Acm Conference on Lisp and Functional Programming*. New York: ACM Press. ISBN 978-0-89791-273-0.

# Disjoint-set data structure

---

In computing, a **disjoint-set data structure** is a data structure that keeps track of a set of elements partitioned into a number of disjoint (nonoverlapping) subsets. A **union-find algorithm** is an algorithm that performs two useful operations on such a data structure:

- *Find*: Determine which subset a particular element is in. This can be used for determining if two elements are in the same subset.
- *Union*: Join two subsets into a single subset.

Because it supports these two operations, a disjoint-set data structure is sometimes called a *union-find data structure* or *merge-find set*. The other important operation, *MakeSet*, which makes a set containing only a given element (a singleton), is generally trivial. With these three operations, many practical partitioning problems can be solved (see the *Applications* section).

In order to define these operations more precisely, some way of representing the sets is needed. One common approach is to select a fixed element of each set, called its *representative*, to represent the set as a whole. Then, *Find*( $x$ ) returns the representative of the set that  $x$  belongs to, and *Union* takes two set representatives as its arguments.

## Disjoint-set linked lists

A simple approach to creating a disjoint-set data structure is to create a linked list for each set. The element at the head of each list is chosen as its representative.

*MakeSet* creates a list of one element. *Union* appends the two lists, a constant-time operation. The drawback of this implementation is that *Find* requires  $\Omega(n)$  or linear time to traverse the list backwards from a given element to the head of the list.

This can be avoided by including in each linked list node a pointer to the head of the list; then *Find* takes constant time, since this pointer refers directly to the set representative. However, *Union* now has to update each element of the list being appended to make it point to the head of the new combined list, requiring  $\Omega(n)$  time.

When the length of each list is tracked, the required time can be improved by always appending the smaller list to the longer. Using this *weighted-union heuristic*, a sequence of  $m$  *MakeSet*, *Union*, and *Find* operations on  $n$  elements requires  $O(m + n \log n)$  time.<sup>[1]</sup> For asymptotically faster operations, a different data structure is needed.

## Analysis of the naive approach

We now explain the bound  $O(n \log(n))$  above.

Suppose you have a collection of lists, each node of a list contains an object, the name of the list to which it belongs, and the number of elements in that list. Also assume that the sum of the number of elements in all lists is  $n$  (i.e. there are  $n$  elements overall). We wish to be able to merge any two of these lists, and update all of their nodes so that they still contain the name of the list to which they belong. The rule for merging the lists  $A$  and  $B$  is that if  $A$  is larger than  $B$  then merge the elements of  $B$  into  $A$  and update the elements that used to belong to  $B$ , and vice versa.

Choose an arbitrary element of list  $L$ , say  $x$ . We wish to count how many times in the worst case will  $x$  need to have the name of the list to which it belongs updated. The element  $x$  will only have its name updated when the list it

belongs to is merged with another list of the same size or of greater size. Each time that happens, the size of the list to which  $x$  belongs at least doubles. So finally, the question is "how many times can a number double before it is the size of  $n$ ?" (then the list containing  $x$  will contain all  $n$  elements). The answer is exactly  $\log_2(n)$ . So for any given element of any given list in the structure described, it will need to be updated  $\log_2(n)$  times in the worst case. Therefore updating a list of  $n$  elements stored in this way takes  $O(n \log(n))$  time in the worst case. A find operation can be done in  $O(1)$  for this structure because each node contains the name of the list to which it belongs.

A similar argument holds for merging the trees in the data structures discussed below, additionally it helps explain the time analysis of some operations in the binomial heap and Fibonacci heap data structures.

## Disjoint-set forests

Disjoint-set forests are data structures where each set is represented by a tree data structure, in which each node holds a reference to its parent node (see spaghetti stack). They were first described by Bernard A. Galler and Michael J. Fischer in 1964,<sup>[2]</sup> although their precise analysis took years.

In a disjoint-set forest, the representative of each set is the root of that set's tree. *Find* follows parent nodes until it reaches the root. *Union* combines two trees into one by attaching the root of one to the root of the other. One way of implementing these might be:

```
function MakeSet(x)
    x.parent := x

function Find(x)
    if x.parent == x
        return x
    else
        return Find(x.parent)

function Union(x, y)
    xRoot := Find(x)
    yRoot := Find(y)
    xRoot.parent := yRoot
```

In this naive form, this approach is no better than the linked-list approach, because the tree it creates can be highly unbalanced; however, it can be enhanced in two ways.

The first way, called *union by rank*, is to always attach the smaller tree to the root of the larger tree, rather than vice versa. Since it is the depth of the tree that affects the running time, the tree with smaller depth gets added under the root of the deeper tree, which only increases the depth if the depths were equal. In the context of this algorithm, the term *rank* is used instead of *depth* since it stops being equal to the depth if path compression (described below) is also used. One-element trees are defined to have a rank of zero, and whenever two trees of the same rank  $r$  are united, the rank of the result is  $r+1$ . Just applying this technique alone yields a worst-case running-time of  $O(\log n)$  per *MakeSet*, *Union*, or *Find* operation. Pseudocode for the improved *MakeSet* and *Union*:

```
function MakeSet(x)
    x.parent := x
    x.rank := 0

function Union(x, y)
    xRoot := Find(x)
    yRoot := Find(y)
    if xRoot == yRoot
```

```

    return

    // x and y are not already in same set. Merge them.
    if xRoot.rank < yRoot.rank
        xRoot.parent := yRoot
    else if xRoot.rank > yRoot.rank
        yRoot.parent := xRoot
    else
        yRoot.parent := xRoot
        xRoot.rank := xRoot.rank + 1

```

The second improvement, called *path compression*, is a way of flattening the structure of the tree whenever *Find* is used on it. The idea is that each node visited on the way to a root node may as well be attached directly to the root node; they all share the same representative. To effect this, as *Find* recursively traverses up the tree, it changes each node's parent reference to point to the root that it found. The resulting tree is much flatter, speeding up future operations not only on these elements but on those referencing them, directly or indirectly. Here is the improved *Find*:

```

function Find(x)
    if x.parent != x
        x.parent := Find(x.parent)
    return x.parent

```

These two techniques complement each other; applied together, the amortized time per operation is only  $O(\alpha(n))$ , where  $\alpha(n)$  is the inverse of the function  $f(n) = A(n, n)$ , and  $A$  is the extremely fast-growing Ackermann function. Since  $\alpha(n)$  is the inverse of this function,  $\alpha(n)$  is less than 5 for all remotely practical values of  $n$ .

Thus, the amortized running time per operation is effectively a small constant.

In fact, this is asymptotically optimal: Fredman and Saks showed in 1989 that  $\Omega(\alpha(n))$  words must be accessed by any disjoint-set data structure per operation on average.<sup>[3]</sup>

## Applications

Disjoint-set data structures model the partitioning of a set, for example to keep track of the connected components of an undirected graph. This model can then be used to determine whether two vertices belong to the same component, or whether adding an edge between them would result in a cycle. The Union-Find algorithm is used in high-performance implementations of Unification.<sup>[4]</sup>

This data structure is used by the Boost Graph Library to implement its Incremental Connected Components<sup>[5]</sup> functionality. It is also used for implementing Kruskal's algorithm to find the minimum spanning tree of a graph.

Note that the implementation as disjoint-set forests doesn't allow deletion of edges—even without path compression or the rank heuristic.

## History

While the ideas used in disjoint-set forests have long been familiar, Robert Tarjan was the first to prove the upper bound (and a restricted version of the lower bound) in terms of the inverse Ackermann function, in 1975.<sup>[6]</sup> Until this time the best bound on the time per operation, proven by Hopcroft and Ullman,<sup>[7]</sup> was  $O(\log^* n)$ , the iterated logarithm of  $n$ , another slowly growing function (but not quite as slow as the inverse Ackermann function).

Tarjan and Van Leeuwen also developed one-pass *Find* algorithms that are more efficient in practice while retaining the same worst-case complexity.<sup>[8]</sup>

In 2007, Sylvain Conchon and Jean-Christophe Filliâtre developed a persistent version of the disjoint-set forest data structure, allowing previous versions of the structure to be efficiently retained, and formalized its correctness using the proof assistant Coq.<sup>[9]</sup>

## References

- [1] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7. Chapter 21: Data structures for Disjoint Sets, pp. 498–524.
- [2] Bernard A. Galler and Michael J. Fischer. An improved equivalence algorithm. *Communications of the ACM*, Volume 7, Issue 5 (May 1964), pp. 301–303. The paper originating disjoint-set forests. ACM Digital Library (<http://portal.acm.org/citation.cfm?doid=364099.364331>)
- [3] M. Fredman and M. Saks. The cell probe complexity of dynamic data structures. *Proceedings of the Twenty-First Annual ACM Symposium on Theory of Computing*, pp. 345–354. May 1989. "Theorem 5: Any CPROBE( $\log n$ ) implementation of the set union problem requires  $\Omega(m \alpha(m, n))$  time to execute  $m$  Find's and  $n-1$  Union's, beginning with  $n$  singleton sets."
- [4] Knight, Kevin (1989). "Unification: A multidisciplinary survey" (<http://portal.acm.org/citation.cfm?id=62030>). *ACM Computing Surveys* **21**: 93–124. doi:10.1145/62029.62030. .
- [5] [http://www.boost.org/libs/graph/doc/incremental\\_components.html](http://www.boost.org/libs/graph/doc/incremental_components.html)
- [6] Tarjan, Robert Endre (1975). "Efficiency of a Good But Not Linear Set Union Algorithm" (<http://portal.acm.org/citation.cfm?id=321884>). *Journal of the ACM* **22** (2): 215–225. doi:10.1145/321879.321884. .
- [7] Hopcroft, J. E.; Ullman, J. D. (1973). "Set Merging Algorithms". *SIAM Journal on Computing* **2** (4): 294–303. doi:10.1137/0202024.
- [8] Robert E. Tarjan and Jan van Leeuwen. Worst-case analysis of set union algorithms. *Journal of the ACM*, 31(2):245–281, 1984.
- [9] Sylvain Conchon and Jean-Christophe Filliâtre. A Persistent Union-Find Data Structure. In ACM SIGPLAN Workshop on ML, Freiburg, Germany, October 2007.

## External links

- C++ implementation ([http://www.boost.org/libs/disjoint\\_sets/disjoint\\_sets.html](http://www.boost.org/libs/disjoint_sets/disjoint_sets.html)), part of the Boost C++ libraries
- A Java implementation with an application to color image segmentation, Statistical Region Merging (SRM), IEEE Trans. Pattern Anal. Mach. Intell. 26(11): 1452–1458 (2004) (<http://www.lix.polytechnique.fr/~nielsen/Srmjava.java>)
- Java applet: A Graphical Union-Find Implementation (<http://www.cs.unm.edu/~rlpm/499/uf.html>), by Rory L. P. McGuire
- *Wait-free Parallel Algorithms for the Union-Find Problem* (<http://citeseer.ist.psu.edu/anderson94waitfree.html>), a 1994 paper by Richard J. Anderson and Heather Woll describing a parallelized version of Union-Find that never needs to block
- Python implementation (<http://code.activestate.com/recipes/215912-union-find-data-structure/>)

---

# Space-partitioning trees

---

## Space partitioning

---

In mathematics, **space partitioning** is the process of dividing a space (usually a Euclidean space) into two or more disjoint subsets (see also partition of a set). In other words, space partitioning divides a space into non-overlapping regions. Any point in the space can then be identified to lie in exactly one of the regions.

### Overview

Space-partitioning systems are often hierarchical, meaning that a space (or a region of space) is divided into several regions, and then the same space-partitioning system is recursively applied to each of the regions thus created. The regions can be organized into a tree, called a **space-partitioning tree**.

Most space-partitioning systems use planes (or, in higher dimensions, hyperplanes) to divide space: points on one side of the plane form one region, and points on the other side form another. Points exactly on the plane are usually arbitrarily assigned to one or the other side. Recursively partitioning space using planes in this way produces a BSP tree, one of the most common forms of space partitioning.

Space partitioning is particularly important in computer graphics, where it is frequently used to organize the objects in a virtual scene. Storing objects in a space-partitioning data structure makes it easy and fast to perform certain kinds of geometry queries — for example, determining whether two objects are close to each other in collision detection, or determining whether a ray intersects an object in ray tracing.<sup>[1]</sup>

### Use in computer graphics

Space partitioning is heavily used in ray tracing. A typical scene may contain millions of polygons. Performing a ray/polygon intersection test with each would be a very computationally expensive task. The use of a proper space partitioning data structure (kd-tree or BVH for example) can reduce the number of intersection test to just a few per primary ray, yielding a logarithmic time complexity with respect to the number of polygons.<sup>[2][3]</sup>

Space partitioning is also often used in scanline algorithms to eliminate the polygons out of the camera's viewing frustum, limiting the number of polygons processed by the pipeline.

### Other uses

In integrated circuit design, an important step is design rule check. This step ensures that the completed design is manufacturable. The check involves rules that specify widths and spacings and other geometry patterns. A modern design can have billions of polygons that represent wires and transistors. Efficient checking relies heavily on geometry query. For example, a rule may specify that any polygon must be at least  $n$  nanometers from any other polygon. This is converted into a geometry query by enlarging a polygon by  $n$  at all sides and query to find all intersecting polygons.

---



## Types of space partitioning data structures

Common space partitioning systems include:

- BSP trees
- Quadtrees
- Octrees
- *kd*-trees
- Bins
- R-trees
- Bounding volume hierarchies
- SEADSs.

## References

- [1] Ray Tracing - Auxiliary Data Structures (<http://undergraduate.csse.uwa.edu.au/units/CITS4241/Handouts/Lecture14.html>)
- [2] Tomas Nikodym (2010). "Ray Tracing Algorithm For Interactive Applications" ([https://dip.felk.cvut.cz/browse/pdfcache/nikodtom\\_2010bach.pdf](https://dip.felk.cvut.cz/browse/pdfcache/nikodtom_2010bach.pdf)). *Czech Technical University, FEE*. .
- [3] Ingo Wald, William R. Mark, et al. (2007). "State of the Art in Ray Tracing Animated Scenes" (<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.108.8495&rep=rep1&type=pdf>). *EUROGRAPHICS*. .

# Binary space partitioning

---

In computer science, **binary space partitioning (BSP)** is a method for recursively subdividing a space into convex sets by hyperplanes. This subdivision gives rise to a representation of objects within the space by means of a tree data structure known as a **BSP tree**.

Binary space partitioning was developed in the context of 3D computer graphics<sup>[1][2]</sup>, where the structure of a BSP tree allows spatial information about the objects in a scene that is useful in rendering, such as their ordering from front-to-back with respect to a viewer at a given location, to be accessed rapidly. Other applications include performing geometrical operations with shapes (constructive solid geometry) in CAD<sup>[3]</sup>, collision detection in robotics and 3-D video games, ray tracing and other computer applications that involve handling of complex spatial scenes.

## Overview

Binary space partitioning is a generic process of recursively dividing a scene into two until the partitioning satisfies one or more requirements. It can be seen as a generalisation of other spatial tree structures such as *k*-d trees and quadtrees, one in which the hyperplanes which partition the space may have any orientation, rather than being aligned with the coordinate axes as they are in *k*-d trees or quadtrees. When used in computer graphics to render scenes composed of planar polygons, the partitioning planes are frequently (but not always) chosen to coincide with the planes defined by polygons in the scene.

The specific choice of partitioning plane and criterion for terminating the partitioning process varies depending on the purpose of the BSP tree. For example, in computer graphics rendering, the scene is divided until each node of the BSP tree contains only polygons which can be rendered in arbitrary order. When back-face culling is used, each node therefore contains a convex set of polygons, whereas when rendering double-sided polygons, each node of the BSP tree contains only polygons which lie in a single plane. In collision detection or ray tracing, a scene may be divided up into primitives on which collision or ray intersection tests are straightforward.

Binary space partitioning arose from the requirement in computer graphics to rapidly draw three dimensional scenes composed of polygons. A simple way to draw such scenes is the painter's algorithm, in which polygons are drawn in

---

order of distance from the viewer, from back to front, painting over the background and previous polygons with each closer object. This approach has two disadvantages: the time required to sort polygons in order from back to front, and the possibility of errors when drawing overlapping polygons. Fuchs and co-authors<sup>[2]</sup> showed that construction of a BSP tree solved both of these problems, by providing a rapid method of sorting polygons with respect to a given viewpoint (linear in the number of polygons in the scene) and by subdividing overlapping polygons to avoid the errors that can occur when using the painter's algorithm. A disadvantage of using binary space partitioning is that the generation of a BSP tree can be a time-consuming operation. Typically, it is therefore performed once on static geometry, as a pre-calculation step, prior its use in rendering or other realtime operations on a scene. The expense of constructing a BSP tree makes it difficult and inefficient to directly implement moving objects into a tree.

BSP trees are often used by 3D video games, particularly first-person shooters and those with indoor environments. Game engines utilising BSP trees include the Doom engine (probably the earliest game to use a BSP data structure was *Doom*), the Quake engine and its descendants. In video games, BSP trees containing the static geometry of a scene are often used together with a Z-buffer, to correctly merge movable objects such as doors and characters onto the background scene. While binary space partitioning provides a convenient way to store and retrieve spatial information about polygons in a scene, it does not solve the problem of visible surface determination.

## Generation

The canonical use of a BSP tree is for rendering polygons (that are double-sided, that is, without back-face culling) with the painter's algorithm<sup>[2]</sup>. Such a tree is constructed from an unsorted list of all the polygons in a scene. The recursive algorithm for construction of a BSP tree from that list of polygons is<sup>[2]</sup>:

1. Choose a polygon  $P$  from the list.
2. Make a node  $N$  in the BSP tree, and add  $P$  to the list of polygons at that node.
3. For each other polygon in the list:
  1. If that polygon is wholly in front of the plane containing  $P$ , move that polygon to the list of nodes in front of  $P$ .
  2. If that polygon is wholly behind the plane containing  $P$ , move that polygon to the list of nodes behind  $P$ .
  3. If that polygon is intersected by the plane containing  $P$ , split it into two polygons and move them to the respective lists of polygons behind and in front of  $P$ .
  4. If that polygon lies in the plane containing  $P$ , add it to the list of polygons at node  $N$ .
4. Apply this algorithm to the list of polygons in front of  $P$ .
5. Apply this algorithm to the list of polygons behind  $P$ .

The following diagram illustrates the use of this algorithm in converting a list of lines or polygons into a BSP tree. At each of the eight steps (i.-viii.), the algorithm above is applied to a list of lines, and one new node is added to the tree.

	Start with a list of lines, (or in 3-D, polygons) making up the scene. In the tree diagrams, lists are denoted by rounded rectangles and nodes in the BSP tree by circles. In the spatial diagram of the lines, direction chosen to be the 'front' of a line is denoted by an arrow.	
i.	Following the steps of the algorithm above, <ol style="list-style-type: none"> <li>1. We choose a line, A, from the list and,...</li> <li>2. ...add it to a node.</li> <li>3. We split the remaining lines in the list into those which lie in front of A (i.e. B2, C2, D2), and those which lie behind (B1, C1, D1).</li> <li>4. We process first the lines lying in front of A (in steps ii–v),...</li> <li>5. ...followed by those behind it (in steps vi–vii).</li> </ol>	
ii.	We now apply the algorithm to the list of lines in front of A (containing B2, C2, D2). We choose a line, B2, add it to a node and split the rest of the list into those lines that are in front of B2 (D2), and those that are behind it (C2, D3).	

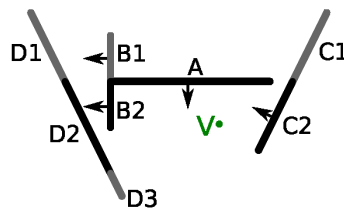
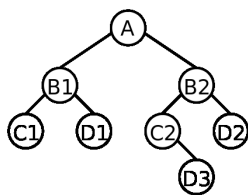
iii.	Choose a line, D2, from the list of lines in front of B2. It is the only line in the list, so after adding it to a node, nothing further needs to be done.	
iv.	We are done with the lines in front of B2, so consider the lines behind B2 (C2 and D3). Choose one of these (C2), add it to a node, and put the other line in the list (D3) into the list of lines in front of C2.	
v.	Now look at the list of lines in front of C2. There is only one line (D3), so add this to a node and continue.	
vi.	We have now added all of the lines in front of A to the BSP tree, so we now start on the list of lines behind A. Choosing a line (B1) from this list, we add B1 to a node and split the remainder of the list into lines in front of B1 (i.e. D1), and lines behind B1 (i.e. C1).	
vii.	Processing first the list of lines in front of B1, D1 is the only line in this list, so add this to a node and continue.	
viii.	Looking next at the list of lines behind B1, the only line in this list is C1, so add this to a node, and the BSP tree is complete.	

The final number of polygons or lines in a tree will often be larger (sometimes much larger<sup>[2]</sup>) than that in the original list, since lines or polygons that cross the partitioning plane must be split into two. It is desirable that this increase is minimised, but also that the final tree remains reasonably balanced. The choice of which polygon or line is used as a partitioning plane (in step 1 of the algorithm) is therefore important in creating an efficient BSP tree.

## Traversal

A BSP tree is traversed in a linear time, in an order determined by the particular function of the tree. Again using the example of rendering double-sided polygons using the painter's algorithm, for a polygon  $P$  to be drawn correctly, all the polygons which are behind the plane in which  $P$  lies must be drawn first, then polygon  $P$  must be drawn, then finally the polygons in front of  $P$  must be drawn. If this drawing order is satisfied for all polygons in a scene, then the entire scene is rendered in the correct order. This procedure can be implemented by recursively traversing a BSP tree using the following algorithm<sup>[2]</sup>. From a given viewing location  $V$ , to render a BSP tree,

1. If the current node is a leaf node, render the polygons at the current node.
2. Otherwise, if the viewing location  $V$  is in front of the current node:
  1. Render the child BSP tree containing polygons behind the current node
  2. Render the polygons at the current node
  3. Render the child BSP tree containing polygons in front of the current node
3. Otherwise, if the viewing location  $V$  is behind the current node:
  1. Render the child BSP tree containing polygons in front of the current node
  2. Render the polygons at the current node
  3. Render the child BSP tree containing polygons behind the current node
4. Otherwise, the viewing location  $V$  must be exactly on the plane associated with the current node. Then:
  1. Render the child BSP tree containing polygons in front of the current node
  2. Render the child BSP tree containing polygons behind the current node



Applying this algorithm recursively to the BSP tree generated above results in the following steps:

- The algorithm is first applied to the root node of the tree, node *A*. *V* is in front of node *A*, so we apply the algorithm first to the child BSP tree containing polygons behind *A*
- This tree has root node *B1*. *V* is behind *B1* so first we apply the algorithm to the child BSP tree containing polygons in front of *B1*:
  - This tree is just the leaf node *D1*, so the polygon *D1* is rendered.
- We then render the polygon *B1*.
- We then apply the algorithm to the child BSP tree containing polygons behind *B1*:
  - This tree is just the leaf node *C1*, so the polygon *C1* is rendered.
- We then draw the polygons of *A*
- We then apply the algorithm to the child BSP tree containing polygons in front of *A*
- This tree has root node *B2*. *V* is behind *B2* so first we apply the algorithm to the child BSP tree containing polygons in front of *B2*:
  - This tree is just the leaf node *D2*, so the polygon *D2* is rendered.
- We then render the polygon *B2*.
- We then apply the algorithm to the child BSP tree containing polygons behind *B2*:
  - This tree has root node *C2*. *V* is in front of *C2* so first we would apply the algorithm to the child BSP tree containing polygons behind *C2*. There is no such tree, however, so we continue.
  - We render the polygon *C2*.
  - We apply the algorithm to the child BSP tree containing polygons in front of *C2*
    - This tree is just the leaf node *D3*, so the polygon *D3* is rendered.

The tree is traversed in linear time and renders the polygons in a far-to-near ordering (*D1*, *B1*, *C1*, *A*, *D2*, *B2*, *C2*, *D3*) suitable for the painter's algorithm.

## Timeline

- 1969 Schumacker et al.<sup>[1]</sup> published a report that described how carefully positioned planes in a virtual environment could be used to accelerate polygon ordering. The technique made use of depth coherence, which states that a polygon on the far side of the plane cannot, in any way, obstruct a closer polygon. This was used in flight simulators made by GE as well as Evans and Sutherland. However, creation of the polygonal data organization was performed manually by scene designer.
- 1980 Fuchs et al.<sup>[2]</sup> extended Schumacker's idea to the representation of 3D objects in a virtual environment by using planes that lie coincident with polygons to recursively partition the 3D space. This provided a fully automated and algorithmic generation of a hierarchical polygonal data structure known as a Binary Space Partitioning Tree (BSP Tree). The process took place as an off-line preprocessing step that was performed once per environment/object. At run-time, the view-dependent visibility ordering was generated by traversing the tree.
- 1981 Naylor's Ph.D thesis containing a full development of both BSP trees and a graph-theoretic approach using strongly connected components for pre-computing visibility, as well as the connection between the two methods. BSP trees as a dimension independent spatial search structure was emphasized, with applications to visible surface determination. The thesis also included the first empirical data demonstrating that the size of the tree and the number of new polygons was reasonable (using a model of the Space Shuttle).
- 1983 Fuchs et al. describe a micro-code implementation of the BSP tree algorithm on an Ikonas frame buffer system. This was the first demonstration of real-time visible surface determination using BSP trees.
- 1987 Thibault and Naylor<sup>[3]</sup> described how arbitrary polyhedra may be represented using a BSP tree as opposed to the traditional b-rep (boundary representation). This provided a solid representation vs. a surface based-representation. Set operations on polyhedra were described using a tool, enabling Constructive Solid

Geometry (CSG) in real-time. This was the fore runner of BSP level design using brushes, introduced in the Quake editor and picked up in the Unreal Editor.

- 1990 Naylor, Amanatides, and Thibault provide an algorithm for merging two bsp trees to form a new bsp tree from the two original trees. This provides many benefits including: combining moving objects represented by BSP trees with a static environment (also represented by a BSP tree), very efficient CSG operations on polyhedra, exact collisions detection in  $O(\log n * \log n)$ , and proper ordering of transparent surfaces contained in two interpenetrating objects (has been used for an x-ray vision effect).
- 1990 Teller and Séquin proposed the offline generation of potentially visible sets to accelerate visible surface determination in orthogonal 2D environments.
- 1991 Gordon and Chen [CHEN91] described an efficient method of performing front-to-back rendering from a BSP tree, rather than the traditional back-to-front approach. They utilised a special data structure to record, efficiently, parts of the screen that have been drawn, and those yet to be rendered. This algorithm, together with the description of BSP Trees in the standard computer graphics textbook of the day (Computer Graphics: Principles and Practice) was used by John Carmack in the making of *Doom*.
- 1992 Teller's PhD thesis described the efficient generation of potentially visible sets as a pre-processing step to acceleration real-time visible surface determination in arbitrary 3D polygonal environments. This was used in *Quake* and contributed significantly to that game's performance.
- 1993 Naylor answers the question of what characterizes a good BSP tree. He used expected case models (rather than worst case analysis) to mathematically measure the expected cost of searching a tree and used this measure to build good BSP trees. Intuitively, the tree represents an object in a multi-resolution fashion (more exactly, as a tree of approximations). Parallels with Huffman codes and probabilistic binary search trees are drawn.

## References

- [1] Schumacker, Robert A. ; Brand, Brigitta; Gilliland, Maurice G.; Sharp, Werner H (1969). Study for Applying Computer-Generated Images to Visual Simulation (Report). U.S. Air Force Human Resources Laboratory. pp. 142. AFHRL-TR-69-14.
- [2] Fuchs, Henry; Kedem, Zvi. M; Naylor, Bruce F. (1980). "On Visible Surface Generation by A Priori Tree Structures". *SIGGRAPH '80 Proceedings of the 7th annual conference on Computer graphics and interactive techniques*. ACM, New York. p. 124-133. doi:10.1145/965105.807481.
- [3] Thibault, William C.; Naylor, Bruce F. (1987). "Set operations on polyhedra using binary space partitioning trees". *SIGGRAPH '87 Proceedings of the 14th annual conference on Computer graphics and interactive techniques*. ACM, New York. p. 153-162. doi:10.1145/37402.37421.

## Additional references

- [NAYLOR90] B. Naylor, J. Amanatides, and W. Thibault, "Merging BSP Trees Yields Polyhedral Set Operations", Computer Graphics (Siggraph '90), 24(3), 1990.
- [NAYLOR93] B. Naylor, "Constructing Good Partitioning Trees", Graphics Interface (annual Canadian CG conference) May, 1993.
- [CHEN91] S. Chen and D. Gordon. "Front-to-Back Display of BSP Trees." (<http://www.rothschild.haifa.ac.il/~gordon/ftb-bsp.pdf>) IEEE Computer Graphics & Algorithms, pp 79–85. September 1991.
- [RADHA91] H. Radha, R. Leoonardi, M. Vetterli, and B. Naylor "Binary Space Partitioning Tree Representation of Images," Journal of Visual Communications and Image Processing 1991, vol. 2(3).
- [RADHA93] H. Radha, "Efficient Image Representation using Binary Space Partitioning Trees.", Ph.D. Thesis, Columbia University, 1993.
- [RADHA96] H. Radha, M. Vetterli, and R. Leoonardi, "Image Compression Using Binary Space Partitioning Trees," IEEE Transactions on Image Processing, vol. 5, No.12, December 1996, pp. 1610–1624.
- [WINTER99] AN INVESTIGATION INTO REAL-TIME 3D POLYGON RENDERING USING BSP TREES. Andrew Steven Winter. April 1999. available online

- Mark de Berg, Marc van Kreveld, Mark Overmars, and Otfried Schwarzkopf (2000). *Computational Geometry* (2nd revised edition ed.). Springer-Verlag. ISBN 3-540-65620-0. Section 12: Binary Space Partitions: pp. 251–265. Describes a randomized Painter's Algorithm.
- Christer Ericson: *Real-Time Collision Detection (The Morgan Kaufmann Series in Interactive 3-D Technology)*. Verlag *Morgan Kaufmann*, S. 349-382, Jahr 2005, ISBN 1-55860-732-3

## External links

- BSP trees presentation (<http://www.cs.wpi.edu/~matt/courses/cs563/talks/bsp/bsp.html>)
- Another BSP trees presentation ([http://web.archive.org/web/20110719195212/http://www.cc.gatech.edu/classes/AY2004/cs4451a\\_fall/bsp.pdf](http://web.archive.org/web/20110719195212/http://www.cc.gatech.edu/classes/AY2004/cs4451a_fall/bsp.pdf))
- A Java applet which demonstrates the process of tree generation (<http://symbolcraft.com/graphics/bsp/>)
- A Master Thesis about BSP generating (<http://archive.gamedev.net/archive/reference/programming/features/bsptree/bsp.pdf>)
- BSP Trees: Theory and Implementation (<http://www.devmaster.net/articles/bsp-trees/>)
- BSP in 3D space (<http://www.euclideanspace.com/threed/solidmodel/spatialdecomposition/bsp/index.htm>)

## Segment tree

In computer science, a **segment tree** is a tree data structure for storing intervals, or segments. It allows querying which of the stored segments contain a given point. It is, in principle, a static structure; that is, its content cannot be modified once the structure is built. A similar data structure is the interval tree.

A segment tree for a set  $I$  of  $n$  intervals uses  $O(n \log n)$  storage and can be built in  $O(n \log n)$  time. Segment trees support searching for all the intervals that contain a query point in  $O(\log n + k)$ ,  $k$  being the number of retrieved intervals or segments.<sup>[1]</sup>

Applications of the segment tree are in the areas of computational geometry, and geographic information systems.

The segment tree can be generalized to higher dimension spaces as well.

## Structure description

*This section describes the structure of a segment tree in a one-dimensional space.*

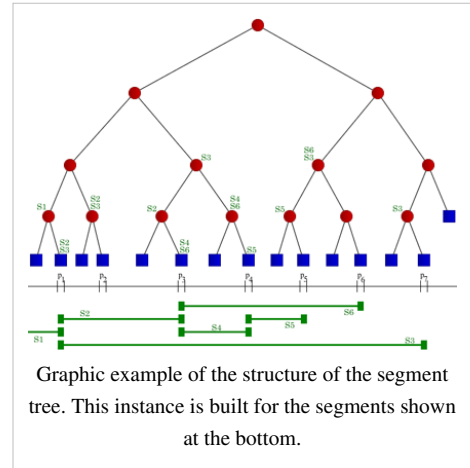
Let  $S$  be a set of intervals, or segments. Let  $p_1, p_2, \dots, p_m$  be the list of distinct interval endpoints, sorted from left to right. Consider the partitioning of the real line induced by those points. The regions of this partitioning are called *elementary intervals*. Thus, the elementary intervals are, from left to right:

$$(-\infty, p_1), [p_1, p_1], (p_1, p_2), [p_2, p_2], \dots, (p_{m-1}, p_m), [p_m, p_m], (p_m, +\infty)$$

That is, the list of elementary intervals consists of open intervals between two consecutive endpoints  $p_i$  and  $p_{i+1}$ , alternated with closed intervals consisting of a single endpoint. Single points are treated themselves as intervals because the answer to a query is not necessarily the same at the interior of an elementary interval and its endpoints.<sup>[2]</sup>

Given a set  $I$  of intervals, or segments, a segment tree  $T$  for  $I$  is structured as follows:

- $T$  is a binary tree.
- Its leafs correspond to the elementary intervals induced by the endpoints in  $I$ , in an ordered way: the leftmost leaf corresponds to the leftmost interval, and so on. The elementary interval corresponding to a leaf  $v$  is denoted  $\text{Int}(v)$ .
- The internal nodes of  $T$  corresponds to intervals that are the union of elementary intervals: the interval  $\text{Int}(N)$  corresponding to node  $N$  is the union of the intervals corresponding to the leafs of the tree rooted at  $N$ . That implies that  $\text{Int}(N)$  is the union of the intervals of its two children.
- Each node or leaf  $v$  in  $T$  stores the interval  $\text{Int}(v)$  and a set of intervals, in some data structure. This canonical subset of node  $v$  contains the intervals  $[x, x']$  from  $I$  such that  $[x, x']$  contains  $\text{Int}(v)$  and does not contain  $\text{Int}(\text{parent}(v))$ . That is, each segment in  $I$  stores the segments that span through its interval, but do not span through the interval of its parent.<sup>[3]</sup>



## Storage requirements

*This section analyzes the storage cost of a segment tree in a one-dimensional space.*

A segment tree  $T$  on a set  $I$  of  $n$  intervals uses  $O(n \log n)$  storage.

*Proof:*

*Lemma:* Any interval  $[x, x']$  of  $I$  is stored in the canonical set for at most two nodes at the same depth.

*Proof:* Let  $v_1, v_2, v_3$  be the three nodes at the same depth, numbered from left to right; and let  $w$  be the parent node of  $v_2$ . Suppose  $[x, x']$  is stored at  $v_1$  and  $v_3$ . This means that  $[x, x']$  spans the whole interval from the left endpoint of  $\text{Int}(v_1)$  to the right endpoint of  $\text{Int}(v_3)$ . Because  $v_2$  lies between  $v_1$  and  $v_3$ ,  $\text{Int}(w)$  must be contained in  $[x, x']$ . Hence,  $[x, x']$  will not be stored at  $v_2$ .

The set  $I$  has at most  $4n + 1$  elementary intervals. Because  $T$  is a binary balanced tree with at most  $4n + 1$  leaves, its height is  $O(\log n)$ . Since any interval is stored at most twice at a given depth of the tree, that the total amount of storage is  $O(n \log n)$ .<sup>[4]</sup>

## Construction

*This section describes the construction of a segment tree in a one-dimensional space.*

A segment tree from the set of segments  $I$ , can be built as follows. First, the endpoints of the intervals in  $I$  are sorted. The elementary intervals are obtained from that. Then, a balanced binary tree is built on the elementary intervals, and for each node  $v$  it is determined the interval  $\text{Int}(v)$  it represents. It remains to compute the canonical subsets for the nodes. To achieve this, the intervals in  $I$  are inserted one by one into the segment tree. An interval  $X = [x, x']$  can be inserted in a subtree rooted at  $T$ , using the following procedure<sup>[5]</sup>:

- If  $\text{Int}(T)$  is contained in  $X$  then store  $X$  at  $T$ , and finish.
- Else:
- If  $X$  intersects the canonical subset of the left child of  $T$ , then insert  $X$  in that child, recursively.
- If  $X$  intersects the canonical subset of the right child of  $T$ , then insert  $X$  in that child, recursively.

The complete construction operation takes  $O(n \log n)$  time, being  $n$  the amount of segments in  $I$ .

*Proof*

Sorting the endpoints takes  $O(n \log n)$ . Building a balanced binary tree from the sorted endpoints, takes linear time on  $n$ .

The insertion of an interval  $X = [x, x']$  into the tree, costs  $O(\log n)$ .

*Proof:* Visiting every node takes constant time (assuming that canonical subsets are stored in a simple data structure like a linked list). When we visit node  $v$ , we either store  $X$  at  $v$ , or  $\text{Int}(v)$  contains an endpoint of  $X$ . As proved above, an interval is stored at most twice at each level of the tree. There is also at most one node at every level whose corresponding interval contains  $x$ , and one node whose interval contains  $x'$ . So, at most four nodes per level are visited. Since there are  $O(\log n)$  levels, the total cost of the insertion is  $O(\log n)$ .<sup>[1]</sup>

## Query

*This section describes the query operation of a segment tree in a one-dimensional space.*

A query for a segment tree, receives a point  $q_x$ , and retrieves a list of all the segments stored which contain the point  $q_x$ .

Formally stated; given a node (subtree)  $v$  and a query point  $q_x$ , the query can be done using the following algorithm:

- Report all the intervals in  $I(v)$ .
- If  $v$  is not a leaf:
  - If  $q_x$  is in  $\text{Int}(\text{left child of } v)$  then
    - Perform a query in the left child of  $v$ .
  - Else
    - Perform a query in the right child of  $v$ .

In a segment tree that contains  $n$  intervals, those containing a given query point can be reported in  $O(\log n + k)$  time, where  $k$  is the number of reported intervals.

*Proof:* The query algorithm visits one node per level of the tree, so  $O(\log n)$  nodes in total. In the other hand, at a node  $v$ , the segments in  $I$  are reported in  $O(1 + k_v)$  time, where  $k_v$  is the number of intervals at node  $v$ , reported. The sum of all the  $k_v$  for all nodes  $v$  visited, is  $k$ , the number of reported segments.<sup>[4]</sup>

## Generalization for higher dimensions

The segment tree can be generalized to higher dimension spaces, in the form of multi-level segment trees. In higher dimension versions, the segment tree stores a collection of axis-parallel (hyper-)rectangles, and can retrieve the rectangles that contain a given query point. The structure uses  $O(n \log^{d-1} n)$  storage, and answers queries in  $O(\log^d n)$ .

The use of fractional cascading lowers the query time bound by a logarithmic factor. The use of the interval tree on the deepest level of associated structures lowers the storage bound with a logarithmic factor.<sup>[6]</sup>



## Notes

The query that asks for all the intervals containing a given point, is often referred as *stabbing query*.<sup>[7]</sup>

The segment tree is less efficient than the interval tree for range queries in one dimension, due to its higher storage requirement:  $O(n \log n)$  against the  $O(n)$  of the interval tree. The importance of the segment tree is that the segments within each node's canonical subset can be stored in any arbitrary manner.<sup>[7]</sup>

Another advantage of the segment tree is that it can easily be adapted to counting queries; that is, to report the number of segments containing a given point, instead of reporting the segments themselves. Instead of storing the intervals in the canonical subsets, it can simply store the number of them. Such a segment tree uses linear storage, and requires an  $O(\log n)$  query time, so it is optimal.<sup>[8]</sup>

A version for higher dimensions of the interval tree and the priority search tree does not exist, that is, there is no clear extension of these structures that solves the analogous problem in higher dimensions. But the structures can be used as associated structure of segment trees.<sup>[6]</sup>

## History

The segment tree was discovered by J. L. Bentley in 1977; in "Solutions to Klee's rectangle problems".<sup>[7]</sup>

## References

- [1] (de Berg, van Kreveld, Overmars, Schwarzkopf 2000, p. 227)
- [2] (de Berg, van Kreveld, Overmars, Schwarzkopf 2000, p. 224)
- [3] (de Berg, van Kreveld, Overmars, Schwarzkopf 2000, pp. 225–226)
- [4] (de Berg, van Kreveld, Overmars, Schwarzkopf 2000, p. 226)
- [5] (de Berg, van Kreveld, Overmars, Schwarzkopf 2000, pp. 226–227)
- [6] (de Berg, van Kreveld, Overmars, Schwarzkopf 2000, p. 230)
- [7] (de Berg, van Kreveld, Overmars, Schwarzkopf 2000, p. 229)
- [8] (de Berg, van Kreveld, Overmars, Schwarzkopf 2000, pp. 229–230)

---

de Berg, Mark; van Kreveld, Marc; Overmars, Mark; Schwarzkopf, Otfried (2000), *Computational Geometry: algorithms and applications* (2nd ed.), Springer-Verlag Berlin Heidelberg New York, ISBN 3-540-65620-0

<http://www.cs.nthu.edu.tw/~wkhon/ds/ds10/tutorial/tutorial6.pdf>

---

# Interval tree

---

In computer science, an **interval tree** is an ordered tree data structure to hold intervals. Specifically, it allows one to efficiently find all intervals that overlap with any given interval or point. It is often used for windowing queries, for instance, to find all roads on a computerized map inside a rectangular viewport, or to find all visible elements inside a three-dimensional scene. A similar data structure is the segment tree.

The trivial solution is to visit each interval and test whether it intersects the given point or interval, which requires  $\Theta(n)$  time, where  $n$  is the number of intervals in the collection. Since a query may return all intervals, for example if the query is a large interval intersecting all intervals in the collection, this is asymptotically optimal; however, we can do better by considering output-sensitive algorithms, where the runtime is expressed in terms of  $m$ , the number of intervals produced by the query. Interval trees are dynamic, i.e., they allow insertion and deletion of intervals. They obtain a query time of  $O(\log n)$  while the preprocessing time to construct the data structure is  $O(n \log n)$  (but the space consumption is  $O(n)$ ).

## Naive approach

In a simple case, the intervals do not overlap and they can be inserted into a simple binary search tree and queried in  $O(\log n)$  time. However, with arbitrarily overlapping intervals, there is no way to compare two intervals for insertion into the tree since orderings sorted by the beginning points or the ending points may be different. A naive approach might be to build two parallel trees, one ordered by the beginning point, and one ordered by the ending point of each interval. This allows discarding half of each tree in  $O(\log n)$  time, but the results must be merged, requiring  $O(n)$  time. This gives us queries in  $O(n + \log n) = O(n)$ , which is no better than brute-force.

Interval trees solve this problem. This article describes two alternative designs for an interval tree, dubbed the *centered interval tree* and the *augmented tree*.

## Centered interval tree

Queries require  $O(\log n + m)$  time, with  $n$  being the total number of intervals and  $m$  being the number of reported results. Construction requires  $O(n \log n)$  time, and storage requires  $O(n)$  space.

### Construction

Given a set of  $n$  intervals on the number line, we want to construct a data structure so that we can efficiently retrieve all intervals overlapping another interval or point.

We start by taking the entire range of all the intervals and dividing it in half at  $x_{center}$  (in practice,  $x_{center}$  should be picked to keep the tree relatively balanced). This gives three sets of intervals, those completely to the left of  $x_{center}$  which we'll call  $S_{left}$ , those completely to the right of  $x_{center}$  which we'll call  $S_{right}$ , and those overlapping  $x_{center}$  which we'll call  $S_{center}$ .

The intervals in  $S_{left}$  and  $S_{right}$  are recursively divided in the same manner until there are no intervals left.

The intervals in  $S_{center}$  that overlap the center point are stored in a separate data structure linked to the node in the interval tree. This data structure consists of two lists, one containing all the intervals sorted by their beginning points, and another containing all the intervals sorted by their ending points.

The result is a tertiary tree with each node storing:

- A center point
  - A pointer to another node containing all intervals completely to the left of the center point
  - A pointer to another node containing all intervals completely to the right of the center point
  - All intervals overlapping the center point sorted by their beginning point
-

- All intervals overlapping the center point sorted by their ending point

## Intersecting

Given the data structure constructed above, we receive queries consisting of ranges or points, and return all the ranges in the original set overlapping this input.

### With an Interval

First, we can reduce the case where an interval  $R$  is given as input to the simpler case where a single point is given as input. We first find all ranges with beginning or end points inside the input interval  $R$  using a separately constructed tree. In the one-dimensional case, we can use a simple tree containing all the beginning and ending points in the interval set, each with a pointer to its corresponding interval.

A binary search in  $O(\log n)$  time for the beginning and end of  $R$  reveals the minimum and maximum points to consider. Each point within this range references an interval that overlaps our range and is added to the result list. Care must be taken to avoid duplicates, since an interval might both begin and end within  $R$ . This can be done using a binary flag on each interval to mark whether or not it has been added to the result set.

The only intervals not yet considered are those overlapping  $R$  that do not have an endpoint inside  $R$ , in other words, intervals that enclose it. To find these, we pick any point inside  $R$  and use the algorithm below to find all intervals intersecting that point (again, being careful to remove duplicates).

### With a Point

The task is to find all intervals in the tree that overlap a given point  $x$ . The tree is walked with a similar recursive algorithm as would be used to traverse a traditional binary tree, but with extra affordance for the intervals overlapping the "center" point at each node.

For each tree node,  $x$  is compared to  $x\_center$ , the midpoint used in node construction above. If  $x$  is less than  $x\_center$ , the leftmost set of intervals,  $S\_left$ , is considered. If  $x$  is greater than  $x\_center$ , the rightmost set of intervals,  $S\_right$ , is considered.

As each node is processed as we traverse the tree from the root to a leaf, the ranges in its  $S\_center$  are processed. If  $x$  is less than  $x\_center$ , we know that all intervals in  $S\_center$  end after  $x$ , or they could not also overlap  $x\_center$ . Therefore, we need only find those intervals in  $S\_center$  that begin before  $x$ . We can consult the lists of  $S\_center$  that have already been constructed. Since we only care about the interval beginnings in this scenario, we can consult the list sorted by beginnings. Suppose we find the closest number no greater than  $x$  in this list. All ranges from the beginning of the list to that found point overlap  $x$  because they begin before  $x$  and end after  $x$  (as we know because they overlap  $x\_center$  which is larger than  $x$ ). Thus, we can simply start enumerating intervals in the list until the endpoint value exceeds  $x$ .

Likewise, if  $x$  is greater than  $x\_center$ , we know that all intervals in  $S\_center$  must begin before  $x$ , so we find those intervals that end after  $x$  using the list sorted by interval endings.

If  $x$  exactly matches  $x\_center$ , all intervals in  $S\_center$  can be added to the results without further processing and tree traversal can be stopped.

## Higher Dimensions

The interval tree data structure can be generalized to a higher dimension  $N$  with identical query and construction time and  $O(n \log n)$  space.

First, a range tree in  $N$  dimensions is constructed that allows efficient retrieval of all intervals with beginning and end points inside the query region  $R$ . Once the corresponding ranges are found, the only thing that is left are those ranges that enclose the region in some dimension. To find these overlaps,  $N$  interval trees are created, and one axis intersecting  $R$  is queried for each. For example, in two dimensions, the bottom of the square  $R$  (or any other horizontal line intersecting  $R$ ) would be queried against the interval tree constructed for the horizontal axis. Likewise, the left (or any other vertical line intersecting  $R$ ) would be queried against the interval tree constructed on the vertical axis.

Each interval tree also needs an addition for higher dimensions. At each node we traverse in the tree,  $x$  is compared with  $S\_center$  to find overlaps. Instead of two sorted lists of points as was used in the one-dimensional case, a range tree is constructed. This allows efficient retrieval of all points in  $S\_center$  that overlap region  $R$ .

## Deletion

If after deleting an interval from the tree, the node containing that interval contains no more intervals, that node may be deleted from the tree. This is more complex than a normal binary tree deletion operation.

An interval may overlap the center point of several nodes in the tree. Since each node stores the intervals that overlap it, with all intervals completely to the left of its center point in the left subtree, similarly for the right subtree, it follows that each interval is stored in the node closest to the root from the set of nodes whose center point it overlaps.

Normal deletion operations in a binary tree (for the case where the node being deleted has two children) involve promoting a node further from the root to the position of the node being deleted (usually the leftmost child of the right subtree, or the rightmost child of the left subtree). As a result of this promotion, some nodes that were above the promoted node will become descendants of it; it is necessary to search these nodes for intervals that also overlap the promoted node, and move those intervals into the promoted node. As a consequence, this may result in new empty nodes, which must be deleted, following the same algorithm again.

## Balancing

The same issues that affect deletion also affect rotation operations; rotation must preserve the invariant that intervals are stored as close to the root as possible.

## Augmented tree

Another way to represent intervals is described in Cormen et al. (2001, Section 14.3: Interval trees, pp. 311–317).

Both insertion and deletion require  $O(\log n)$  time, with  $n$  being the total number of intervals.

Use a simple ordered tree, for example a binary search tree or self-balancing binary search tree, where the tree is ordered by the 'low' values of the intervals, and an extra annotation is added to every node recording the maximum high value of both its subtrees. It is simple to maintain this attribute in only  $O(h)$  steps during each addition or removal of a node, where  $h$  is the height of the node added or removed in the tree, by updating all ancestors of the node from the bottom up. Additionally, the tree rotations used during insertion and deletion may require updating the high value of the affected nodes.

Now, it's known that two intervals  $A$  and  $B$  overlap only when both  $A.\text{low} \leq B.\text{high}$  and  $A.\text{high} \geq B.\text{low}$ . When searching the trees for nodes overlapping with a given interval, you can immediately skip:

- all nodes to the right of nodes whose low value is past the end of the given interval.
- all nodes that have their maximum 'high' value below the start of the given interval.

A total order can be defined on the intervals by ordering them first by their 'low' value and finally by their 'high' value. This ordering can be used to prevent duplicate intervals from being inserted into the tree in  $O(\log n)$  time, versus the  $O(k + \log n)$  time required to find duplicates if  $k$  intervals overlap a new interval.

### Java Example: Adding a new interval to the tree

The key of each node is the interval itself and the value of each node is the end point of the interval:

```
public void add(Interval i) {
    put(i, i.getEnd());
}
```

### Java Example: Searching a point or an interval in the tree

To search for an interval, you walk the tree, omitting those branches which can't contain what you're looking for. The simple case is looking for a point:

```
// Search for all intervals which contain "p", starting with the
// node "n" and adding matching intervals to the list "result"
public void search(IntervalNode n, Point p, List<Interval> result) {
    // Don't search nodes that don't exist
    if (n == null)
        return;

    // If p is to the right of the rightmost point of any interval
    // in this node and all children, there won't be any matches.
    if (p.compareTo(n.getValue()) > 0)
        return;

    // Search left children
    if (n.getLeft() != null)
        search(IntervalNode (n.getLeft()), p, result);

    // Check this node
    if (n.getKey().contains(p))
        result.add(n.getKey());

    // If p is to the left of the start of this interval,
    // then it can't be in any child to the right.
    if (p.compareTo(n.getKey().getStart()) < 0)
        return;

    // Otherwise, search right children
    if (n.getRight() != null)
        search(IntervalNode (n.getRight()), p, result);
}
```

The code to search for an interval is exactly the same except for the check in the middle:

```
// Check this node
if (n.getKey().overlapsWith(i))
```

```
result.add (n.getKey());
```

`overlapsWith()` is defined as:

```
public boolean overlapsWith(Interval other) {
    return start.compareTo(other.getEnd()) <= 0 &&
           end.compareTo(other.getStart()) >= 0;
}
```

## Higher dimension

This can be extended to higher dimensions by cycling through the dimensions at each level of the tree. For example, for two dimensions, the odd levels of the tree might contain ranges for the  $x$  coordinate, while the even levels contain ranges for the  $y$  coordinate. However, it is not quite obvious how the rotation logic will have to be extended for such cases to keep the tree balanced.

A much simpler solution is to use nested interval trees. First, create a tree using the ranges for the  $y$  coordinate. Now, for each node in the tree, add another interval tree on the  $x$  ranges, for all elements whose  $y$  range intersect that node's  $y$  range.

The advantage of this solution is that it can be extended to an arbitrary amount of dimensions using the same code base.

At first, the cost for the additional trees might seem prohibitive but that is usually not the case. As with the solution above, you need one node per  $x$  coordinate, so this cost is the same in both solutions. The only difference is that you need an additional tree structure per vertical interval. This structure is typically very small (a pointer to the root node plus maybe the number of nodes and the height of the tree).

## References

- Mark de Berg, Marc van Kreveld, Mark Overmars, and Otfried Schwarzkopf. *Computational Geometry*, Second Revised Edition. Springer-Verlag 2000. Section 10.1: Interval Trees, pp. 212–217.
- Cormen, Thomas H.; Leiserson, Charles E.; Rivest, Ronald L.; Stein, Clifford (2001), *Introduction to Algorithms* (2nd ed.), MIT Press and McGraw-Hill, ISBN 0-262-03293-7
- Franco P. Preparata and Michael Ian Shamos. *Computational Geometry: An Introduction*. Springer-Verlag, 1985

## External links

- CGAL : Computational Geometry Algorithms Library in C++<sup>[1]</sup> contains a robust implementation of Range Trees
- Interval Tree (an augmented self balancing avl tree implementation)<sup>[2]</sup>

## References

[1] <http://www.cgal.org/>

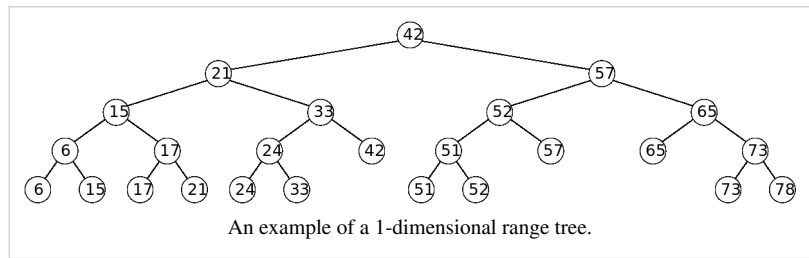
[2] <http://code.google.com/p/intervaltree/>

# Range tree

In computer science, a **range tree** is an ordered tree data structure to hold a list of points. It allows all points within a given range to be reported efficiently, and is typically used in two or higher dimensions. Range trees were introduced by Jon Louis Bentley in 1979<sup>[1]</sup>. Similar data structures were discovered independently by Lueker<sup>[2]</sup>, Lee and Wong<sup>[3]</sup>, and Willard<sup>[4]</sup>. The range tree is an alternative to the  $k$ -d tree. Compared to  $k$ -d trees, range trees offer faster query times of  $O(\log^d n + k)$  but worse storage of  $O(n \log^{(d-1)} n)$ , where  $n$  is the number of points stored in the tree,  $d$  is the dimension of each point and  $k$  is the number of points reported by a given query.

## Description

A range tree on a set of 1-dimensional points is a balanced binary search tree on those points. The points stored in the tree are stored in the leaves of the tree; each internal node stores the largest value contained in its left subtree. A range tree on a set of points



in  $d$ -dimensions is a recursively defined multi-level binary search tree. Each level of the data structure is a binary search tree on one of the  $d$ -dimensions. The first level is a binary search tree on the first of the  $d$ -coordinates. Each vertex  $v$  of this tree contains an associated structure that is a  $(d-1)$ -dimensional range tree on the last  $(d-1)$ -coordinates of the points stored in the subtree of  $v$ .

## Operations

### Construction

A 1-dimensional range tree on a set of  $n$  points is a binary search tree, which can be constructed in  $O(n \log n)$  time. Range trees in higher dimensional are constructed recursively by constructing a balanced binary search tree on the first coordinate of the points, and then, for each vertex  $v$  in this tree, constructing a  $(d-1)$ -dimensional range tree on the points contained in the subtree of  $v$ . Constructing a range tree this way would require  $O(n \log^d n)$  time.

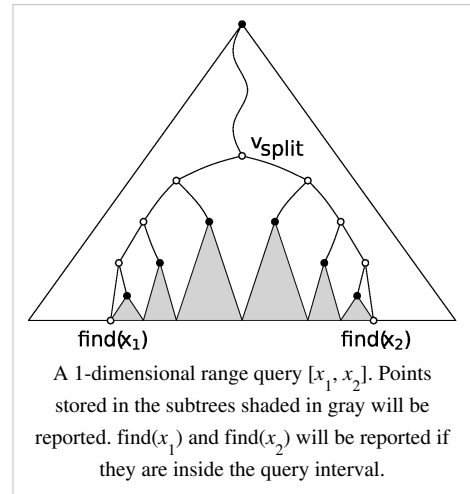
This can be improved by noticing that a range tree on a set of 2-dimensional points can be constructed in  $O(n \log n)$  time<sup>[5]</sup>. Let  $S$  be a set of  $n$  2-dimensional points. If  $S$  contains only one point, return a leaf containing that point. Otherwise, construct the associated structure of  $S$ , a 1-dimensional range tree on the  $y$ -coordinates of the points in  $S$ . Let  $x_m$  be the median  $x$ -coordinate of the points. Let  $S_L$  be the set of points with  $x$ -coordinate less than or equal to  $x_m$  and let  $S_R$  be the set of points with  $x$ -coordinate greater than  $x_m$ . Recursively construct  $v_L$ , a 2-dimensional range tree on  $S_L$ , and  $v_R$ , a 2-dimensional range tree on  $S_R$ . Create a vertex  $v$  with left-child  $v_L$  and right-child  $v_R$ . If we sort the points by their  $y$ -coordinates at the start of the algorithm, and maintain this ordering when splitting the points by their  $x$ -coordinate, we can construct the associated structures of each subtree in linear time. This reduces the time to construct a 2-dimensional range tree to  $O(n \log n)$ , which also reduces the time to construct a  $d$ -dimensional range tree to  $O(n \log^{(d-1)} n)$ .

## Range queries

Range trees can be used to find the set of points that lie inside a given interval. To report the points that lie in the interval  $[x_1, x_2]$ , we start by searching for  $x_1$  and  $x_2$ . At some vertex in the tree, the search paths to  $x_1$  and  $x_2$  will diverge. Let  $v_{\text{split}}$  be the last vertex that these two search paths have in common. Continue searching for  $x_1$  in the range tree. For every vertex  $v$  in the search path from  $v_{\text{split}}$  to  $x_1$ , if the value stored at  $v$  is greater than  $x_1$ , report every point in the right-subtree of  $v$ . If  $v$  is a leaf, report the value stored at  $v$  if it is inside the query interval. Similarly, reporting all of the points stored in the left-subtrees of the vertices with values less than  $x_2$  along the search path from  $v_{\text{split}}$  to  $x_2$ , and report the leaf of this path if it lies within the query interval.

Since the range tree is a balanced binary tree, the search paths to  $x_1$  and  $x_2$  have length  $O(\log n)$ . Reporting all of the points stored in the subtree of a vertex can be done in linear time using any tree traversal algorithm. It follows that the time to perform a range query is  $O(\log n + k)$ , where  $k$  is the number of points in the query interval.

Range queries in  $d$ -dimensions are similar. Instead of reporting all of the points stored in the subtrees of the search paths, perform a  $(d-1)$ -dimensional range query on the associated structure of each subtree. Eventually, a 1-dimensional range query will be performed and the correct points will be reported. Since a  $d$ -dimensional query consists of  $O(\log n)$   $(d-1)$ -dimensional range queries, it follows that the time required to perform a  $d$ -dimensional range query is  $O(\log^d n + k)$ , where  $k$  is the number of points in the query interval. This can be reduced to  $O(\log^{(d-1)} n + k)$  using the technique of fractional cascading<sup>[2][4][5]</sup>.



## References

- [1] Bentley, J. L. (1979). "Decomposable searching problems". *Information Processing Letters* **8** (5): 244–201. doi:10.1016/0020-0190(79)90117-0.
- [2] Lueker, G. S. (1978). "A data structure for orthogonal range queries". *19th Annual Symposium on Foundations of Computer Science (sfcs 1978)*. pp. 28–21. doi:10.1109/SFCS.1978.1.
- [3] Lee, D. T.; Wong, C. K. (1980). "Quintary trees: A file structure for multidimensional database systems". *ACM Transactions on Database Systems* **5** (3): 339. doi:10.1145/320613.320618.
- [4] Willard, Dan E.. *The super-b-tree algorithm* (Technical report). Cambridge, MA: Aiken Computer Lab, Harvard University. TR-03-79.
- [5] de Berg, Mark; Cheong, Otfried; van Kreveld, Marc; Overmars, Mark (2008). *Computational Geometry*. doi:10.1007/978-3-540-77974-2. ISBN 978-3-540-77973-5.

## External links

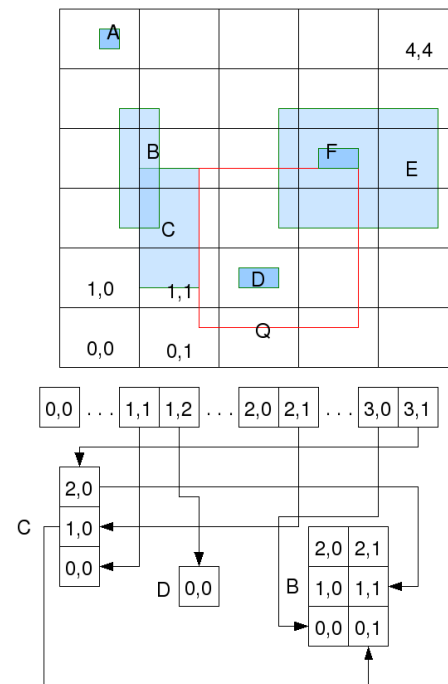
- Range and Segment Trees ([http://www.cgal.org/Manual/latest/doc\\_html/cgal\\_manual/SearchStructures/Chapter\\_main.html](http://www.cgal.org/Manual/latest/doc_html/cgal_manual/SearchStructures/Chapter_main.html)) in CGAL, the Computational Geometry Algorithms Library.
- Lecture 8: Range Trees (<http://www.cs.uu.nl/docs/vakken/ga/slides5b.pdf>), Marc van Kreveld.



# Bin

In computational geometry, the **bin** data structure allows efficient region queries, i.e., if there are some axis-aligned rectangles on a 2D plane, answer the question *Given a query rectangle, return all rectangles intersecting it*. kd-tree is another data structure that can answer this question efficiently. In the example in the figure, *A*, *B*, *C*, *D*, *E*, and *F* are existing rectangles, the query with the rectangle *Q* should return *C*, *D*, *E* and *F*, if we define all rectangles as closed intervals.

The data structure partitions a region of the 2D plane into uniform-sized *bins*. The bounding box of the bins encloses all *candidate* rectangles to be queried. All the bins are arranged in a 2D array. All the candidates are represented also as 2D arrays. The size of a candidate's array is the number of bins it intersects. For example, in the figure, candidate *B* has 6 elements arranged in a 3 row by 2 column array because it intersects 6 bins in such an arrangement. Each bin contains the head of a singly linked list. If a candidate intersects a bin, it is chained to the bin's linked list. Each element in a candidate's array is a link node in the corresponding bin's linked list.



The bin data structure

## Operations

### Query

From the query rectangle *Q*, we can find out which bin its lower-left corner intersects efficiently by simply subtracting the bin's bounding box's lower-left corner from the lower-left corner of *Q* and dividing the result by the width and height of a bin respectively. We then iterate the bins *Q* intersects and examine all the candidates in the linked-lists of these bins. For each candidate we check if it does indeed intersect *Q*. If so and it is not previously reported, then we report it. We can use the convention that we only report a candidate the first time we find it. This can be done easily by clipping the candidate against the query rectangle and comparing its lower-left corner against the current location. If it is a match then we report, otherwise we skip.

## Insertion and deletion

Insertion is linear to the number of bins a candidate intersects because inserting a candidate into 1 bin is constant time. Deletion is more expensive because we need to search the singly linked list of each bin the candidate intersects.

In a multithread environment, insert, delete and query are mutually exclusive. However, instead of locking the whole data structure, a sub-range of bins may be locked. Detailed performance analysis should be done to justify the overhead.

## Efficiency and tuning

The analysis is similar to a hash table. The worst-case scenario is that all candidates are concentrated in one bin. Then query is  $O(n)$ , delete is  $O(n)$ , and insert is  $O(1)$ , where  $n$  is the number of candidates. If the candidates are evenly spaced so that each bin has a constant number of candidates, The query is  $O(k)$  where  $k$  is the number of bins the query rectangle intersects. Insert and delete are  $O(m)$  where  $m$  is the number of bins the inserting candidate intersects. In practice delete is much slower than insert.

Like a hash table, bin's efficiency depends a lot on the distribution of both location and size of candidates and queries. In general, the smaller the query rectangle, the more efficient the query. The bin's size should be such that it contains as few candidates as possible but large enough so that candidates do not span too many bins. If a candidate span many bins, a query has to skip this candidate over and over again after it is reported at the first bin of intersection. For example, in the figure,  $E$  is visited 4 times in the query of  $Q$  and so has to be skipped 3 times.

To further speed up the query, divisions can be replaced by right shifts. This requires the number of bins along an axis direction to be an exponent of 2.

## Compared to other range query data structures

Against kd-tree, the bin structure allows efficient insertion and deletion without the complexity of rebalancing. This can be very useful in algorithms that need to incrementally add shapes to the search data structure.

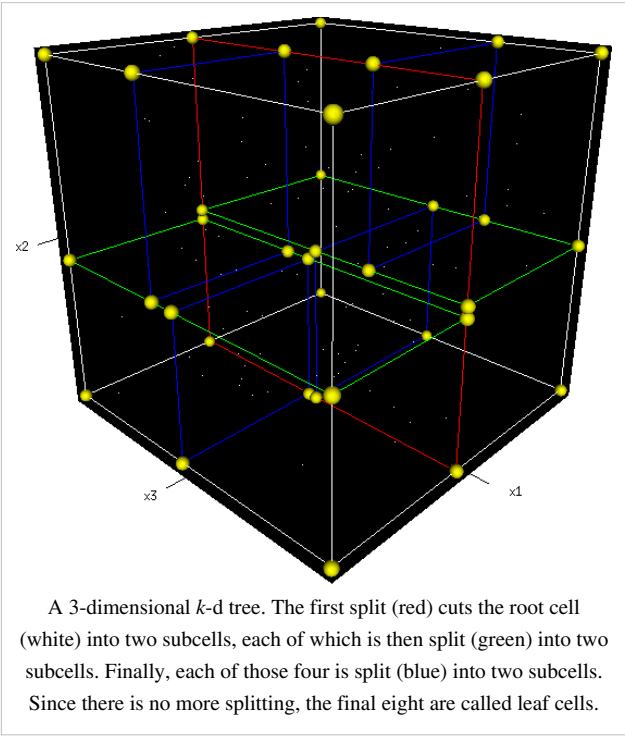
# *k*-d tree

KD-tree		
Type	Multidimensional BST	
Invented	1975	
Invented by	Jon Louis Bentley	
Time complexity in big O notation		
	Average	Worst case
Space	$O(n)$	$O(n)$
Search	$O(\log n)$	$O(n)$
Insert	$O(\log n)$	$O(n)$
Delete	$O(\log n)$	$O(n)$

In computer science, a ***k*-d tree** (short for *k-dimensional tree*) is a space-partitioning data structure for organizing points in a *k*-dimensional space. *k*-d trees are a useful data structure for several applications, such as searches involving a multidimensional search key (e.g. range searches and nearest neighbor searches). *k*-d trees are a special case of binary space partitioning trees.

## Informal description

The *k*-d tree is a binary tree in which every node is a *k*-dimensional point. Every non-leaf node can be thought of as implicitly generating a splitting hyperplane that divides the space into two parts, known as half-spaces. Points to the left of this hyperplane represent the left subtree of that node and points right of the hyperplane are represented by the right subtree. The hyperplane direction is chosen in the following way: every node in the tree is associated with one of the *k*-dimensions, with the hyperplane perpendicular to that dimension's axis. So, for example, if for a particular split the "x" axis is chosen, all points in the subtree with a smaller "x" value than the node will appear in the left subtree and all points with larger "x" value will be in the right subtree. In such a case, the hyperplane would be set by the x-value of the point, and its normal would be the unit x-axis.<sup>[1]</sup>



## Operations on *k*-d trees

### Construction

Since there are many possible ways to choose axis-aligned splitting planes, there are many different ways to construct *k*-d trees. The canonical method of *k*-d tree construction has the following constraints:

- As one moves down the tree, one cycles through the axes used to select the splitting planes. (For example, in a 3-dimensional tree, the root would have an *x*-aligned plane, the root's children would both have *y*-aligned planes, the root's grandchildren would all have *z*-aligned planes, the root's great-grandchildren would all have *x*-aligned planes, the root's great-great-grandchildren would all have *y*-aligned planes, and so on.)
- Points are inserted by selecting the median of the points being put into the subtree, with respect to their coordinates in the axis being used to create the splitting plane. (Note the assumption that we feed the entire set of *n* points into the algorithm up-front.)

This method leads to a balanced *k*-d tree, in which each leaf node is about the same distance from the root. However, balanced trees are not necessarily optimal for all applications.

Note also that it is not *required* to select the median point. In that case, the result is simply that there is no guarantee that the tree will be balanced. A simple heuristic to avoid coding a complex linear-time median-finding algorithm, or using an  $O(n \log n)$  sort of all *n* points, is to use sort to find the median of a *fixed* number of *randomly* selected points to serve as the splitting plane. In practice, this technique often results in nicely balanced trees.

Given a list of *n* points, the following algorithm uses a median-finding sort to construct a balanced *k*-d tree containing those points.

```
function kdtree (list of points pointList, int depth)
{
    // Select axis based on depth so that axis cycles through all valid values
    var int axis := depth mod k;

    // Sort point list and choose median as pivot element
    select median by axis from pointList;

    // Create node and construct subtrees
    var tree_node node;
    node.location := median;
    node.leftChild := kdtree(points in pointList before median, depth+1);
    node.rightChild := kdtree(points in pointList after median, depth+1);
    return node;
}
```

It is common that points "after" the median include only the ones that are strictly greater than the median. For points that lie on the median, it is possible to define a "superkey" function that compares the points in all dimensions. In some cases, it is acceptable to let points equal to the median lie on one side of the median, for example, by splitting the points into a "less than" subset and a "greater than or equal to" subset.

The above algorithm implemented in the Python programming language is as follows:

```
class Node: pass

def kdtree(point_list, depth=0):

    if not point_list:
```

```

    return None

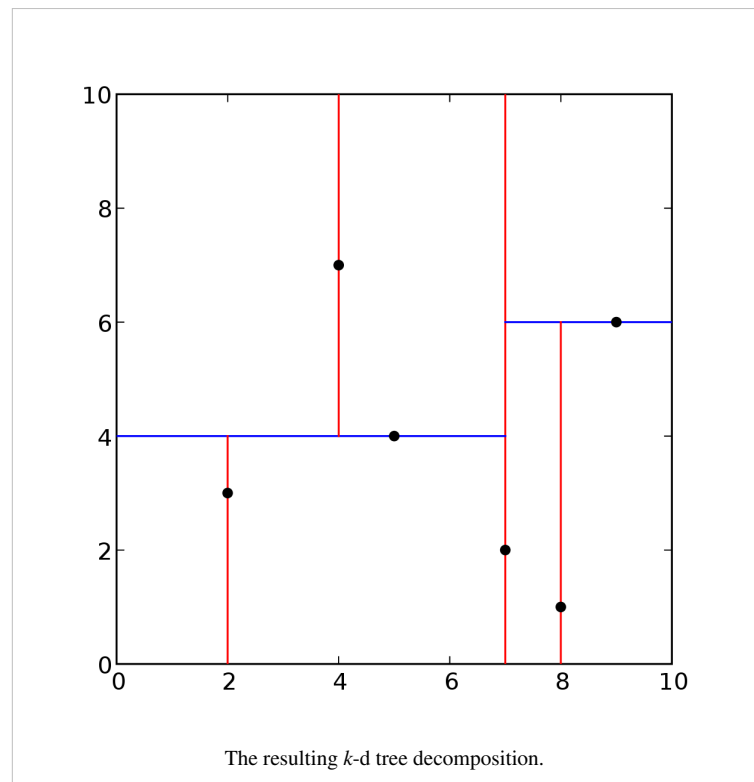
    # Select axis based on depth so that axis cycles through all valid
values
    k = len(point_list[0]) # assumes all points have the same dimension
    axis = depth % k

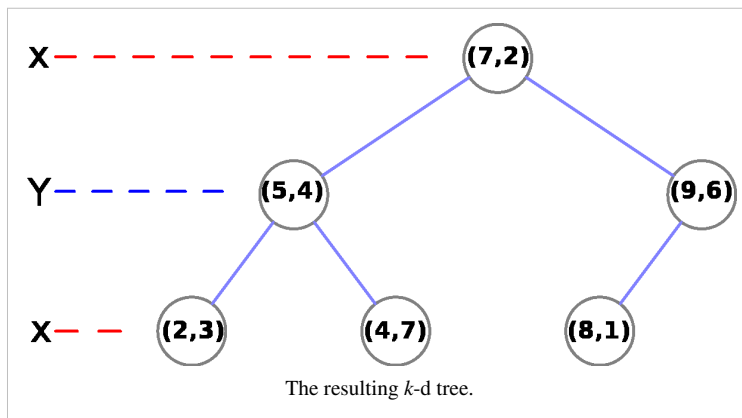
    # Sort point list and choose median as pivot element
    point_list.sort(key=lambda point: point[axis])
    median = len(point_list) // 2 # choose median

    # Create node and construct subtrees
    node = Node()
    node.location = point_list[median]
    node.left_child = kdtree(point_list[:median], depth + 1)
    node.right_child = kdtree(point_list[median + 1:], depth + 1)
    return node

```

Example usage would be:





```
point_list = [(2,3), (5,4), (9,6), (4,7), (8,1), (7,2)]
tree = kdtree(point_list)
```

The tree generated is shown on the right.

This algorithm creates the invariant that for any node, all the nodes in the left subtree are on one side of a splitting plane, and all the nodes in the right subtree are on the other side. Points that lie on the splitting plane may appear on either side. The splitting plane of a node goes through the point associated with that node (referred to in the code as *node.location*).

A novel tree-building algorithm builds a balanced  $k$ -d tree in  $O(kn \log n)$  time by sorting  $n$  points in  $k$  dimensions independently and *prior* to building the  $k$ -d tree.<sup>[2][3]</sup> A suitable sorting algorithm is Heapsort that creates a sorted array in  $O(n \log n)$  time. Application of Heapsort to  $n$  points in each of  $k$  dimensions requires  $O(kn \log n)$  time, and produces  $k$  sorted arrays of length  $n$  that contain references (or pointers) to the  $n$  points. These arrays are numbered from 0 to  $k-1$ . Each array represents the result of sorting the points in one of the  $k$  dimensions. For example, the elements of array 0, from first to last, reference the  $n$  points in order of increasing  $x$ -coordinate. Similarly, the elements of arrays 1, 2, and 3, from first to last, reference the  $n$  points in order of increasing  $y$ -,  $z$ - and  $w$ -coordinates, respectively.

In order to insert the first node into the  $k$ -d tree, the median element of array 0 is chosen and stored in the tree node. This median element splits array 0 into two subarrays. One subarray lies above the median element, and the other subarray lies below it. Also, the  $x$ -coordinate of the point that this median element references defines an  $x$ -aligned splitting plane that may be used to split each of the other  $k-1$  arrays into two subarrays. The following procedure splits an array into two subarrays:

- Consider each element of the array in order from first to last.
- Test against the splitting plane the  $x$ -coordinate of the point that is referenced by the array element, and assign that element to one of two subarrays, depending on which side of the splitting plane the point lies.
- Ignore the array element that references the same point that the median element of array 0 references, because this point defines the splitting plane.

This procedure splits the arrays into two sets of subarrays while preserving the original sorted order within each subarray. These subarrays may then be used to insert nodes into the two subtrees at the next level of the tree in a recursive manner. However, if the subarrays comprise only one or two array elements, no further recursion is required because these cases may be solved trivially.

Attention to few details simplifies creation of the  $k$ -d tree:

- The arrays should be split into subarrays that represent "less than" and "greater than or equal to" partitioning. This convention requires that, after choosing the median element of array 0, the element of array 0 that lies immediately below the median element be examined to ensure that this adjacent element references a point whose

$x$ -coordinate is less than and not equal to the  $x$ -coordinate of the splitting plane. If this adjacent element references a point whose  $x$ -coordinate is equal to the  $x$ -coordinate of the splitting plane, continue searching towards the beginning of array 0 until the first instance of an array element is found that references a point whose  $x$ -coordinate is less than and not equal to the  $x$ -coordinate of the splitting plane. When this array element is found, the element that lies immediately above this element is the correct choice for the median element. Apply this method of choosing the median element at each level of recursion.

- The procedure for producing subarrays guarantees that the two subarrays comprise one less array element than the array from which these subarrays were produced. This characteristic permits re-use of the  $k$  arrays at each level of recursion as follows: (1) copy array 0 into a temporary array, (2) build the subarrays that are produced from array 1 in array 0, (3) build the subarrays that are produced from array 2 in array 1, (4) continue this pattern, and build the subarrays that are produced from array  $k-1$  in array  $k-2$ , and finally (4) copy the temporary array into array  $k-1$ . This method permutes the subarrays so that at successive levels of the  $k$ -d tree, the median element is chosen from  $x$ -,  $y$ -,  $z$ -,  $w$ -, ... sorted arrays.
- The addresses of the first and last elements of the  $2k$  subarrays can be passed to the next level of recursion in order to designate where these subarrays lie within the  $k$  arrays. Each of the two sets of  $k$  subarrays have identical addresses for their first and last elements.

This novel tree-building algorithm requires at most  $O([k-1]n)$  tests of coordinates against splitting planes to build each of the  $\log n$  levels of a balanced  $k$ -d tree. Hence, building the entire  $k$ -d tree requires less than  $O([k-1]n \log n)$  time, which is less than the  $O(kn \log n)$  time that is required to sort the  $n$  points in  $k$  dimensions prior to building the  $k$ -d tree.

## Adding elements

One adds a new point to a  $k$ -d tree in the same way as one adds an element to any other search tree. First, traverse the tree, starting from the root and moving to either the left or the right child depending on whether the point to be inserted is on the "left" or "right" side of the splitting plane. Once you get to the node under which the child should be located, add the new point as either the left or right child of the leaf node, again depending on which side of the node's splitting plane contains the new node.

Adding points in this manner can cause the tree to become unbalanced, leading to decreased tree performance. The rate of tree performance degradation is dependent upon the spatial distribution of tree points being added, and the number of points added in relation to the tree size. If a tree becomes too unbalanced, it may need to be re-balanced to restore the performance of queries that rely on the tree balancing, such as nearest neighbour searching.

## Removing elements

To remove a point from an existing  $k$ -d tree, without breaking the invariant, the easiest way is to form the set of all nodes and leaves from the children of the target node, and recreate that part of the tree.

Another approach is to find a replacement for the point removed.<sup>[4]</sup> First, find the node  $R$  that contains the point to be removed. For the base case where  $R$  is a leaf node, no replacement is required. For the general case, find a replacement point, say  $p$ , from the subtree rooted at  $R$ . Replace the point stored at  $R$  with  $p$ . Then, recursively remove  $p$ .

For finding a replacement point, if  $R$  discriminates on  $x$  (say) and  $R$  has a right child, find the point with the minimum  $x$  value from the subtree rooted at the right child. Otherwise, find the point with the maximum  $x$  value from the subtree rooted at the left child.

## Balancing

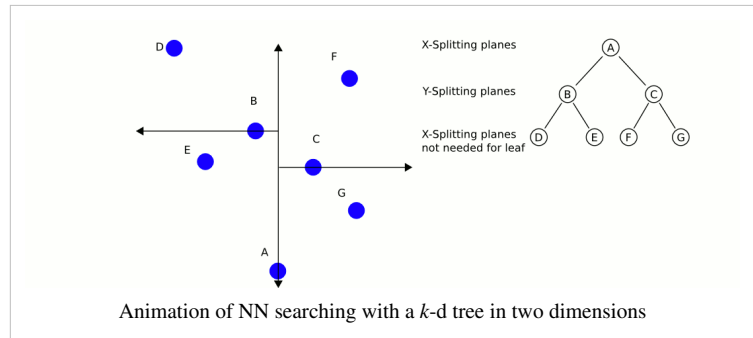
Balancing a  $k$ -d tree requires care. Because  $k$ -d trees are sorted in multiple dimensions, the tree rotation technique cannot be used to balance them — this may break the invariant.

Several variants of balanced  $k$ -d trees exist. They include divided  $k$ -d tree, pseudo  $k$ -d tree,  $k$ -d B-tree, hB-tree and Bkd-tree. Many of these variants are adaptive  $k$ -d trees.

## Nearest neighbour search

The nearest neighbour search (NN) algorithm aims to find the point in the tree that is nearest to a given input point. This search can be done efficiently by using the tree properties to quickly eliminate large portions of the search space.

Searching for a nearest neighbour in a  $k$ -d tree proceeds as follows:



1. Starting with the root node, the algorithm moves down the tree recursively, in the same way that it would if the search point were being inserted (i.e. it goes left or right depending on whether the point is less than or greater than the current node in the split dimension).
2. Once the algorithm reaches a leaf node, it saves that node point as the "current best"
3. The algorithm unwinds the recursion of the tree, performing the following steps at each node:
  1. If the current node is closer than the current best, then it becomes the current best.
  2. The algorithm checks whether there could be any points on the other side of the splitting plane that are closer to the search point than the current best. In concept, this is done by intersecting the splitting hyperplane with a hypersphere around the search point that has a radius equal to the current nearest distance. Since the hyperplanes are all axis-aligned this is implemented as a simple comparison to see whether the difference between the splitting coordinate of the search point and current node is less than the distance (overall coordinates) from the search point to the current best.
    1. If the hypersphere crosses the plane, there could be nearer points on the other side of the plane, so the algorithm must move down the other branch of the tree from the current node looking for closer points, following the same recursive process as the entire search.
    2. If the hypersphere doesn't intersect the splitting plane, then the algorithm continues walking up the tree, and the entire branch on the other side of that node is eliminated.
4. When the algorithm finishes this process for the root node, then the search is complete.

Generally the algorithm uses squared distances for comparison to avoid computing square roots. Additionally, it can save computation by holding the squared current best distance in a variable for comparison.

Finding the nearest point is an  $O(\log N)$  operation in the case of randomly distributed points. Analyses of binary search trees has found that the worst case search time for a  $k$ -dimensional KD tree containing  $N$  nodes is given by the following equation.<sup>[5]</sup>

$$t_{worst} = O(k \cdot N^{1-\frac{1}{k}})$$

In very high dimensional spaces, the curse of dimensionality causes the algorithm to need to visit many more branches than in lower dimensional spaces. In particular, when the number of points is only slightly higher than the number of dimensions, the algorithm is only slightly better than a linear search of all of the points.

The algorithm can be extended in several ways by simple modifications. It can provide the  $k$ -Nearest Neighbours to a point by maintaining  $k$  current bests instead of just one. Branches are only eliminated when they can't have points



closer than any of the  $k$  current bests.

It can also be converted to an approximation algorithm to run faster. For example, approximate nearest neighbour searching can be achieved by simply setting an upper bound on the number points to examine in the tree, or by interrupting the search process based upon a real time clock (which may be more appropriate in hardware implementations). Nearest neighbour for points that are in the tree already can be achieved by not updating the refinement for nodes that give zero distance as the result, this has the downside of discarding points that are not unique, but are co-located with the original search point.

Approximate nearest neighbour is useful in real-time applications such as robotics due to the significant speed increase gained by not searching for the best point exhaustively. One of its implementations is best-bin-first search.

## High-dimensional data

$k$ -d trees are not suitable for efficiently finding the nearest neighbour in high dimensional spaces. As a general rule, if the dimensionality is  $k$ , the number of points in the data,  $N$ , should be  $N \gg 2^k$ . Otherwise, when  $k$ -d trees are used with high-dimensional data, most of the points in the tree will be evaluated and the efficiency is no better than exhaustive search,<sup>[6]</sup> and approximate nearest-neighbour methods should be used instead.

## Complexity

- Building a static  $k$ -d tree from  $n$  points takes:
  - $O(n \log^2 n)$  time if an  $O(n \log n)$  sort such as Heapsort is used to compute the median at each level;
  - $O(n \log n)$  time if a complex linear-time median-finding algorithm such as the one described in Cormen *et al.*<sup>[7]</sup> is used;
  - $O(kn \log n)$  plus  $O([k-1]n \log n)$  time if  $n$  points are sorted in each of  $k$  dimensions using an  $O(n \log n)$  sort prior to building the  $k$ -d tree.
- Inserting a new point into a balanced  $k$ -d tree takes  $O(\log n)$  time.
- Removing a point from a balanced  $k$ -d tree takes  $O(\log n)$  time.
- Querying an axis-parallel range in a balanced  $k$ -d tree takes  $O(n^{1-1/k} + m)$  time, where  $m$  is the number of the reported points, and  $k$  the dimension of the  $k$ -d tree.

## Variations

### Volumetric objects

Instead of points, a  $k$ -d tree can also contain rectangles or hyperrectangles.<sup>[8][9]</sup> Thus range search becomes the problem of returning all rectangles intersecting the search rectangle. The tree is constructed the usual way with all the rectangles at the leaves. In an orthogonal range search, the *opposite* coordinate is used when comparing against the median. For example, if the current level is split along  $x_{\text{high}}$ , we check the  $x_{\text{low}}$  coordinate of the search rectangle. If the median is less than the  $x_{\text{low}}$  coordinate of the search rectangle, then no rectangle in the left branch can ever intersect with the search rectangle and so can be pruned. Otherwise both branches should be traversed. See also interval tree, which is a 1-dimensional special case.

### Points only in leaves

It is also possible to define a  $k$ -d tree with points stored solely in leaves.<sup>[10]</sup> This form of  $k$ -d tree allows a variety of split mechanics other than the standard median split. The midpoint splitting rule<sup>[11]</sup> selects on the middle of the longest axis of the space being searched, regardless of the distribution of points. This guarantees that the aspect ratio will be at most 2:1, but the depth is dependent on the distribution of points. A variation, called sliding-midpoint, only splits on the middle if there are points on both sides of the split. Otherwise, it splits on point nearest to the middle.

Maneewongvatana and Mount show that this offers "good enough" performance on common data sets. Using sliding-midpoint, an approximate nearest neighbour query can be answered in  $O\left(\frac{1}{\epsilon^d} \log n\right)$ . Approximate range counting can be answered in  $O\left(\log n + \left(\frac{1}{\epsilon}\right)^d\right)$  with this method.

## References

- [1] J. L. Bentley. Multidimensional binary search trees used for associative searching (<http://doi.acm.org/10.1145/361002.361007>). Communications of the ACM, 18(9):509-517, 1975.
- [2] I. Wald and V. Havran. On building fast kd-trees for ray tracing, and on doing that in  $O(N \log N)$  (<http://dcgi.felk.cvut.cz/cgg/~havran/ARTICLES/ingo06rtKdtree.pdf>) On building fast kd-trees for ray tracing, and on doing that in  $O(N \log N)$  ([https://docs.google.com/viewer?a=v&q=cache:sJ7XuH1grawJ:citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.140.2707&rep=rep1&type=pdf+&hl=en&gl=us&pid=bl&srcid=ADGEESg3NWyi\\_-vPPrvNBR3Z12qs0mswaOSFMUzBtHD37xiN23FJvNyzZwQ56BV2xJmnGks0uhf\\_scVVszu\\_eP\\_y-OE4NtXUS2rIfgB3bmzlvawtkASig=AHIEtbSjJndZKfV7dZyBNXINZZgSBkd9A](https://docs.google.com/viewer?a=v&q=cache:sJ7XuH1grawJ:citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.140.2707&rep=rep1&type=pdf+&hl=en&gl=us&pid=bl&srcid=ADGEESg3NWyi_-vPPrvNBR3Z12qs0mswaOSFMUzBtHD37xiN23FJvNyzZwQ56BV2xJmnGks0uhf_scVVszu_eP_y-OE4NtXUS2rIfgB3bmzlvawtkASig=AHIEtbSjJndZKfV7dZyBNXINZZgSBkd9A)). IEEE Symposium on Interactive Ray Tracing, pp. 61-69, 2006.
- [3] H. M. Kakde. Range searching using kd tree. pp. 1-12, 2005 (<http://www.cs.fsu.edu/~lifeifei/cis5930/kdtree.pdf>) Range searching using kd tree. pp. 1-12, 2005 (<http://www.scribd.com/doc/26091526/Range-Searching-Using-Kd-Tree>)
- [4] Chandran, Sharat. Introduction to kd-trees (<http://www.cs.umd.edu/class/spring2002/cmsc420-0401/pbasic.pdf>). University of Maryland Department of Computer Science.
- [5] Lee, D. T.; Wong, C. K. (1977). "Worst-case analysis for region and partial region searches in multidimensional binary search trees and balanced quad trees". *Acta Informatica* **9** (1): 23–29. doi:10.1007/BF00263763.
- [6] Jacob E. Goodman, Joseph O'Rourke and Piotr Indyk (Ed.) (2004). "Chapter 39 : Nearest neighbours in high-dimensional spaces". *Handbook of Discrete and Computational Geometry* (2nd ed.). CRC Press.
- [7] Cormen, Thomas H.; Leiserson, Charles E., Rivest, Ronald L.. *Introduction to Algorithms*. MIT Press and McGraw-Hill. Chapter 10.
- [8] Rosenberg J. Geographical Data Structures Compared: A Study of Data Structures Supporting Region Queries. IEEE Transaction on CAD Integrated Circuits Systems 4(1):53-67
- [9] Houthuys P. Box Sort, a multidimensional binary sorting method for rectangular boxes, used for quick range searching. The Visual Computer, 1987, 3:236-249
- [10] de Berg, Mark et al. Computational Geometry: Algorithms and Applications, 3rd Edition, pages 99-101. Springer, 2008.
- [11] S. Maneewongvatana and D. M. Mount. It's okay to be skinny, if your friends are fat (<http://www.cs.umd.edu/~mount/Papers/cgc99-smpack.pdf>). 4th Annual CGC Workshop on Computational Geometry, 1999.

## External links

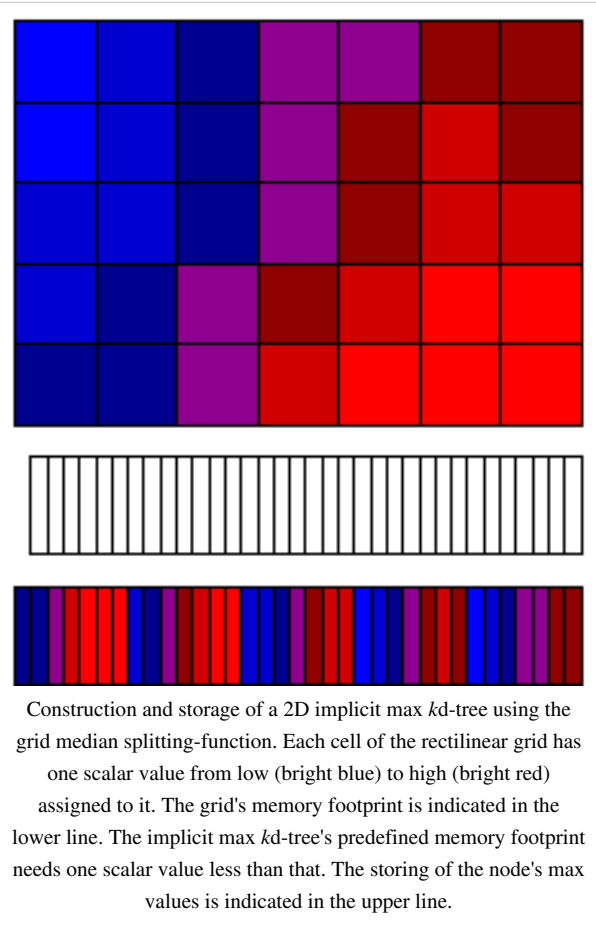
- [libkdtree++](http://libkdtree.alioth.debian.org) (<http://libkdtree.alioth.debian.org>), an open-source STL-like implementation of  $k$ -d trees in C++.
- A tutorial on KD Trees (<http://www.autonlab.org/autonweb/14665/version/2/part/5/data/moore-tutorial.pdf?branch=main&language=en>)
- FLANN (<http://people.cs.ubc.ca/~mariusm/index.php/FLANN/FLANN>) and its fork nanoflann (<http://code.google.com/p/nanoflann/>), efficient C++ implementations of  $k$ -d tree algorithms.
- kdtree (<http://code.google.com/p/kdtree/>) A simple C library for working with KD-Trees
- K-D Tree Demo, Java applet (<http://donar.umiacs.umd.edu/quadtree/points/kdtree.html>)
- libANN (<http://www.cs.umd.edu/~mount/ANN/>) Approximate Nearest Neighbour Library includes a  $k$ -d tree implementation
- Caltech Large Scale Image Search Toolbox (<http://www.vision.caltech.edu/malaa/software/research/image-search/>): a Matlab toolbox implementing randomized  $k$ -d tree for fast approximate nearest neighbour search, in addition to LSH, Hierarchical K-Means, and Inverted File search algorithms.
- Heuristic Ray Shooting Algorithms (<http://dcgi.felk.cvut.cz/home/havran/phdthesis.html>), pp. 11 and after
- Into (<http://intopii.com/into/>) contains open source implementations of exact and approximate (k)NN search methods using  $k$ -d trees in C++.
- Math::Vector::Real::kdTree (<http://https://metacpan.org/module/Math::Vector::Real::kdTree>) Perl implementation of  $k$ -d trees.

# Implicit $k$ -d tree

An **implicit  $k$ -d tree** is a  $k$ -d tree defined implicitly above a rectilinear grid. Its split planes' positions and orientations are not given explicitly but implicitly by some recursive splitting-function defined on the hyperrectangles belonging to the tree's nodes. Each inner node's split plane is positioned on a grid plane of the underlying grid, partitioning the node's grid into two subgrids.

## Nomenclature and references

The terms "min/max  $k$ -d tree" and "implicit  $k$ -d tree" are sometimes mixed up. This is because the first publication using the term "implicit  $k$ -d tree" <sup>[1]</sup> did actually use explicit min/max  $k$ -d trees but referred to them as "implicit  $k$ -d trees" to indicate that they may be used to ray trace implicitly given iso surfaces. Nevertheless this publication used also slim  $k$ -d trees which are a subset of the implicit  $k$ -d trees with the restriction that they can only be built over integer hyperrectangles with sidelengths that are powers of two. Implicit  $k$ -d trees as defined here have recently been introduced, with applications in computer graphics.<sup>[2][3]</sup> As it is possible to assign attributes to implicit  $k$ -d tree nodes, one may refer to an implicit  $k$ -d tree which has min/max values assigned to its nodes as an "implicit min/max  $k$ -d tree".



## Construction

Implicit  $k$ -d trees are in general not constructed explicitly. When accessing a node, its split plane orientation and position are evaluated using the specific splitting-function defining the tree. Different splitting-functions may result in different trees for the same underlying grid.

## Splitting-functions

Splitting-functions may be adapted to special purposes. Underneath two specifications of special splitting-function classes.

- **Non-degenerated splitting-functions** do not allow the creation of degenerated nodes (nodes whose corresponding integer hyperrectangle's volume is equal zero). Their corresponding implicit  $k$ -d trees are full binary trees, which have for  $n$  leaf nodes  $n - 1$  inner nodes. Their corresponding implicit  $k$ -d trees are **non-degenerated implicit  $k$ -d trees**.
- **complete splitting-functions** are non-degenerated splitting-functions whose corresponding implicit  $k$ -d tree's leaf nodes are single grid cells such that they have one inner node less than the amount of gridcells given in the grid. The corresponding implicit  $k$ -d trees are **complete implicit  $k$ -d trees**.

A complete splitting function is for example the **grid median splitting-function**. It creates fairly balanced implicit  $k$ -d trees by using  $k$ -dimensional integer hyperrectangles  $hyprec[2][k]$  belonging to each node of the implicit  $k$ -d tree. The hyperrectangles define which gridcells of the rectilinear grid belong to their corresponding node. If the volume of this hyperrectangle equals one, the corresponding node is a single grid cell and is therefore not further subdivided and marked as leaf node. Otherwise the hyperrectangle's longest extend is chosen as orientation  $o$ . The corresponding split plane  $p$  is positioned onto the grid plane that is closest to the hyperrectangle's grid median along that orientation.

Split plane orientation  $o$ :

```
o = min{argmax(i = 1 ... k: (hyprec[1][i] - hyprec[0][i]))}
```

Split plane position  $p$ :

```
p = roundDown((hyprec[0][o] + hyprec[1][o]) / 2)
```

### Assigning attributes to implicit $k$ -d tree nodes

An obvious advantage of implicit  $k$ -d trees is that their split plane's orientations and positions need not to be stored explicitly.

But some applications require besides the split plane's orientations and positions further attributes at the inner tree nodes. These attributes may be for example single bits or single scalar values, defining if the subgrids belonging to the nodes are of interest or not. For complete implicit  $k$ -d trees it is possible to pre-allocate a correctly sized array of attributes and to assign each inner node of the tree to a unique element in that allocated array.

The amount of gridcells in the grid is equal the volume of the integer hyperrectangle belonging to the grid. As a complete implicit  $k$ -d tree has one inner node less than grid cells, it is known in advance how many attributes need to be stored. The relation "*Volume of integer hyperrectangle to inner nodes*" defines together with the complete splitting-function a recursive formula assigning to each split plane a unique element in the allocated array. The corresponding algorithm is given in C-pseudo code underneath.

```
// Assigning attributes to inner nodes of a complete implicit k-d tree

// create an integer help hyperrectangle hyprec (its volume vol(hyprec)
// is equal the amount of leaves)
int hyprec[2][k] = { { 0, ..., 0 }, { length_1, ..., length_k } };
// allocate once the array of attributes for the entire implicit k-d
// tree
attr *a = new attr[volume(hyprec) - 1];

attr implicitKdTreeAttributes(int hyprec[2][k], attr *a)
{
    if(vol(hyprec) > 1) // the current node is an inner node
    {
        // evaluate the split plane's orientation o and its position p using
        // the underlying complete split-function
        int o, p;
        completeSplittingFunction(hyprec, &o, &p);
        // evaluate the children's integer hyperrectangles hyprec_l and
        // hyprec_r
        int hyprec_l[2][k], hyprec_r[2][k];
        hyprec_l = hyprec;
```

```

hyprec_l[1][o] = p;
hyprec_r      = hyprec;
hyprec_r[0][o] = p;
// evaluate the children's memory location a_l and a_r
attr* a_l = a + 1;
attr* a_r = a + vol(hyprec_l);
// evaluate recursively the children's attributes c_l and c_r
attr c_l = implicitKdTreeAttributes(hyprec_l, a_l);
attr c_r = implicitKdTreeAttributes(hyprec_r, a_r);
// merge the children's attributes to the current attribute c
attr c = merge(c_l, c_r);
// store the current attribute and return it
a[0] = c;
return c;
}
// The current node is a leaf node. Return the attribute belonging to
the corresponding gridcell
return attribute(hyprec);
}

```

It is worth mentioning that this algorithm works for all rectilinear grids. The corresponding integer hyperrectangle does not necessarily have to have sidelengths that are powers of two.

## Applications

Implicit max- $k$ -d trees are used for ray casting isosurfaces/MIP (maximum intensity projection). The attribute assigned to each inner node is the maximal scalar value given in the subgrid belonging to the node. Nodes are not traversed if their scalar values are smaller than the searched iso-value/current maximum intensity along the ray. The low storage requirements of the implicit max  $kd$ -tree and the favorable visualization complexity of ray casting allow to ray cast (and even change the isosurface for) very large scalar fields at interactive framerates on commodity PCs. Similarly an implicit min/max  $kd$ -tree may be used to efficiently evaluate queries such as terrain line of sight.<sup>[4]</sup>

## Complexity

Given an implicit  $k$ -d tree spanned over an  $k$ -dimensional grid with  $n$  gridcells.

- Assigning attributes to the nodes of the tree takes  $O(kn)$  time.
- Storing attributes to the nodes takes  $O(n)$  memory.
- Ray casting iso-surfaces/MIP an underlying scalar field using the corresponding implicit max  $k$ -d tree takes roughly  $O(\log(n))$  time.

## References

- [1] Ingo Wald, Heiko Friedrich, Gerd Marmitt, Philipp Slusallek and Hans-Peter Seidel "Faster Isosurface Ray Tracing using Implicit KD-Trees" IEEE Transactions on Visualization and Computer Graphics (2005)
- [2] Matthias Groß, Carsten Lojewski, Martin Bertram and Hans Hagen "Fast Implicit  $k$ -d Trees: Accelerated Isosurface Ray Tracing and Maximum Intensity Projection for Large Scalar Fields" CGIM07: Proceedings of Computer Graphics and Imaging (2007) 67-74
- [3] Matthias Groß (PhD, 2009) Towards Scientific Applications for Interactive Ray Casting (<http://kluedo.ub.uni-kl.de/volltexte/2009/2361/>)
- [4] Bernardt Duvénhage "Using An Implicit Min/Max KD-Tree for Doing Efficient Terrain Line of Sight Calculations" in "Proceedings of the 6th International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa", 2009.

## min/max kd-tree

---

A **min/max kd-tree** is a  $kd$ -tree with two scalar values - a minimum and a maximum - assigned to its nodes. The minimum/maximum of an inner node is equal the minimum/maximum of its children's minima/maxima.

### Construction

Min/max  $kd$ -trees may be constructed recursively. Starting with the root node, the splitting plane orientation and position is evaluated. Then the children's splitting planes and min/max values are evaluated recursively. The min/max value of the current node is simply the minimum/maximum of its children's minima/maxima.

### Properties

The min/max  $kd$ tree has - besides the properties of an  $kd$ -tree - the special property that an inner node's min/max values coincide each with a min/max value of either one child. This allows to discard the storage of min/max values at the leaf nodes by storing two bits at inner nodes, assigning min/max values to the children: Each inner node's min/max values will be known in advance, where the root node's min/max values are stored separately. Each inner node has besides two min/max values also two bits given, defining to which child those min/max values are assigned (0: to the left child 1: to the right child). The non-assigned min/max values of the children are the from the current node already known min/max values. The two bits may also be stored in the least significant bits of the min/max values which have therefore to be approximated by fractioning them down/up.

The resulting memory reduction is not minor, as the leaf nodes of full binary  $kd$ -trees are one half of the tree's nodes.

### Applications

Min/max  $kd$ -trees are used for ray casting isosurfaces/MIP (maximum intensity projection). Isosurface ray casting only traverses nodes for which the chosen isovalue lies between the min/max values of the current node. Nodes that do not fulfill this requirement do not contain an isosurface to the given isovalue and are therefore skipped (empty space skipping). For MIP, nodes are not traversed if their maximum is smaller than the current maximum intensity along the ray. The favorable visualization complexity of ray casting allows to ray cast (and even change the isosurface for) very large scalar fields at interactive framerates on commodity PCs. Especially implicit max  $kd$ -trees are an optimal choice for visualizing scalar fields defined on rectilinear grids (see <sup>[1][2][3]</sup>). Similarly an implicit min/max  $kd$ -tree may be used to efficiently evaluate queries such as terrain line of sight<sup>[4]</sup>.

## References

- [1] Matthias Groß, Carsten Lojewski, Martin Bertram and Hans Hagen "Fast Implicit KD-Trees: Accelerated Isosurface Ray Tracing and Maximum Intensity Projection for Large Scalar Fields" CGIM07: Proceedings of Computer Graphics and Imaging (2007) 67-74
- [2] Ingo Wald, Heiko Friedrich, Gerd Marmitt, Philipp Slusallek and Hans-Peter Seidel "Faster Isosurface Ray Tracing using Implicit KD-Trees" IEEE Transactions on Visualization and Computer Graphics (2005)
- [3] Matthias Groß (PhD, 2009) Towards Scientific Applications for Interactive Ray Casting (<http://kluedo.ub.uni-kl.de/volltexte/2009/2361/>)
- [4] Bernardt Duvenhage "Using An Implicit Min/Max KD-Tree for Doing Efficient Terrain Line of Sight Calculations" in "Proceedings of the 6th International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa", 2009.

# Adaptive k-d tree

---

An **adaptive k-d tree** is a tree for multidimensional points where successive levels may be split along different dimensions.

## References

- Samet, Hanan (2006). *Foundations of multidimensional and metric data structures* <sup>[1]</sup>. Morgan Kaufmann. ISBN 978-0-12-369446-1.

Paul E. Black, Adaptive k-d tree <sup>[2]</sup> at the NIST Dictionary of Algorithms and Data Structures.

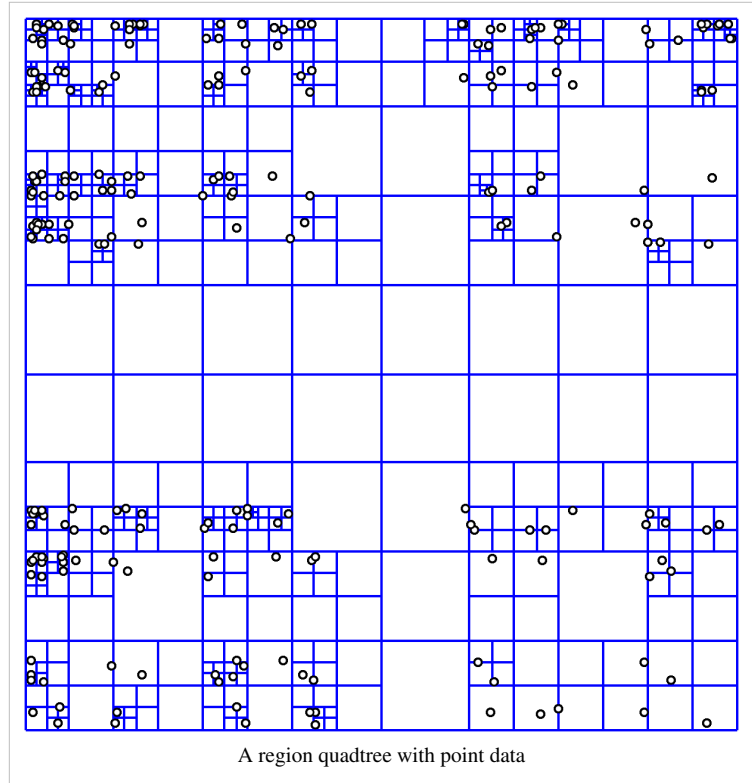
## References

- [1] <http://books.google.dk/books?id=KrQdmLjTSaQC>
- [2] <http://www.nist.gov/dads/HTML/adaptkdtree.html>

# Quadtree

A **quadtree** is a tree data structure in which each internal node has exactly four children. Quadtrees are most often used to partition a two dimensional space by recursively subdividing it into four quadrants or regions. The regions may be square or rectangular, or may have arbitrary shapes. This data structure was named a quadtree by Raphael Finkel and J.L. Bentley in 1974. A similar partitioning is also known as a *Q-tree*. All forms of Quadtrees share some common features:

- They decompose space into adaptable cells
- Each cell (or bucket) has a maximum capacity. When maximum capacity is reached, the bucket splits
- The tree directory follows the spatial decomposition of the Quadtree.



## Types

Quadtrees may be classified according to the type of data they represent, including areas, points, lines and curves. Quadtrees may also be classified by whether the shape of the tree is independent of the order data is processed. Some common types of quadtrees are:

### The region quadtree

The region quadtree represents a partition of space in two dimensions by decomposing the region into four equal quadrants, subquadrants, and so on with each leaf node containing data corresponding to a specific subregion. Each node in the tree either has exactly four children, or has no children (a leaf node). The region quadtree is not strictly a 'tree' - as the positions of subdivisions are independent of the data. They are more precisely called 'tries'.

A region quadtree with a depth of  $n$  may be used to represent an image consisting of  $2^n \times 2^n$  pixels, where each pixel value is 0 or 1. The root node represents the entire image region. If the pixels in any region are not entirely 0s or 1s, it is subdivided. In this application, each leaf node represents a block of pixels that are all 0s or all 1s.

A region quadtree may also be used as a variable resolution representation of a data field. For example, the temperatures in an area may be stored as a quadtree, with each leaf node storing the average temperature over the subregion it represents.

If a region quadtree is used to represent a set of point data (such as the latitude and longitude of a set of cities), regions are subdivided until each leaf contains at most a single point.



## Point quadtree

The point quadtree is an adaptation of a binary tree used to represent two dimensional point data. It shares the features of all quadtrees but is a true tree as the center of a subdivision is always on a point. The tree shape depends on the order data is processed. It is often very efficient in comparing two dimensional ordered data points, usually operating in  $O(\log n)$  time.

### Node structure for a point quadtree

A node of a point quadtree is similar to a node of a binary tree, with the major difference being that it has four pointers (one for each quadrant) instead of two ("left" and "right") as in an ordinary binary tree. Also a key is usually decomposed into two parts, referring to x and y coordinates. Therefore a node contains following information:

- 4 Pointers: quad['NW'], quad['NE'], quad['SW'], and quad['SE']
- point; which in turn contains:
  - key; usually expressed as x, y coordinates
  - value; for example a name

## Edge quadtree

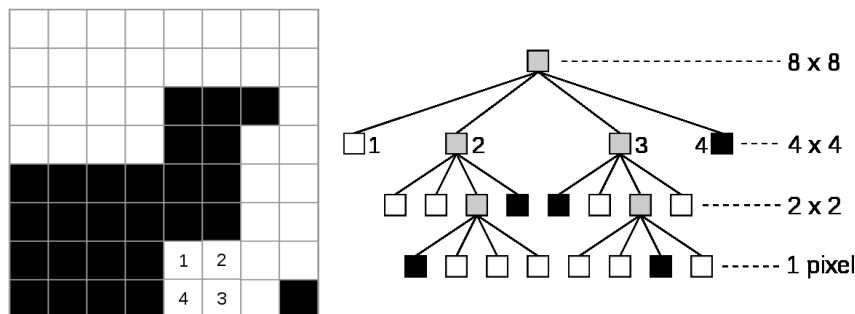
Edge quadtrees are specifically used to store lines rather than points. Curves are approximated by subdividing cells to a very fine resolution. This can result in extremely unbalanced trees which may defeat the purpose of indexing.

## Polygonal Map Quadtree

The Polygonal Map Quadtree (or PMQuadtree) is a variation of quadtrees which are used to store collections of polygons that may be degenerate (meaning that they have isolated vertices or edges).<sup>[1]</sup>

## Some common uses of quadtrees

- Image representation



- Spatial indexing
- Efficient collision detection in two dimensions
- View frustum culling of terrain data
- Storing sparse data, such as a formatting information for a spreadsheet or for some matrix calculations
- Solution of multidimensional fields (computational fluid dynamics, electromagnetism)
- Conway's Game of Life simulation program.<sup>[2]</sup>
- State estimation<sup>[3]</sup>
- Quadtrees are also used in the area of fractal image analysis

Quadtrees are the two-dimensional analog of octrees.

## Pseudo code

The following pseudo code shows one means of implementing a quadtree which handles only points. There are other approaches available.

## Prerequisites

It is assumed these structures are used.

```
// Simple coordinate object to represent points and vectors
struct XY
{
    float x;
    float y;

    function __construct(float _x, float _y) {...}
}

// Axis-aligned bounding box with half dimension and center
struct AABB
{
    XY center;
    XY halfDimension;

    function __construct(XY center, XY halfDimension) {...}
    function containsPoint(XY p) {...}
    function intersectsAABB(AABB other) {...}
}
```

## QuadTree class

This class represents both one quad tree and the node where it is rooted.

```
class QuadTree
{
    // Arbitrary constant to indicate how many elements can be stored in this quad tree node
    constant int QT_NODE_CAPACITY = 4;

    // Axis-aligned bounding box stored as a center with half-dimensions
    // to represent the boundaries of this quad tree
    AABB boundary;

    // Points in this quad tree node
    Array of XY [size = QT_NODE_CAPACITY] points;

    // Children
    QuadTree* northWest;
    QuadTree* northEast;
    QuadTree* southWest;
    QuadTree* southEast;
```

```

// Methods

function __construct(AABB _boundary) {...}

function insert(XY p) {...}

function subdivide() {...} // create four children which fully divide this quad into four quads of equal area

function queryRange(AABB range) {...}
}

```

## Insertion

The following method inserts a point into the appropriate quad of a quadtree, splitting if necessary.

```

class QuadTree
{
...

// Insert a point into the QuadTree
function insert(XY p)
{
// Ignore objects which do not belong in this quad tree
if (!boundary.containsPoint(p))
    return false; // object cannot be added

// If there is space in this quad tree, add the object here
if (points.size < QT_NODE_CAPACITY)
{
    points.append(p);
    return true;
}

// Otherwise, we need to subdivide then add the point to whichever node will accept it
if (northWest != null)
    subdivide();

if (northWest->insert(p)) return true;
if (northEast->insert(p)) return true;
if (southWest->insert(p)) return true;
if (southEast->insert(p)) return true;

// Otherwise, the point cannot be inserted for some unknown reason (which should never happen)
return false;
}
}

```

## Query range

The following method finds all points which are contained within a range.

```
class QuadTree
{
    ...

    // Find all points which appear within a range
    function queryRange(AABB range)
    {
        // Prepare an array of results
        Array of XY pointsInRange;

        // Automatically abort if the range does not collide with this quad
        if (!boundary.intersectsAABB(range))
            return pointsInRange; // empty list

        // Check objects at this quad level
        for (int p := 0; p < points.size; p++)
        {
            if (range.containsPoint(points[p]))
                pointsInRange.append(points[p]);
        }

        // Terminate here, if there are no children
        if (northWest == null)
            return pointsInRange;

        // Otherwise, add the points from the children
        pointsInRange.appendArray(northWest->queryRange(range));
        pointsInRange.appendArray(northEast->queryRange(range));
        pointsInRange.appendArray(southWest->queryRange(range));
        pointsInRange.appendArray(southEast->queryRange(range));

        return pointsInRange;
    }
}
```

## References

### Notes

- [1] Hanan Samet and Robert Webber. "Storing a Collection of Polygons Using Quadtrees". *ACM Transactions on Graphics* July 1985: 182-222. *InfoLAB*. Web. 23 March 2012
- [2] Tomas G. Rokicki (2006-04-01). "An Algorithm for Compressing Space and Time" (<http://www.ddj.com/hpc-high-performance-computing/184406478>). . Retrieved 2009-05-20.
- [3] Henning Eberhardt, Vesa Klumpp, Uwe D. Hanebeck, *Density Trees for Efficient Nonlinear State Estimation*, Proceedings of the 13th International Conference on Information Fusion, Edinburgh, United Kingdom, July, 2010.

### General references

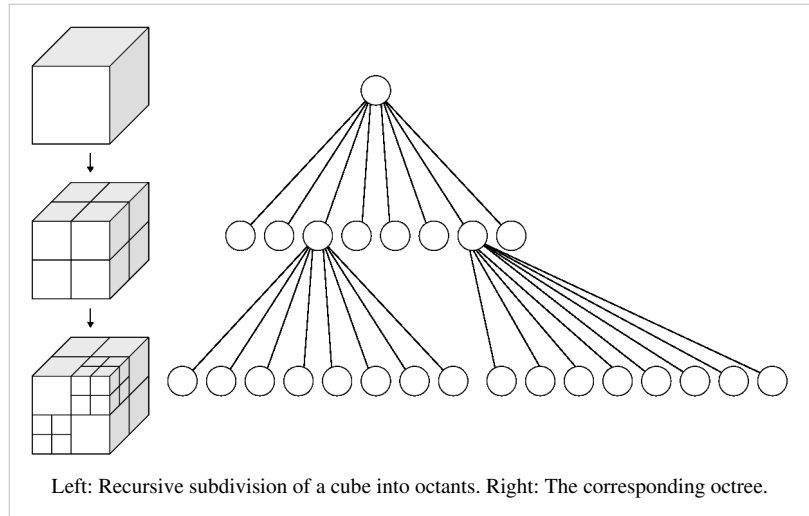
1. Raphael Finkel and J.L. Bentley (1974). "Quad Trees: A Data Structure for Retrieval on Composite Keys". *Acta Informatica* **4** (1): 1–9. doi:10.1007/BF00288933.
2. Mark de Berg, Marc van Kreveld, Mark Overmars, and Otfried Schwarzkopf (2000). *Computational Geometry* (2nd revised ed.). Springer-Verlag. ISBN 3-540-65620-0. Chapter 14: Quadrees: pp. 291–306.
3. Samet, Hanan; Webber, Robert (July 1985). "Storing a Collection of Polygons Using Quadtrees" ([http://infolab.usc.edu/csci585/Spring2008/den\\_ar/p182-samet.pdf](http://infolab.usc.edu/csci585/Spring2008/den_ar/p182-samet.pdf)). Retrieved 23 March 2012.

### External links

- A discussion of the Quadtree and an application (<http://www.cs.berkeley.edu/~demmell/cs267/lecture26/lecture26.html>)
  - Considerable discussion and demonstrations of Spatial Indexing (<http://homepages.ge.ucl.ac.uk/~mhaklay/java.htm>)
  - Example C# code for a quad tree (<http://digitseven.com/QuadTree.aspx>)
  - Javascript Implementation of the QuadTree used for collision detection (<http://www.mikechambers.com/blog/2011/03/21/javascript-quadtree-implementation/>)
  - C++ Implementation of a QuadTree used for spatial indexing of triangles (<http://sourceforge.net/projects/quadtreesim/>)
-

# Octree

An **octree** is a tree data structure in which each internal node has exactly eight children. Octrees are most often used to partition a three dimensional space by recursively subdividing it into eight octants. Octrees are the three-dimensional analog of quadrees. The name is formed from *oct* + *tree*, but note that it is normally written "*octree*" with only one "t". Octrees are often used in 3D graphics and 3D game engines.



## Octrees for spatial representation

Each node in an octree subdivides the space it represents into eight octants. In a point region (PR) octree, the node stores an explicit 3-dimensional point, which is the "center" of the subdivision for that node; the point defines one of the corners for each of the eight children. In an MX octree, the subdivision point is implicitly the center of the space the node represents. The root node of a PR octree can represent infinite space; the root node of an MX octree must represent a finite bounded space so that the implicit centers are well-defined. Octrees are never considered kD-trees, as kD-trees split along a dimension and octrees split around a point. kD-trees are also always binary, which is not true of octrees. By using DFS the nodes are to be traversed and only required surfaces are to be viewed.

## Common uses of octrees

- Spatial indexing
- Nearest neighbor search
- Efficient collision detection in three dimensions
- View frustum culling
- Fast Multipole Method
- Unstructured grid
- Finite element analysis
- Sparse voxel octree
- State estimation<sup>[1]</sup>

## Application to color quantization

The octree color quantization algorithm, invented by Gervautz and Purgathofer in 1988, encodes image color data as an octree up to nine levels deep. Octrees are used because  $2^3 = 8$  and there are three color components in the RGB system. The node index to branch out from at the top level is determined by a formula that uses the most significant bits of the red, green, and blue color components, e.g.  $4r + 2g + b$ . The next lower level uses the next bit significance, and so on. Less significant bits are sometimes ignored to reduce the tree size.

The algorithm is highly memory efficient because the tree's size can be limited. The bottom level of the octree consists of leaf nodes that accrue color data not represented in the tree; these nodes initially contain single bits. If

much more than the desired number of palette colors are entered into the octree, its size can be continually reduced by seeking out a bottom-level node and averaging its bit data up into a leaf node, pruning part of the tree. Once sampling is complete, exploring all routes in the tree down to the leaf nodes, taking note of the bits along the way, will yield approximately the required number of colors.

## References

- [1] Henning Eberhardt, Vesa Klumpp, Uwe D. Hanebeck, *Density Trees for Efficient Nonlinear State Estimation*, Proceedings of the 13th International Conference on Information Fusion, Edinburgh, United Kingdom, July, 2010. ([http://isas.uka.de/Publikationen/Fusion10\\_EberhardtKlumpp.pdf](http://isas.uka.de/Publikationen/Fusion10_EberhardtKlumpp.pdf))

## External links

- Octree Quantization in Microsoft Systems Journal (<http://www.microsoft.com/msj/archive/S3F1.aspx>)
- Color Quantization using Octrees in Dr. Dobb's (<http://www.ddj.com/184409805>)
- Color Quantization using Octrees in Dr. Dobb's Source Code (<ftp://ftp.drdobbs.com/sourcecode/ddj/1996/9601.zip>)
- Octree Color Quantization Overview ([http://web.cs.wpi.edu/~matt/courses/cs563/talks/color\\_quant/CQoctree.html](http://web.cs.wpi.edu/~matt/courses/cs563/talks/color_quant/CQoctree.html))
- Parallel implementation of octree generation algorithm, P. Sojan Lal, A Unnikrishnan, K Poulose Jacob, ICIP 1997, IEEE Digital Library ([http://ieeexplore.ieee.org/xpl/freeabs\\_all.jsp?arnumber=727419](http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=727419))
- Generation of Octrees from Raster Scan with Reduced Information Loss, P. Sojan Lal, A Unnikrishnan, K Poulose Jacob, IASTED International conference VIIP 2001 (<http://dblp.uni-trier.de/db/conf/viip/viip2001.html#LalUJ01>) ([http://www.actapress.com/catalogue2009/proc\\_series13.html#viip2001](http://www.actapress.com/catalogue2009/proc_series13.html#viip2001))
- C++ implementation (GPL license) (<http://nomis80.org/code/octree.html>)
- Parallel Octrees for Finite Element Applications (<http://sc07.supercomputing.org/schedule/pdf/pap117.pdf>)
- Cube 2: Sauerbraten - a game written in the octree-heavy Cube 2 engine (<http://www.sauerbraten.org/>)
- Ogre - A 3d Object-oriented Graphics Rendering Engine with a Octree Scene Manager Implementation (MIT license) (<http://www.ogre3d.org>)
- Dendro: parallel multigrid for octree meshes (MPI/C++ implementation) (<http://www.cc.gatech.edu/csela/dendro>)
- **Video:** Use of an octree in state estimation (<http://www.youtube.com/watch?v=Jw4VAgcWruY>)

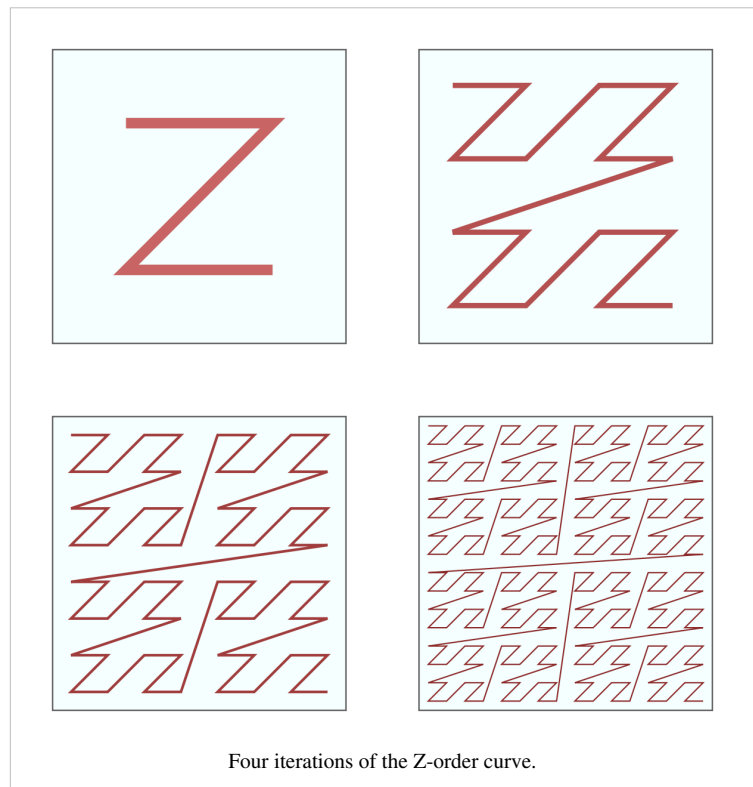
# Linear octrees

A **linear octree** is an octree that is represented by a linear array instead of a tree data structure.

To simplify implementation, a linear octree is usually complete (that is, every internal node has exactly 8 child nodes) and where the maximum permissible depth is fixed a priori (making it sufficient to store the complete list of leaf nodes). That is, all the nodes of the octree can be generated from the list of its leaf nodes. Space filling curves are often used to represent linear octrees.

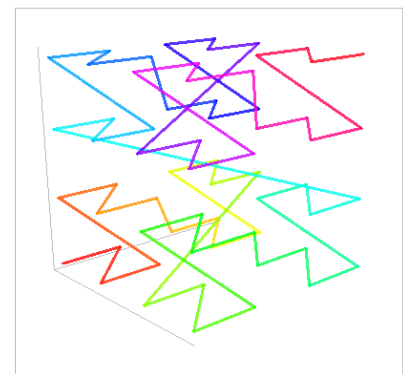
## Z-order

In mathematical analysis and computer science, **Z-order**, **Morton order**, or **Morton code** is a function which maps multidimensional data to one dimension while preserving locality of the data points. It was introduced in 1966 by G. M. Morton.<sup>[1]</sup> The z-value of a point in multidimensions is simply calculated by interleaving the binary representations of its coordinate values. Once the data are sorted into this ordering, any one-dimensional data structure can be used such as binary search trees, B-trees, skip lists or (with low significant bits truncated) hash tables. The resulting ordering can equivalently be described as the order one would get from a depth-first traversal of a quadtree; because of its close connection with quadtrees, the Z-ordering can be used to efficiently construct quadtrees and related higher dimensional data structures.<sup>[2]</sup>

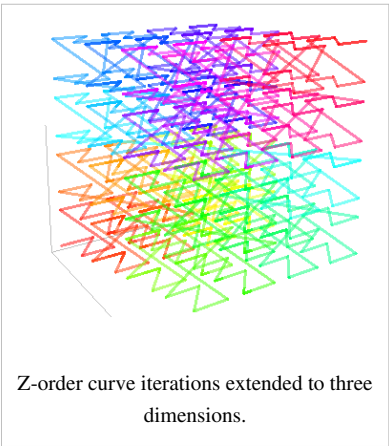


### Coordinate values

The figure below shows the Z-values for the two dimensional case with integer coordinates  $0 \leq x \leq 7$ ,  $0 \leq y \leq 7$  (shown both in decimal and binary). Interleaving the binary coordinate values yields binary z-values as shown. Connecting the z-values in their numerical order produces the recursively Z-shaped curve.







	x:	0	1	2	3	4	5	6	7
		000	001	010	011	100	101	110	111
y: 0		000000	000001	000100	000101	010000	010001	010100	010101
1		000010	000011	000110	000111	010010	010011	010110	010111
2		001000	001001	001100	001101	011000	011001	011100	011101
3		001010	001011	001110	001111	011010	011011	011110	011111
4		100000	100001	100100	100101	110000	110001	110100	110101
5		100010	100011	100110	100111	110010	110011	110110	110111
6		101000	101001	101100	101101	111000	111001	111100	111101
7		101010	101011	101110	101111	111010	111011	111110	111111

Efficiently building quadtrees

As mentioned, the Z-ordering can be used to efficiently build a quadtree for a set of points. The basic idea is to sort the input set according to Z-order. Once sorted, the points can either be stored in a binary search tree and used directly, which is called a linear quadtree,<sup>[3]</sup> or they can be used to build a pointer based quadtree.

The input points are usually scaled in each dimension to be positive integers, either as a fixed point representation over the unit range [0, 1] or corresponding to the machine word size. Both representations are equivalent and allow for the highest order non-zero bit to be found in constant time. Each square in the quadtree has a side length which is a power of two, and corner coordinates which are multiples of the side length. Given any two points, the *derived square* for the two points is the smallest square covering both points. The interleaving of bits from the x and y components of each point is called the *shuffle* of x and y, and can be extended to higher dimensions.<sup>[2]</sup>

Points can be sorted according to their shuffle without explicitly interleaving the bits. To do this, for each dimension, the most significant bit of the exclusive or of the coordinates of the two points for that dimension is examined. The dimension for which the most significant bit is largest is then used to compare the two points to determine their shuffle order.

The exclusive or operation masks off the higher order bits for which the two coordinates are identical. Since the shuffle interleaves bits from higher order to lower order, identifying the coordinate with the largest most significant bit, identifies the first bit in the shuffle order which differs, and that coordinate can be used to compare the two points.<sup>[4]</sup> This is shown in the following Python code:

```
def cmp_zorder(a, b):
    j = 0
    k = 0
    x = 0
    for k in range(dim):
        y = a[k] ^ b[k]
        if less_msb(x, y):
            j = k
            x = y
    return a[j] - b[j]
```

One way to determine whether the most significant smaller is to compare the floor of the base-2 logarithm of each point. It turns out the following operation is equivalent, and only requires exclusive or operations<sup>[4]</sup>:

```
def less_msb(x, y):
    return x < y and x < (x ^ y)
```

It is also possible to compare floating point numbers using the same technique. The *less\_msb* function is modified to first compare the exponents. Only when they are equal is the standard *less\_msb* function used on the mantissas.<sup>[5]</sup>

Once the points are in sorted order, two properties make it easy to build a quadtree: The first is that the points contained in a square of the quadtree form a contiguous interval in the sorted order. The second is that if more than one child of a square contains an input point, the square is the *derived square* for two adjacent points in the sorted order.

For each adjacent pair of points, the derived square is computed and its side length determined. For each derived square, the interval containing it is bounded by the first larger square to the right and to the left in sorted order.<sup>[2]</sup> Each such interval corresponds to a square in the quadtree. The result of this is a compressed quadtree, where only nodes containing input points or two or more children are present. A non-compressed quadtree can be built by restoring the missing nodes, if desired.

Rather than building a pointer based quadtree, the points can be maintained in sorted order in a data structure such as a binary search tree. This allows points to be added and deleted in  $O(\log n)$  time. Two quadtrees can be merged by merging the two sorted sets of points, and removing duplicates. Point location can be done by searching for the points preceding and following the query point in the sorted order. If the quadtree is compressed, the predecessor node found may be an arbitrary leaf inside the compressed node of interest. In this case, it is necessary to find the predecessor of the least common ancestor of the query point and the leaf found.<sup>[6]</sup>

## Use with one-dimensional data structures for range searching

Although preserving locality well, for efficient range searches an algorithm is necessary for calculating, from a point encountered in the data structure, the next Z-value which is in the multidimensional search range:

x=	0	1	2	3	4	5	6	7
y= 0	0	1	4	5	16	17	<b>20</b>	<del>21</del>
y= 1	2	3	6	7	18	<b>19</b>	22	23
y= 2	8	9	12	13	24	25	28	29
y= 3	10	11	14	15	26	27	30	31
y= 4	<del>32</del>	<del>33</del>	<b>36</b>	<b>37</b>	48	49	52	53
y= 5	<del>34</del>	<del>35</del>	38	39	50	51	54	55
y= 6	<del>40</del>	<del>41</del>	<del>44</del>	<del>45</del>	56	57	60	61
y= 7	<del>42</del>	<del>43</del>	46	47	58	59	62	63

In this example, the range being queried ( $x = 2, \dots, 3, y = 2, \dots, 6$ ) is indicated by the dotted rectangle. Its highest Z-value (MAX) is 45. In this example, the value  $F = 19$  is encountered when searching a data structure in increasing Z-value direction, so we would have to search in the interval between  $F$  and MAX (hatched area). To speed up the search, one would calculate the next Z-value which is in the search range, called BIGMIN (36 in the example) and only search in the interval between BIGMIN and MAX (bold values), thus skipping most of the hatched area. Searching in decreasing direction is analogous with LITMAX which is the highest Z-value in the query range lower than  $F$ . The BIGMIN problem has first been stated and its solution shown in Tropf and Herzog.<sup>[7]</sup> This solution is also used in UB-trees ("GetNextZ-address"). As the approach does not depend on the one dimensional data structure chosen, there is still free choice of structuring the data, so well known methods such as balanced trees can be used to cope with dynamic data (in contrast for example to R-trees where special considerations are necessary). Similarly, this independence makes it easier to incorporate the method into existing databases.

Applying the method hierarchically (according to the data structure at hand), optionally in both increasing and decreasing direction, yields highly efficient multidimensional range search which is important in both commercial and technical applications, e.g. as a procedure underlying nearest neighbour searches. Z-order is one of the few multidimensional access methods that has found its way into commercial database systems (Oracle database 1995,<sup>[8]</sup> Transbase 2000<sup>[9]</sup>).

As long ago as 1966, G.M.Morton proposed Z-order for file sequencing of a static two dimensional geographical database. Areal data units are contained in one or a few quadratic frames represented by their sizes and lower right corner Z-values, the sizes complying with the Z-order hierarchy at the corner position. With high probability, changing to an adjacent frame is done with one or a few relatively small scanning steps.

## Related structures

As an alternative, the Hilbert curve has been suggested as it has a better order-preserving behaviour, but here the calculations are much more complicated, leading to significant processor overhead. BIGMIN source code for both Z-curve and Hilbert-curve were described in a patent by H. Tropf.<sup>[10]</sup>

For a recent overview on multidimensional data processing, including e.g. nearest neighbour searches, see Hanan Samet's textbook.<sup>[11]</sup>

## Applications in linear algebra

The Strassen algorithm for matrix multiplication is based on splitting the matrices in four blocks, and then recursively each of these blocks in four smaller blocks, until the blocks are single elements (or more practically: until reaching matrices so small that the trivial algorithm is faster). Arranging the matrix elements in Z-order then improves locality, and has the additional advantage (compared to row- or column-major ordering) that the subroutine for multiplying two blocks does not need to know the total size of the matrix, but only the size of the blocks and their location in memory. Effective use of Strassen multiplication with Z-order has been demonstrated, see Valsalam and Skjellum's 2002 paper<sup>[12]</sup>.

## References

- [1] Morton, G. M. (1966), *A computer Oriented Geodetic Data Base; and a New Technique in File Sequencing*, Technical Report, Ottawa, Canada: IBM Ltd.
- [2] Bern, M.; Eppstein, D.; Teng, S.-H. (1999), "Parallel construction of quadrees and quality triangulations", *Int. J. Comp. Geom. & Appl.* **9** (6): 517–532, doi:10.1142/S0218195999000303.
- [3] Gargantini, I. (1982), "An effective way to represent quadrees", *Communications of the ACM* **25** (12): 905–910.
- [4] Chan, T. (2002), "Closest-point problems simplified on the RAM" ([http://www.cs.uwaterloo.ca/~tmchan/ram\\_soda.ps.gz](http://www.cs.uwaterloo.ca/~tmchan/ram_soda.ps.gz)), *ACM-SIAM Symposium on Discrete Algorithms*, .
- [5] Connor, M.; Kumar, P (2009), "Fast construction of k-nearest neighbour graphs for point clouds" ([http://compgeom.com/~piyush/papers/tvcg\\_stann.pdf](http://compgeom.com/~piyush/papers/tvcg_stann.pdf)), *IEEE Transactions on Visualization and Computer Graphics*,
- [6] Har-Peled, S. (2010), *Data structures for geometric approximation* ([http://www.madalgo.au.dk/img/SS2010/Course Material/Data-Structures for Geometric Approximation by Sariel Har-Peled.pdf](http://www.madalgo.au.dk/img/SS2010/Course%20Material/Data-Structures%20for%20Geometric%20Approximation%20by%20Sariel%20Har-Peled.pdf)),
- [7] Tropf, H.; Herzog, H. (1981), "Multidimensional Range Search in Dynamically Balanced Trees" (<http://www.vision-tools.com/h-tropf/multidimensionalrangequery.pdf>), *Angewandte Informatik* **2**: 71–77, .
- [8] Gaede, Volker; Guenther, Oliver (1998), "Multidimensional access methods" (<http://www-static.cc.gatech.edu/computing/Database/readinggroup/articles/p170-gaede.pdf>), *ACM Computing Surveys* **30** (2): 170–231, doi:10.1145/280277.280279, .
- [9] Ramsak, Frank; Markl, Volker; Fenk, Robert; Zirkel, Martin; Elhardt, Klaus; Bayer, Rudolf (2000), "Integrating the UB-tree into a Database System Kernel" (<http://www.mistral.in.tum.de/results/publications/RMF+00.pdf>), *Int. Conf. on Very Large Databases (VLDB)*, pp. 263–272, .
- [10] US 7321890 (<http://worldwide.espacenet.com/textdoc?DB=EPODOC&IDX=US7321890>), Tropf, H., "Database system and method for organizing data elements according to a Hilbert curve", issued January 22, 2008.
- [11] Samet, H. (2006), *Foundations of Multidimensional and Metric Data Structures*, San Francisco: Morgan-Kaufmann.
- [12] Vinod Valsalam, Anthony Skjellum: A framework for high-performance matrix multiplication based on hierarchical abstractions, algorithms and optimized low-level kernels. *Concurrency and Computation: Practice and Experience* 14(10): 805-839 (2002)

## External links

- STANN: A library for approximate nearest neighbor search, using Z-order curve (<http://www.compgeom.com/~stann>)
- Methods for programming bit interleaving (<http://graphics.stanford.edu/~seander/bithacks.html#InterleaveTableObvious>), Sean Eron Anderson, Stanford University

# UB-tree

---

The **UB-tree** as proposed by Rudolf Bayer and Volker Markl is a balanced tree for storing and efficiently retrieving multidimensional data. It is basically a B+ tree (information only in the leaves) with records stored according to Z-order, also called Morton order. Z-order is simply calculated by bitwise interlacing the keys.

Insertion, deletion, and point query are done as with ordinary B+ trees. To perform range searches in multidimensional point data, however, an algorithm must be provided for calculating, from a point encountered in the data base, the next Z-value which is in the multidimensional search range.

The original algorithm to solve this key problem was exponential with the dimensionality and thus not feasible<sup>[1]</sup> ("GetNextZ-address"). A solution to this "crucial part of the UB-tree range query" linear with the z-address bit length has been described later.<sup>[2]</sup> This method has already been described in an older paper<sup>[3]</sup> where using Z-order with search trees has first been proposed.

## References

- [1] Markl, V. (1999). *MISTRAL: Processing Relational Queries using a Multidimensional Access Technique* (<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.32.6487>). .
- [2] Ramsak, Frank; Markl, Volker; Fenk, Robert; Zirkel, Martin; Elhardt, Klaus; Bayer, Rudolf (September 10–14, 2000). "Integrating the UB-tree into a Database System Kernel" (<http://www.vldb.org/dblp/db/conf/vldb/RamsakMFZEB00.html>). 26th International Conference on Very Large Data Bases (<http://www.vldb.org/dblp/db/conf/vldb/vldb2000.html>). pp. 263–272. .
- [3] Tropf, H.; Herzog, H.. "Multidimensional Range Search in Dynamically Balanced Trees" (<http://www.vision-tools.com/h-tropf/multidimensionalrangequery.pdf>) (PDF). *Angewandte Informatik (Applied Informatics)* (2/1981): 71–77. ISSN 0013-5704. .

## External links

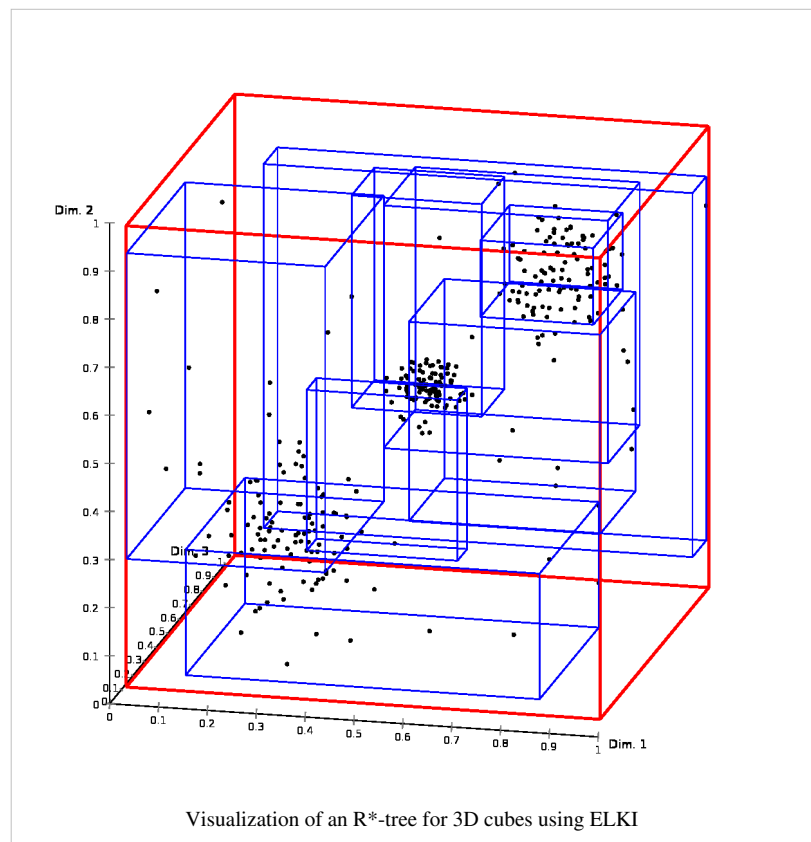
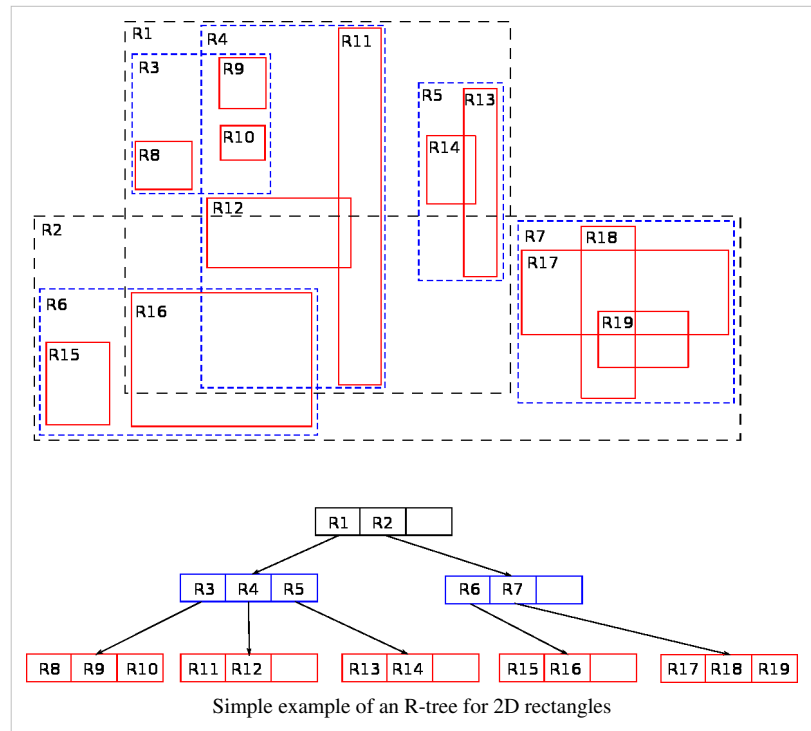
- Mistral (<http://mistral.in.tum.de/>)

# R-tree

**R-trees** are tree data structures used for spatial access methods, i.e., for indexing multi-dimensional information such as geographical coordinates, rectangles or polygons. The R-tree was proposed by Antonin Guttman in 1984<sup>[1]</sup> and has found significant use in both research and real-world applications.<sup>[2]</sup> A common real-world usage for an R-tree might be to store spatial objects such as restaurant locations or the polygons that typical maps are made of: streets, buildings, outlines of lakes, coastlines, etc. and then find answers quickly to queries such as "Find all museums within 2 km of my current location", "retrieve all road segments within 2 km of my location" (to display them in a navigation system) or "find the nearest gas station" (although not taking roads into account).

## R-tree idea

The key idea of the data structure is to group nearby objects and represent them with their minimum bounding rectangle in the next higher level of the tree; the "R" in R-tree is for rectangle. Since all objects lie within this bounding rectangle, a query that does not intersect the bounding rectangle also cannot intersect any of the contained objects. At the leaf level, each rectangle describes a single object; at higher levels the aggregation of an increasing number of objects. This can also be seen as an increasingly coarse approximation of the data set.



Similar to the B-tree, the R-tree is also a balanced search tree (so all leaf nodes are at the same height), organizes the data in pages, and is designed for storage on disk (as used in databases). Each page can contain a maximum number

of entries, often denoted as  $M$ . It also guarantees a minimum fill (except for the root node), however best performance has been experienced with a minimum fill of 30%–40% of the maximum number of entries (B-trees guarantee 50% page fill, and B\*-trees even 66%). The reason for this is the more complex balancing required for spatial data as opposed to linear data stored in B-trees.

As with most trees, the searching algorithms (e.g., intersection, containment, nearest neighbor search) are rather simple. The key idea is to use the bounding boxes to decide whether or not to search inside a subtree. In this way, most of the nodes in the tree are never read during a search. Like B-trees, this makes R-trees suitable for large data sets and databases, where nodes can be paged to memory when needed, and the whole tree cannot be kept in main memory.

The key difficulty of R-trees is to build an efficient tree that on one hand is balanced (so the leaf nodes are at the same height) on the other hand the rectangles do not cover too much empty space and do not overlap too much (so that during search, fewer subtrees need to be processed). For example, the original idea for inserting elements to obtain an efficient tree is to always insert into the subtree that requires least enlargement of its bounding box. Once that page is full, the data is split into two sets that should cover the minimal area each. Most of the research and improvements for R-trees aims at improving the way the tree is built and can be grouped into two objectives: building an efficient tree from scratch (known as bulk-loading) and performing changes on an existing tree (insertion and deletion).

R-trees do not historically guarantee good worst-case performance, but generally perform well with real-world data.<sup>[3]</sup> While more of theoretical interest, the (bulk-loaded) Priority R-Tree variant of the R-tree is also worst-case optimal,<sup>[4]</sup> but due to the increased complexity, has not received much attention in practical applications so far.

When data is organized in an R-Tree, the  $k$  nearest neighbors (for any  $L^p$ -Norm) of all points can efficiently be computed using a spatial join.<sup>[5]</sup> This is beneficial for many algorithms based on the  $k$  nearest neighbors, for example the Local Outlier Factor. DeLi-Clu,<sup>[6]</sup> Density-Link-Clustering is a cluster analysis algorithm that uses the R-tree structure for a similar kind of spatial join to efficiently compute an OPTICS clustering.

## Variants

- R\* tree
- R+ tree
- Hilbert R-tree
- X-tree

## Algorithm

### Data layout

Data in R-trees is organized in pages, that can have a variable number of entries (up to some pre-defined maximum, and usually above a minimum fill). Each entry within a non-leaf node stores two pieces of data: a way of identifying a child node, and the bounding box of all entries within this child node. Leaf nodes store the data required for each child, often a point or bounding box representing the child and an external identifier for the child. For point data, the leaf entries can be just the points themselves. For polygon data (that often requires the storage of large polygons) the common setup is to store only the MBR (minimum bounding rectangle) of the polygon along with a unique identifier in the tree.

## Search

The input is a search rectangle (Query box). Searching is quite similar to searching in a B+tree. The search starts from the root node of the tree. Every internal node contains a set of rectangles and pointers to the corresponding child node and every leaf node contains the rectangles of spatial objects (the pointer to some spatial object can be there). For every rectangle in a node, it has to be decided if it overlaps the search rectangle or not. If yes, the corresponding child node has to be searched also. Searching is done like this in a recursive manner until all overlapping nodes have been traversed. When a leaf node is reached, the contained bounding boxes (rectangles) are tested against the search rectangle and their objects (if there are any) are put into the result set if they lie within the search rectangle.

## Insertion

To insert an object, the tree is traversed recursively from the root node. At each step, all rectangles in the current directory node are examined, and a candidate is chosen using a heuristic such as choosing the rectangle which requires least enlargement. The search then descends into this page, until reaching a leaf node. If the leaf node is full, it must be split before the insertion is made. Again, since an exhaustive search is too expensive, an heuristic is employed to split the node into two. Adding the newly created node to the previous level, this level can again overflow, and these overflows can propagate up to the root node; when this node also overflows, a new root node is created and the tree has increased in height.

### Choosing the insertion subtree

At each level, the algorithm needs to decide in which subtree to insert the new data object. When a data object is fully contained in a single rectangle, the choice is clear. When there are multiple options or rectangles in need of enlargement, the choice can have a significant impact on the performance of the said tree.

In the classic R-tree, objects are inserted into the subtree that needs the least enlargement. In the more advanced R\*-tree, a mixed heuristic is employed. At leaf level, it tries to minimize the overlap (in case of ties, prefer least enlargement and then least area); at the higher levels, it behaves similar to the R-tree, but on ties again preferring the subtree with smaller area. The decreased overlap of rectangles in the R\*-tree is one of the key benefits over the traditional R-Tree (this is also a consequence of the other heuristics used, not only the subtree choosing).

### Splitting an overflowing node

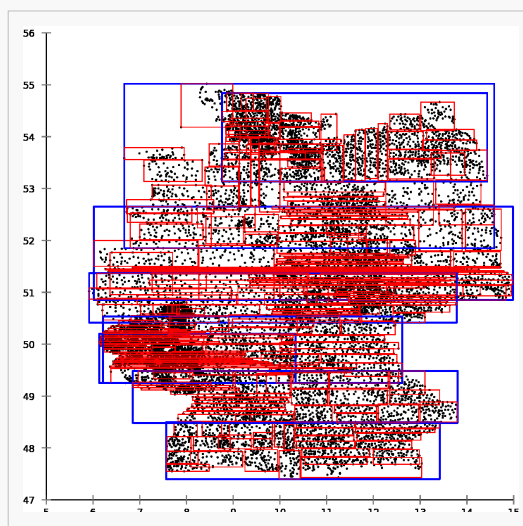
Since redistributing all objects of a node into two nodes has an exponential number of options, an heuristic needs to be employed to find the best split. In the classic R-Tree, Guttman proposed two such heuristics, called QuadraticSplit and LinearSplit. In quadratic split, the algorithm searches the pair of rectangles that is the worst combination to have in the same node, and puts them as initial objects into the two new groups. It then searches the entry which has the strongest preference for one of the groups (in terms of area increase) and assigns the object to this group until all objects are assigned (satisfying the minimum fill).

There are other splitting strategies such as Greene's Split,<sup>[7]</sup> the R\*-tree splitting heuristic<sup>[8]</sup> (which again tries to minimize overlap, but also prefers quadratic pages) or the linear split algorithm proposed by Ang and Tan<sup>[9]</sup> (which however can produce very unregular rectangles, which are less performant for many real world range and window queries). In addition to having a more advanced splitting heuristic, the R\*-tree also tries to avoid splitting a node by reinserting some of the node members, which is similar to the way a B-tree balances overflowing nodes. This was shown to also reduce overlap and thus tree performance.

Finally, the X-tree<sup>[10]</sup> can be seen as a R\*-tree variant that can also decide to not split a node, but construct a so-called super-node containing all the extra entries, when it doesn't find a good split (in particular for high-dimensional data).

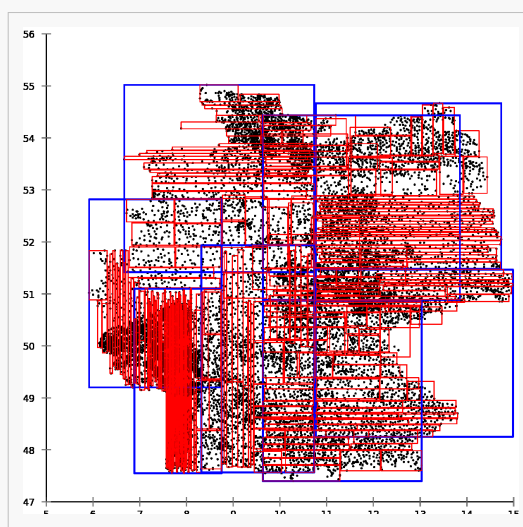
Effect of different splitting heuristics on a database with Germany postal districts





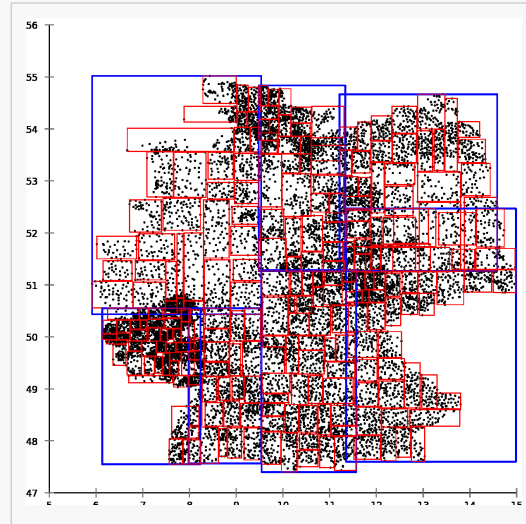
Guttman quadratic split.<sup>[1]</sup>

There are many pages that extend from east to west all over Germany, and pages overlap a lot. This is not beneficial for most applications, that often only need a small rectangular area that intersects with many slices.



Ang-Tan linear split.<sup>[9]</sup>

While the slices do not extend as far as with Guttman, the slicing problem affects almost every leaf page. Leaf pages overlap little, but directory pages do.



R\* tree topological split.<sup>[8]</sup>

The pages overlap very little since the R\*-tree tries to minimize page overlap, and the reinsertions further optimized the tree. The split strategy also does not prefer slices, the resulting pages are much more useful for common map applications.

## Deletion

Deleting an entry from a page may require updating the bounding rectangles of parent pages. However, when a page is underfull, it will not be balanced with its neighbors. Instead, the page will be dissolved and all its children (which may be subtrees, not only leaf objects) will be reinserted. If during this process the root node has a single element, the tree height can decrease.

## Bulk-loading

- Nearest-X - Objects are sorted by their first coordinate ("X") and then split into pages of the desired size.
- Packed Hilbert R-Tree - variation of Nearest-X, but sorting using the Hilbert value of the center of a rectangle instead of using the X coordinate. There is no guarantee the pages will not overlap.
- Sort-Tile-Recursive (STR)<sup>[11]</sup>: Another variation of Nearest-X, that estimates the total number of leaves required as  $l = \lceil \text{number of objects} / \text{capacity} \rceil$ , the required split factor in each dimension to achieve this as  $s = \lceil l^{1/d} \rceil$ , then repeatedly splits each dimensions successively into  $s$  equal sized partitions using 1-dimensional sorting. The resulting pages, if they occupy more than one page, are again bulk-loaded using the same algorithm. For point data, the leaf nodes will not overlap, and "tile" the data space into approximately equal sized pages.
- Priority R-Tree

## References

- [1] doi: 10.1145/602259.602266  
This citation will be automatically completed in the next few minutes. You can jump the queue or expand by hand ([http://en.wikipedia.org/wiki/Template:cite\\_doi/\\_10.1145.2f602259.602266\\_?preload=Template:Cite\\_doi/preload&editintro=Template:Cite\\_doi/editintro&action=edit](http://en.wikipedia.org/wiki/Template:cite_doi/_10.1145.2f602259.602266_?preload=Template:Cite_doi/preload&editintro=Template:Cite_doi/editintro&action=edit))
- [2] Y. Manolopoulos; A. Nanopoulos; Y. Theodoridis (2006). *R-Trees: Theory and Applications* (<http://books.google.com/books?id=1mu099DN9UwC&pg=PR5>). Springer. ISBN 978-1-85233-977-7. . Retrieved 8 October 2011.
- [3] doi: 10.1007/978-3-540-45072-6\_2  
This citation will be automatically completed in the next few minutes. You can jump the queue or expand by hand ([http://en.wikipedia.org/wiki/Template:cite\\_doi/\\_10.1007.2f978-3-540-45072-6\\_2?preload=Template:Cite\\_doi/preload&editintro=Template:Cite\\_doi/editintro&action=edit](http://en.wikipedia.org/wiki/Template:cite_doi/_10.1007.2f978-3-540-45072-6_2?preload=Template:Cite_doi/preload&editintro=Template:Cite_doi/editintro&action=edit))
- [4] doi: 10.1145/1007568.1007608  
This citation will be automatically completed in the next few minutes. You can jump the queue or expand by hand ([http://en.wikipedia.org/wiki/Template:cite\\_doi/\\_10.1145.2f1007568.1007608\\_?preload=Template:Cite\\_doi/preload&editintro=Template:Cite\\_doi/editintro&action=edit](http://en.wikipedia.org/wiki/Template:cite_doi/_10.1145.2f1007568.1007608_?preload=Template:Cite_doi/preload&editintro=Template:Cite_doi/editintro&action=edit))
- [5] doi: 10.1145/170036.170075  
This citation will be automatically completed in the next few minutes. You can jump the queue or expand by hand ([http://en.wikipedia.org/wiki/Template:cite\\_doi/\\_10.1145.2f170036.170075?preload=Template:Cite\\_doi/preload&editintro=Template:Cite\\_doi/editintro&action=edit](http://en.wikipedia.org/wiki/Template:cite_doi/_10.1145.2f170036.170075?preload=Template:Cite_doi/preload&editintro=Template:Cite_doi/editintro&action=edit))
- [6] doi: 10.1007/11731139\_16  
This citation will be automatically completed in the next few minutes. You can jump the queue or expand by hand ([http://en.wikipedia.org/wiki/Template:cite\\_doi/\\_10.1007.2f11731139\\_16?preload=Template:Cite\\_doi/preload&editintro=Template:Cite\\_doi/editintro&action=edit](http://en.wikipedia.org/wiki/Template:cite_doi/_10.1007.2f11731139_16?preload=Template:Cite_doi/preload&editintro=Template:Cite_doi/editintro&action=edit))
- [7] doi: 10.1109/ICDE.1989.47268  
This citation will be automatically completed in the next few minutes. You can jump the queue or expand by hand ([http://en.wikipedia.org/wiki/Template:cite\\_doi/\\_10.1109.2f1109.47268\\_?preload=Template:Cite\\_doi/preload&editintro=Template:Cite\\_doi/editintro&action=edit](http://en.wikipedia.org/wiki/Template:cite_doi/_10.1109.2f1109.47268_?preload=Template:Cite_doi/preload&editintro=Template:Cite_doi/editintro&action=edit))
- [8] doi: 10.1145/93597.98741  
This citation will be automatically completed in the next few minutes. You can jump the queue or expand by hand ([http://en.wikipedia.org/wiki/Template:cite\\_doi/\\_10.1145.2f93597.98741?preload=Template:Cite\\_doi/preload&editintro=Template:Cite\\_doi/editintro&action=edit](http://en.wikipedia.org/wiki/Template:cite_doi/_10.1145.2f93597.98741?preload=Template:Cite_doi/preload&editintro=Template:Cite_doi/editintro&action=edit))
- [9] doi: 10.1007/3-540-63238-7\_38  
This citation will be automatically completed in the next few minutes. You can jump the queue or expand by hand ([http://en.wikipedia.org/wiki/Template:cite\\_doi/\\_10.1007.2f3-540-63238-7\\_38?preload=Template:Cite\\_doi/preload&editintro=Template:Cite\\_doi/editintro&action=edit](http://en.wikipedia.org/wiki/Template:cite_doi/_10.1007.2f3-540-63238-7_38?preload=Template:Cite_doi/preload&editintro=Template:Cite_doi/editintro&action=edit))
- [10] Berchtold, Stefan; Keim, Daniel A.; Kriegel, Hans-Peter (1996). "The X-Tree: An Index Structure for High-Dimensional Data" (<http://www.dbs.ifi.lmu.de/Publicationen/Papers/x-tree.ps>). *Proceedings of the 22nd VLDB Conference* (Mumbai, India): 28–39. .
- [11] Leutenegger, Scott T.; Edgington, Jeffrey M.; Lopez, Mario A. (February 1997). *STR: A Simple and Efficient Algorithm for R-Tree Packing* ([http://archive.org/details/nasa\\_techdoc\\_19970016975](http://archive.org/details/nasa_techdoc_19970016975)). .

## External links

- R-tree portal (<http://www.rtreeportal.org/>)
- R-Tree implementations: C & C++ (<http://superliminal.com/sources/sources.htm#C&C++Code>), Java applet (<http://gis.umb.no/gis/applets/rtree2/jdk1.1/>), Common Lisp (<http://www.cliki.net/spatial-trees>), Python (<http://pypi.python.org/pypi/Rtree/>), Javascript (<http://github.com/imbcmth/RTree>).

# R+ tree

---

An **R+ tree** is a method for looking up data using a location, often (x, y) coordinates, and often for locations on the surface of the earth. Searching on one number is a solved problem; searching on two or more, and asking for locations that are nearby in both x and y directions, requires craftier algorithms.

Fundamentally, an R+ tree is a tree data structure, a variant of the R tree, used for indexing spatial information.

## Difference between R+ trees and R trees

R+ trees are a compromise between R-trees; and kd-trees; they avoid overlapping of internal nodes by inserting an object into multiple leaves if necessary. **Coverage** is the entire area to cover all related rectangles. **Overlap** is the entire area which is contained in two or more nodes.<sup>[1]</sup> Minimal coverage reduces the amount of "dead space" (empty area) which is covered by the nodes of the R-tree. Minimal overlap reduces the set of search paths to the leaves (even more critical for the access time than minimal coverage). Efficient search requires minimal coverage and overlap.

R+ trees differ from R trees in that:

- Nodes are not guaranteed to be at least half filled
- The entries of any internal node do not overlap
- An object ID may be stored in more than one leaf node

## Advantages

- Because nodes are not overlapped with each other, point query performance benefits since all spatial regions are covered by at most one node.
- A single path is followed and fewer nodes are visited than with the R-tree

## Disadvantages

- Since rectangles are duplicated, an R+ tree can be larger than an R tree built on same data set.
- Construction and maintenance of R+ trees is more complex than the construction and maintenance of R trees and other variants of the R tree.

## References

- [1] Härder, Rahm, Theo, Erhard (2007). *Datenbanksysteme*. (2., überarb. Aufl. ed.). Berlin [etc.]: Gardners Books. pp. 285, 286. ISBN 3-540-42133-5.
  - T. Sellis, N. Roussopoulos, and C. Faloutsos. The R+-Tree: A dynamic index for multi-dimensional objects (<http://citeseer.ist.psu.edu/sellis87rtree.html>). In VLDB, 1987.
-

# R\* tree

**R\*-trees** are a variant of R-trees used for indexing spatial information. R\*-trees support point and spatial data at the same time with a slightly higher cost than other R-trees. It was proposed by Norbert Beckmann, Hans-Peter Kriegel, Ralf Schneider, and Bernhard Seeger in 1990.<sup>[1]</sup>

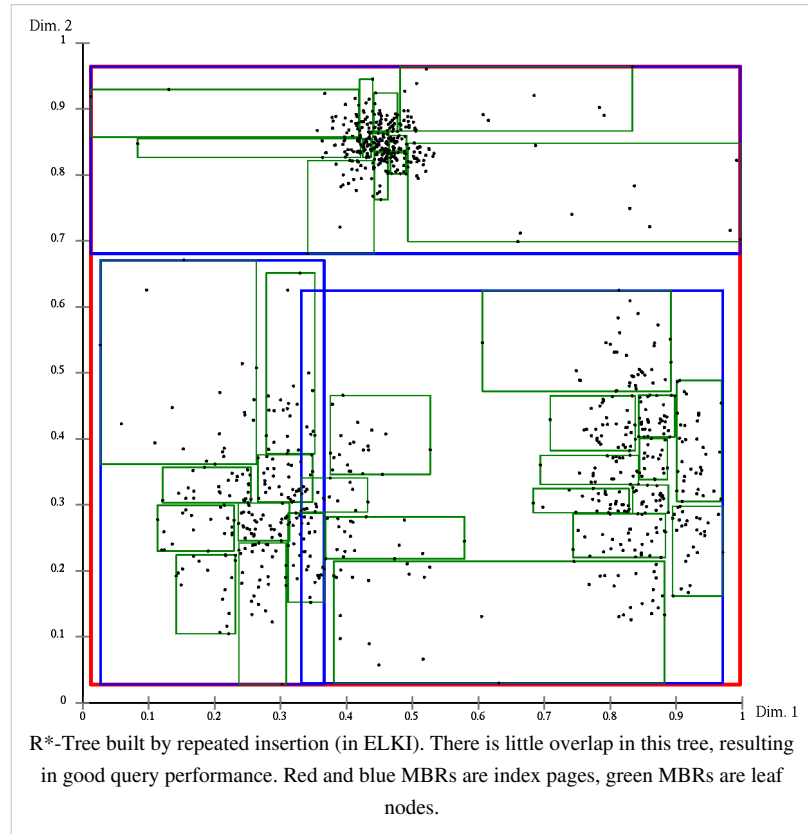
## Difference between R\*-trees and R-trees

Minimization of both coverage and overlap is crucial to the performance of R-trees. Overlap means that, on data query or insertion, more than one branch of the tree needs to be expanded (due to the way data is being split in regions which may overlap). A minimized coverage improves pruning performance, allowing to exclude whole pages from search more often, in particular for negative range queries.

The R\*-tree attempts to reduce both, using a combination of a revised node split algorithm and the concept of forced reinsertion at node overflow. This is based on the observation that R-tree structures are highly susceptible to the order in which their entries are inserted, so an insertion-built (rather than bulk-loaded) structure is likely to be sub-optimal. Deletion and reinsertion of entries allows them to

"find" a place in the tree that may be more appropriate than their original location.

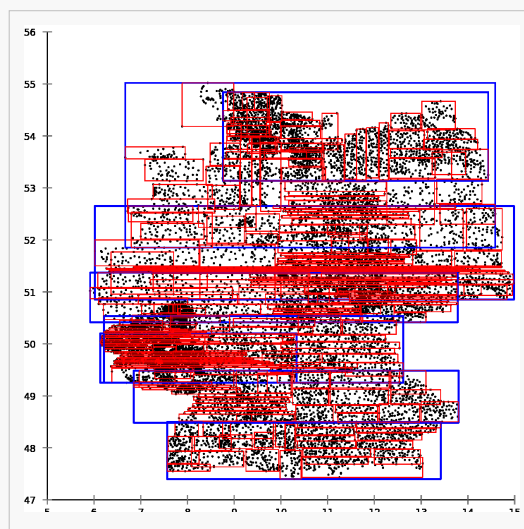
When a node overflows, a portion of its entries are removed from the node and reinserted into the tree. (In order to avoid an indefinite cascade of reinsertions caused by subsequent node overflow, the reinsertion routine may be called only once in each level of the tree when inserting any one new entry.) This has the effect of producing more well-clustered groups of entries in nodes, reducing node coverage. Furthermore, actual node splits are often postponed, causing average node occupancy to rise. Re-insertion can be seen as a method of incremental tree optimization triggered on node overflow.



## Performance

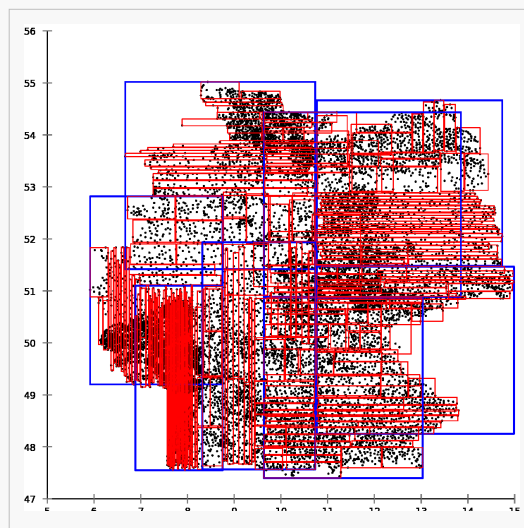
- Improved split heuristic produces pages that are more rectangular and thus better for many applications.
- Reinsertion method optimizes the existing tree, but increases complexity.
- Efficiently supports point and spatial data at the same time.

Effect of different splitting heuristics on a database with Germany postal districts



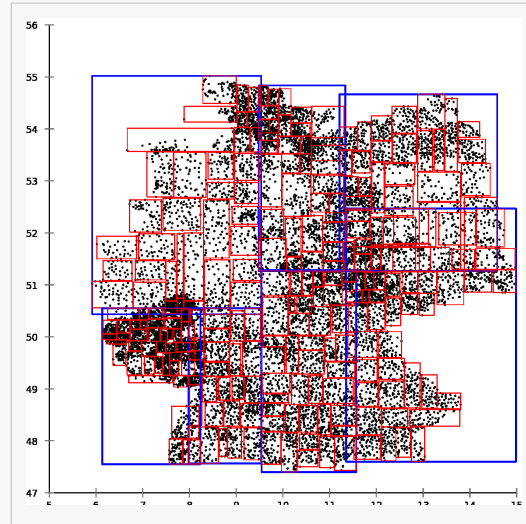
R-Tree with Guttman quadratic split.<sup>[1]</sup>

There are many pages that extend from east to west all over Germany, and pages overlap a lot. This is not beneficial for most applications, that often only need a small rectangular area that intersects with many slices.



R-Tree with Ang-Tan linear split.<sup>[9]</sup>

While the slices do not extend as far as with Guttman, the slicing problem affects almost every leaf page. Leaf pages overlap little, but directory pages do.



### R\* tree topological split.<sup>[1]</sup>

The pages overlap very little since the R\*-tree tries to minimize page overlap, and the reinsertions further optimized the tree. The split strategy also does not prefer slices, the resulting pages are much more useful for common map applications.

## Algorithm

The R\*-tree uses the same algorithm as the R-tree for query and delete operations. The primary difference is the insert algorithm, specifically how it chooses which branch to insert the new node into and the methodology for splitting a node that is full.

## References

[1] doi: 10.1145.2F93597.98741

This citation will be automatically completed in the next few minutes. You can jump the queue or expand by hand ([http://en.wikipedia.org/wiki/Template:cite\\_doi/\\_10.1145.2f93597.98741?preload=Template:Cite\\_doi/preload&editintro=Template:Cite\\_doi/editintro&action=edit](http://en.wikipedia.org/wiki/Template:cite_doi/_10.1145.2f93597.98741?preload=Template:Cite_doi/preload&editintro=Template:Cite_doi/editintro&action=edit))

## External links

- R-tree Demo (<http://donar.umiacs.umd.edu/quadtrees/points/rtrees.html>)
- The TPIE Library contains a C++ R\* tree implementation (<http://www.cs.duke.edu/TPIE/>)
- A header-only C++ R\* Tree Implementation (<http://www.virtualroadside.com/blog/index.php/2008/10/04/r-tree-implementation-for-cpp/>)
- Java and C++ implementation are in the Spatial Index Library (<http://www2.research.att.com/~marioh/spatialindex/>)

# Hilbert R-tree

**Hilbert R-tree**, an R-tree variant, is an index for multidimensional objects like lines, regions, 3-D objects, or high dimensional feature-based parametric objects. It can be thought of as an extension to B+-tree for multidimensional objects.

The performance of R-trees depends on the quality of the algorithm that clusters the data rectangles on a node. Hilbert R-trees use space-filling curves, and specifically the Hilbert curve, to impose a linear ordering on the data rectangles.

There are two types of Hilbert R-tree, one for static databases and one for dynamic databases. In both cases, space filling curves and specifically the Hilbert curve are used to achieve better ordering of multidimensional objects in the node. This ordering has to be 'good', in the sense that it should group 'similar' data rectangles together, to minimize the area and perimeter of the resulting minimum bounding rectangles (MBRs). Packed Hilbert R-trees are suitable for static databases in which updates are very rare or in which there are no updates at all.

The dynamic Hilbert R-tree is suitable for dynamic databases where insertions, deletions, or updates may occur in real time. Moreover, dynamic Hilbert R-trees employ flexible deferred splitting mechanism to increase the space utilization. Every node has a well defined set of sibling nodes. By adjusting the split policy the Hilbert R-tree can achieve a degree of space utilization as high as is desired. This is done by proposing an ordering on the R-tree nodes. The Hilbert R-tree sorts rectangles according to the Hilbert value of the center of the rectangles (i.e., MBR). (The Hilbert value of a point is the length of the Hilbert curve from the origin to the point.) Given the ordering, every node has a well-defined set of sibling nodes; thus, deferred splitting can be used. By adjusting the split policy, the Hilbert R-tree can achieve as high utilization as desired. To the contrary, other R-tree variants have no control over the space utilization.

## The basic idea

Although the following example is for a static environment, it explains the intuitive principles for good R-tree design. These principles are valid for both static and dynamic databases. Roussopoulos and Leifker proposed a method for building a packed R-tree that achieves almost 100% space utilization. The idea is to sort the data on the x or y coordinate of one of the corners of the rectangles. Sorting on any of the four coordinates gives similar results. In this discussion points or rectangles are sorted on the x coordinate of the lower left corner of the rectangle. In the discussion below the Roussopoulos and Leifker's method is referred to as the lowx packed R-tree. The sorted list of rectangles is scanned; successive rectangles are assigned to the same R-tree leaf node until that node is full; a new leaf node is then created and the scanning of the sorted list continues. Thus, the nodes of the resulting R-tree will be fully packed, with the possible exception of the last node at each level. Thus, the utilization is  $\approx 100\%$ . Higher levels of the tree are created in a similar way.

Figure 1 highlights the problem of the lowx packed R-tree. Figure 1[Right] shows the leaf nodes of the R-tree that the lowx packing method will create for the points of Figure 1 [Left]. The fact that the resulting father nodes cover little area explains why the lowx packed R-tree achieves excellent performance for point queries. However, the fact that the fathers have large perimeters, explains the degradation of performance for region queries. This is consistent with the analytical formulas for R-tree performance.<sup>[1]</sup> Intuitively, the packing algorithm should ideally assign nearby points to the same leaf node. Ignorance of the y coordinate by the lowx packed R-tree tends to violate this empirical rule.



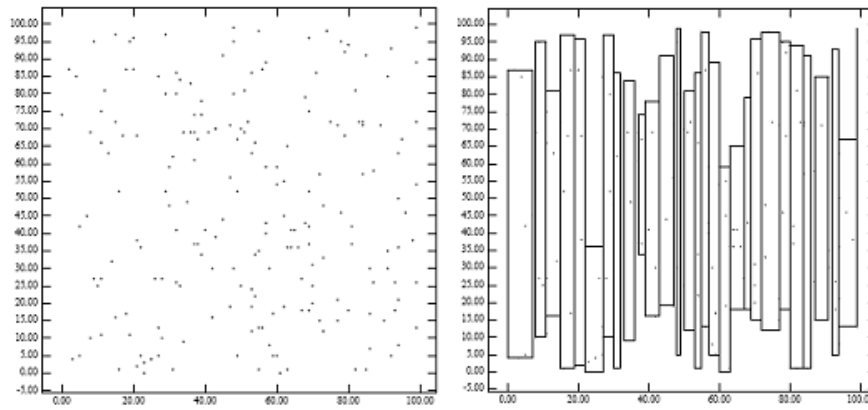


Figure 1: [Left] 200 points uniformly distributed; [Right] MBR of nodes generated by the 'lowx packed R-tree' algorithm

This section describes two variants of the Hilbert R-trees. The first index is suitable for the static database in which updates are very rare or in which there are no updates at all. The nodes of the resulting R-tree will be fully packed, with the possible exception of the last node at each level. Thus, the space utilization is  $\approx 100\%$ ; this structure is called a packed Hilbert R-tree. The second index, called a Dynamic Hilbert R-tree, supports insertions and deletions, and is suitable for a dynamic environment.

## Packed Hilbert R-trees

The following provides a brief introduction to the Hilbert curve. The basic Hilbert curve on a  $2 \times 2$  grid, denoted by  $H_1$  is shown in Figure 2. To derive a curve of order  $i$ , each vertex of the basic curve is replaced by the curve of order  $i - 1$ , which may be appropriately rotated and/or reflected. Figure 2 also shows the Hilbert curves of order two and three. When the order of the curve tends to infinity, like other space filling curves, the resulting curve is a fractal, with a fractal dimension of two.<sup>[1][2]</sup> The Hilbert curve can be generalized for higher dimensionalities. Algorithms for drawing the two-dimensional curve of a given order can be found in <sup>[3]</sup> and <sup>[2]</sup>. An algorithm for higher dimensionalities is given in <sup>[4]</sup>.

The path of a space filling curve imposes a linear ordering on the grid points; this path may be calculated by starting at one end of the curve and following the path to the other end. The actual coordinate values of each point can be calculated. However, for the Hilbert curve this is much harder than for example the Z-order curve. Figure 2 shows one such ordering for a  $4 \times 4$  grid (see curve  $H_2$ ). For example, the point  $(0,0)$  on the  $H_2$  curve has a Hilbert value of 0, while the point  $(1,1)$  has a Hilbert value of 2.

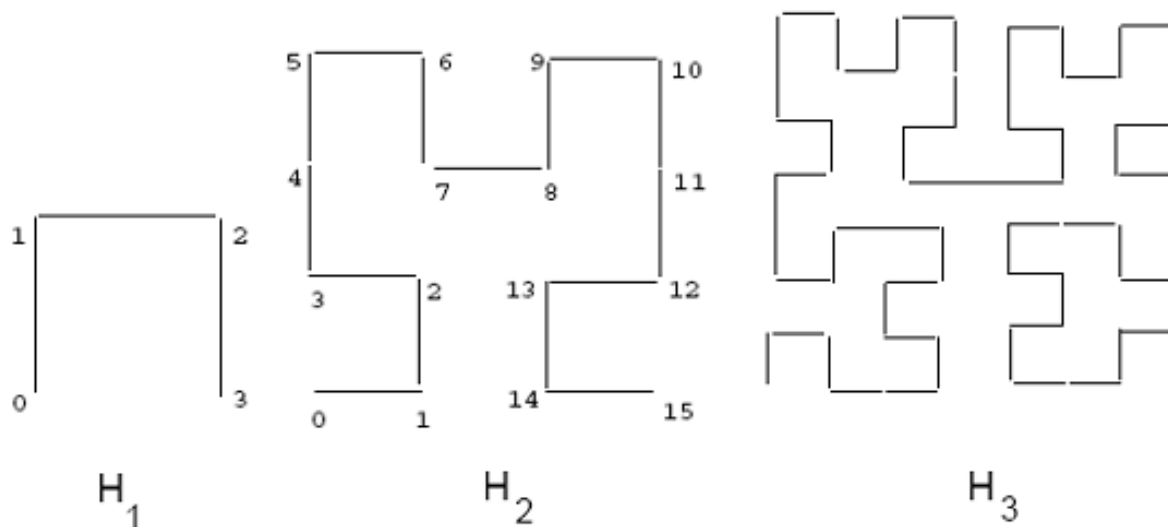


Figure 2: Hilbert curves of order 1, 2, and 3

The Hilbert curve imposes a linear ordering on the data rectangles and then traverses the sorted list, assigning each set of  $C$  rectangles to a node in the R-tree. The final result is that the set of data rectangles on the same node will be close to each other in the linear ordering, and most likely in the native space; thus, the resulting R-tree nodes will have smaller areas. Figure 2 illustrates the intuitive reasons why our Hilbert-based methods will result in good performance. The data is composed of points (the same points as given in Figure 1). By grouping the points according to their Hilbert values, the MBRs of the resulting R-tree nodes tend to be small square-like rectangles. This indicates that the nodes will likely have small area and small perimeters. Small area values result in good performance for point queries; small area and small perimeter values lead to good performance for larger queries.

### Algorithm Hilbert-Pack

(packs rectangles into an R-tree)

Step 1. Calculate the Hilbert value for each data rectangle

Step 2. Sort data rectangles on ascending Hilbert values

Step 3. /\* Create leaf nodes (level  $l=0$ ) \*/

- While (there are more rectangles)
  - generate a new R-tree node
  - assign the next  $C$  rectangles to this node

Step 4. /\* Create nodes at higher level ( $l + 1$ ) \*/

- While (there are  $> 1$  nodes at level  $l$ )
  - sort nodes at level  $l \geq 0$  on ascending creation time
  - repeat Step 3

The assumption here is that the data are static or the frequency of modification is low. This is a simple heuristic for constructing an R-tree with 100% space utilization which at the same time will have as good response time as possible.

### Dynamic Hilbert R-trees

The performance of R-trees depends on the quality of the algorithm that clusters the data rectangles on a node. Hilbert R-trees use space-filling curves, and specifically the Hilbert curve, to impose a linear ordering on the data rectangles. The Hilbert value of a rectangle is defined as the Hilbert value of its center.

#### Tree structure

The Hilbert R-tree has the following structure. A leaf node contains at most  $C_l$  entries each of the form  $(R, \text{obj\_id})$  where  $C_l$  is the capacity of the leaf,  $R$  is the MBR of the real object  $(x_{\text{low}}, x_{\text{high}}, y_{\text{low}}, y_{\text{high}})$  and  $\text{obj\_id}$  is a pointer to the object description record. The main difference between the Hilbert R-tree and the R\*-tree<sup>[5]</sup> is that non-leaf nodes also contain information about the LHVs (Largest Hilbert Value). Thus, a non-leaf node in the Hilbert R-tree contains at most  $C_n$  entries of the form  $(R, \text{ptr}, \text{LHV})$  where  $C_n$  is the capacity of a non-leaf node,  $R$  is the MBR that encloses all the children of that node,  $\text{ptr}$  is a pointer to the child node, and LHV is the largest Hilbert value among the data rectangles enclosed by  $R$ . Notice that since the non-leaf node picks one of the Hilbert values of the children to be the value of its own LHV, there is not extra cost for calculating the Hilbert values of the MBR of non-leaf nodes. Figure 3 illustrates some rectangles organized in a Hilbert R-tree. The Hilbert values of the centers are the numbers near the 'x' symbols (shown only for the parent node 'II'). The LHV's are in [brackets]. Figure 4 shows how the tree of Figure 3 is stored on the disk; the contents of the parent node 'II' are shown in more detail. Every data rectangle in node 'I' has a Hilbert value  $v \leq 33$ ; similarly every rectangle in node 'II' has a Hilbert value greater than 33 and  $\leq 107$ , etc.

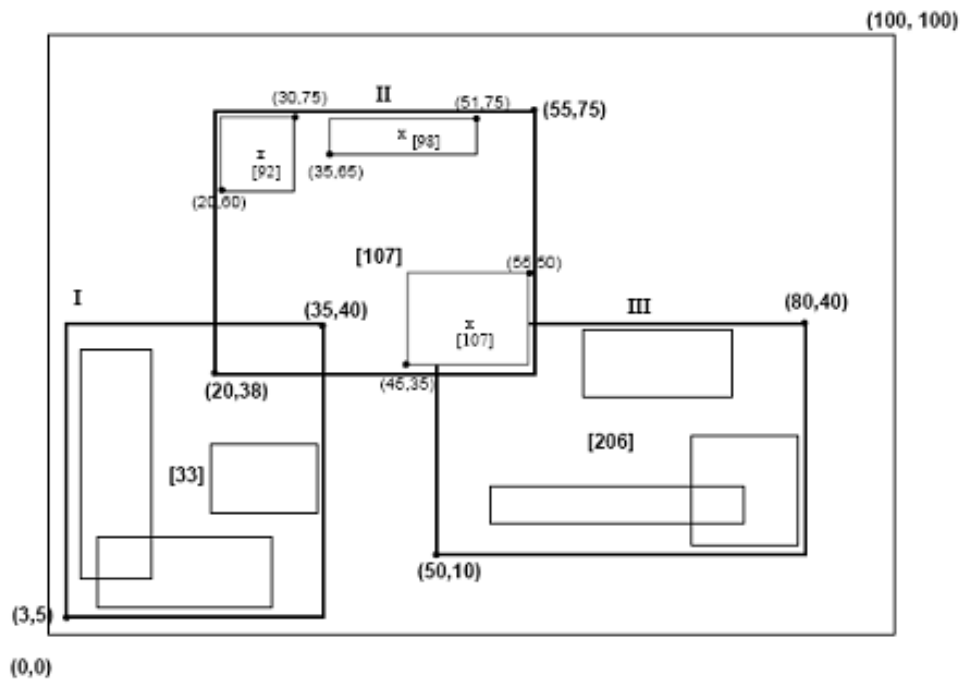


Figure 3: Data rectangles organized in a Hilbert R-tree (Hilbert values and LHV's are in Brackets)

A plain R-tree splits a node on overflow, creating two nodes from the original one. This policy is called a 1-to-2 splitting policy. It is possible also to defer the split, waiting until two nodes split into three. Note that this is similar to the B\*-tree split policy. This method is referred to as the 2-to-3 splitting policy. In general, this can be extended to s-to-(s+1) splitting policy; where s is the order of the splitting policy. To implement the order-s splitting policy, the overflowing node tries to push some of its entries to one of its s - 1 siblings; if all of them are full, then s-to-(s+1) split need to be done. The s - 1 siblings are called the cooperating siblings. Next, the algorithms for searching, insertion, and overflow handling are described in details.

## Searching

The searching algorithm is similar to the one used in other R-tree variants. Starting from the root, it descends the tree and examines all nodes that intersect the query rectangle. At the leaf level, it reports all entries that intersect the query window w as qualified data items.

### Algorithm Search(node Root, rect w):

S1. Search nonleaf nodes:

Invoke Search for every entry whose MBR intersects the query window w.

S2. Search leaf nodes:

Report all entries that intersect the query window w as candidates.

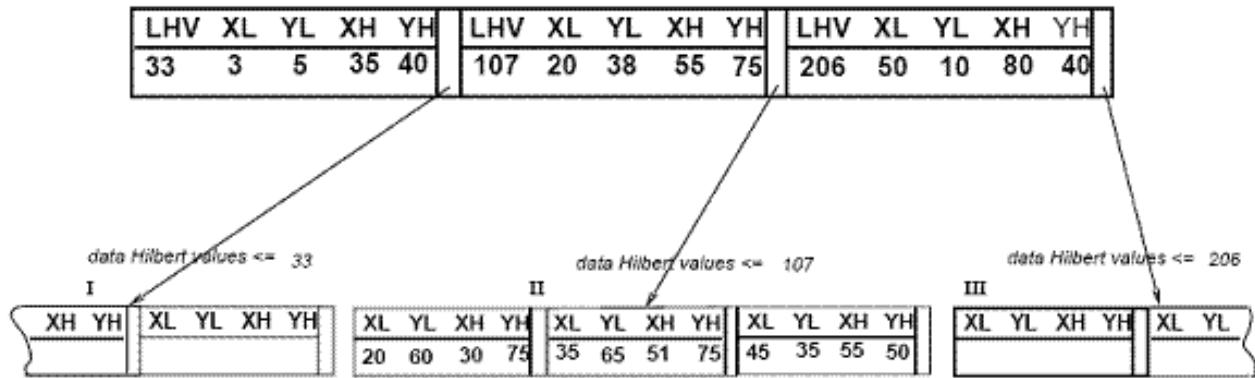


Figure 4: The file structure for the Hilbert R-tree

## Insertion

To insert a new rectangle  $r$  in the Hilbert R-tree, the Hilbert value  $h$  of the center of the new rectangle is used as a key. At each level the node with the minimum LHV value greater than  $h$  of all its siblings is chosen. When a leaf node is reached, the rectangle  $r$  is inserted in its correct order according to  $h$ . After a new rectangle is inserted in a leaf node  $N$ , **AdjustTree** is called to fix the MBR and LHV values in the upper-level nodes.

**Algorithm Insert(node Root, rect r):** /\* Inserts a new rectangle  $r$  in the Hilbert R-tree.  $h$  is the Hilbert value of the rectangle \*/

I1. Find the appropriate leaf node:

    Invoke ChooseLeaf( $r, h$ ) to select a leaf node  $L$  in which to place  $r$ .

I2. Insert  $r$  in a leaf node  $L$ :

    If  $L$  has an empty slot, insert  $r$  in  $L$  in the appropriate place according to the Hilbert order and return.

    If  $L$  is full, invoke HandleOverflow( $L, r$ ), which will return new leaf if split was inevitable,

I3. Propagate changes upward:

    Form a set  $S$  that contains  $L$ , its cooperating siblings and the new leaf (if any)

    Invoke AdjustTree( $S$ ).

I4. Grow tree taller:

    If node split propagation caused the root to split, create a new root whose children are the two resulting nodes.

**Algorithm ChooseLeaf(rect  $r$ , int  $h$ ):**

/\* Returns the leaf node in which to place a new rectangle  $r$ . \*/

C1. Initialize:

    Set  $N$  to be the root node.

C2. Leaf check:

    If  $N$  is a leaf\_ return  $N$ .

C3. Choose subtree:

    If  $N$  is a non-leaf node, choose the entry ( $R, ptr, LHV$ )

with the minimum LHV value greater than  $h$ .

C4. Descend until a leaf is reached:

Set  $N$  to the node pointed by  $ptr$  and repeat from C2.

**Algorithm AdjustTree(set S):**

/\*  $S$  is a set of nodes that contains the node being updated, its cooperating siblings (if overflow has occurred) and the newly

created node  $NN$  (if split has occurred). The routine ascends from the leaf level towards the root, adjusting MBR and LHV of nodes that cover the nodes in  $S$ . It propagates splits (if any) \*/

A1. If root level is reached, stop.

A2. Propagate node split upward:

Let  $N_p$  be the parent node of  $N$ .

If  $N$  has been split, let  $NN$  be the new node.

Insert  $NN$  in  $N_p$  in the correct order according to its Hilbert value if there is room. Otherwise, invoke  $HandleOverflow(N_p, NN)$ .

If  $N_p$  is split, let  $PP$  be the new node.

A3. Adjust the MBR's and LHV's in the parent level:

Let  $P$  be the set of parent nodes for the nodes in  $S$ .

Adjust the corresponding MBR's and LHV's of the nodes in  $P$  appropriately.

A4. Move up to next level:

Let  $S$  become the set of parent nodes  $P$ , with

$NN = PP$ , if  $N_p$  was split.

repeat from A1.

## Deletion

In the Hilbert R-tree there is no need to re-insert orphaned nodes whenever a father node underflows. Instead, keys can be borrowed from the siblings or the underflowing node is merged with its siblings. This is possible because the nodes have a clear ordering (according to Largest Hilbert Value, LHV); in contrast, in R-trees there is no such concept concerning sibling nodes. Notice that deletion operations require  $s$  cooperating siblings, while insertion operations require  $s - 1$  siblings.

**Algorithm Delete(r):**

D1. Find the host leaf:

Perform an exact match search to find the leaf node  $L$  that contains  $r$ .

D2. Delete  $r$  :

Remove  $r$  from node  $L$ .

D3. If  $L$  underflows

borrow some entries from  $s$  cooperating siblings.

if all the siblings are ready to underflow.

merge  $s + 1$  to  $s$  nodes,

adjust the resulting nodes.

D4. Adjust MBR and LHV in parent levels.

form a set  $S$  that contains  $L$  and its cooperating

siblings (if underflow has occurred).

invoke **AdjustTree(S)**.

## Overflow handling

The overflow handling algorithm in the Hilbert R-tree treats the overflowing nodes either by moving some of the entries to one of the  $s - 1$  cooperating siblings or by splitting  $s$  nodes into  $s + 1$  nodes.

**Algorithm HandleOverflow(node N, rect r):**

*/\* return the new node if a split occurred. \*/*

H1. Let  $\epsilon$  be a set that contains all the entries from N

and its  $s - 1$  cooperating siblings.

H2. Add r to  $\epsilon$ .

H3. If at least one of the  $s - 1$  cooperating siblings is not full,

distribute  $\epsilon$  evenly among the  $s$  nodes according to Hilbert values.

H4. If all the  $s$  cooperating siblings are full,

create a new node NN and

distribute  $\epsilon$  evenly among the  $s + 1$  nodes according

to Hilbert values

return NN.

## Notes and references

- [1] I. Kamel and C. Faloutsos, On Packing R-trees, Second International ACM Conference on Information and Knowledge Management (CIKM), pages 490–499, Washington D.C., 1993.
- [2] H. Jagadish. Linear clustering of objects with multiple attributes. In Proc. of ACM SIGMOD Conf., pages 332–342, Atlantic City, NJ, May 1990.
- [3] J. Griffiths. An algorithm for displaying a class of space-filling curves, Software-Practice and Experience 16(5), 403–411, May 1986.
- [4] T. Bially. Space-filling curves. Their generation and their application to bandwidth reduction. IEEE Trans. on Information Theory. IT15(6), 658–664, November 1969.
- [5] doi: 10.1145/93597.98741  
This citation will be automatically completed in the next few minutes. You can jump the queue or expand by hand ([http://en.wikipedia.org/wiki/Template:cite\\_doi/\\_10.1145.2f93597.98741?preload=Template:Cite\\_doi/preload&editintro=Template:Cite\\_doi/editintro&action=edit](http://en.wikipedia.org/wiki/Template:cite_doi/_10.1145.2f93597.98741?preload=Template:Cite_doi/preload&editintro=Template:Cite_doi/editintro&action=edit))
- I. Kamel and C. Faloutsos. Parallel R-Trees. In Proc. of ACM SIGMOD Conf., pages 195–204 San Diego, CA, June 1992. Also available as Tech. Report UMIACS TR 92-1, CS-TR-2820.
- I. Kamel and C. Faloutsos. Hilbert R-tree: An improved R-tree using fractals. In Proc. of VLDB Conf., pages 500–509, Santiago, Chile, September 1994. Also available as Tech\_ Report UMIACS TR 93-12.1 CS-TR-3032.1.
- N. Koudas, C. Faloutsos and I. Kamel. Declustering Spatial Databases on a Multi-Computer Architecture, International Conference on Extending Database Technology (EDBT), pages 592–614, 1996.
- N. Roussopoulos and D. Leifker. Direct spatial search on pictorial databases using Packed R-trees. In Proc. of ACM SIGMOD, pages 17–31, Austin, TX, May 1985.
- M. Schroeder. Fractals, Chaos, Power Laws: Minutes From an Infinite Paradise. W.H. Freeman and Company, NY, 1991.
- T. Sellis, N. Roussopoulos, and C. Faloutsos. The R+-Tree: a dynamic index for multi-dimensional objects. In Proc. 13th International Conference on VLDB, pages 507–518, England, September 1987.

# X-tree

---

In computer science, an **X-tree** is an index tree structure based on the R-tree used for storing data in many dimensions. It differs from R-trees, R+-trees and R\*-trees because it emphasizes prevention of overlap in the bounding boxes, which increasingly becomes a problem in high dimensions. In cases where nodes cannot be split without preventing overlap, the node split will be deferred, resulting in **super-nodes**. In extreme cases, the tree will linearize, which defends against worst-case behaviors observed in some other data structures.

## References

Berchtold, Stefan; Keim, Daniel A.; Kriegel, Hans-Peter (1996). "The X-Tree: An Index Structure for High-Dimensional Data"<sup>[1]</sup>. *Proceedings of the 22nd VLDB Conference* (Mumbai, India): 28–39.

## References

[1] <http://www.dbs.ifi.lmu.de/Publikationen/Papers/x-tree.ps>

# Metric tree

---

A **metric tree** is any tree data structure specialized to index data in metric spaces. Metric trees exploit properties of metric spaces such as the triangle inequality to make accesses to the data more efficient. Examples include the M-tree, vp-trees, cover trees, MVP Trees, and bk trees.<sup>[1]</sup>

## Multidimensional search

Most algorithms and data structures for searching a dataset are based on the classical binary search algorithm, and generalizations such as the k-d tree or range tree work by interleaving the binary search algorithm over the separate coordinates and treating each spatial coordinate as an independent search constraint. These data structures are well-suited for range query problems asking for every point  $(x, y)$  that satisfies  $\min_x \leq x \leq \max_x$  and  $\min_y \leq y \leq \max_y$ .

A limitation of these multidimensional search structures is that they are only defined for searching over objects that can be treated as vectors. They aren't applicable for the more general case in which the algorithm is given only a collection of objects and a function for measuring the distance or similarity between two objects. If, for example, someone were to create a function that returns a value indicating how similar one image is to another, a natural algorithmic problem would be to take a dataset of images and find the ones that are similar according to the function to a given query image.

---

## Metric data structures

If there is no structure to the similarity measure then a brute force search requiring the comparison of the query image to every image in the dataset is the best that can be done. If, however, the similarity function satisfies the triangle inequality then it is possible to use the result of each comparison to prune the set of candidates to be examined.

The first article on metric trees, as well as the first use of the term "metric tree", published in the open literature was by Jeffrey Uhlmann in 1991.<sup>[2]</sup> Other researchers were working independently on similar data structures, and research on metric tree data structures blossomed in the late 1990s and included an examination by Google co-founder Sergey Brin of their use for very large databases.<sup>[3]</sup> The first textbook on metric data structures was published in 2006.<sup>[1]</sup>

## References

- [1] Samet, Hanan (2006). *Foundations of multidimensional and metric data structures* (<http://books.google.dk/books?id=KrQdmLjTSaQC>). Morgan Kaufmann. ISBN 978-0-12-369446-1. .
- [2] Uhlmann, Jeffrey (1991). "Satisfying General Proximity/Similarity Queries with Metric Trees". *Information Processing Letters* **40** (4).
- [3] Brin, Sergey (1995). "Near Neighbor Search in Large Metric Spaces". *21st International Conference on Very Large Data Bases (VLDB)*.

## vP-tree

---

A **vantage-point tree**, or **vp-tree** is a BSP tree that segregates data in a metric space by choosing a position in the space (the "vantage point") and dividing the data points into two partitions: those that are nearer to the vantage point than a threshold, and those that are not. By repeatedly applying this procedure to partition the data into smaller and smaller sets, a tree data structure is created where neighbors in the tree are likely to be neighbors in the space.<sup>[1]</sup>

This iterative partitioning process is similar to that of a k-d tree, but uses circular (or spherical, hyperspherical, etc.) rather than rectilinear partitions. In 2D Euclidean space, this can be visualized as a series of circles segregating the data.

The vp-tree is particularly useful in dividing data in a non-standard metric space into a BSP tree.

## References

- [1] Yianilos, Peter N. (1993). "Data structures and algorithms for nearest neighbor search in general metric spaces" (<http://pnylab.com/pny/papers/vptree/vptree/>). *Proceedings of the fourth annual ACM-SIAM Symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics Philadelphia, PA, USA. pp. 311–321. pny93. . Retrieved 2008-08-22.



# BK-tree

---

A **BK-tree** is a metric tree suggested by Walter Austin Burkhard and Robert M. Keller<sup>BK73</sup> specifically adapted to discrete metric spaces. For simplicity, let us consider **integer** discrete metric  $d(x, y)$ . Then, BK-tree is defined in the following way. An arbitrary element  $a$  is selected as root node. The root node may have zero or more subtrees. The  $k$ -th subtree is recursively built of all elements  $b$  such that  $d(a, b) = k$ . BK-trees can be used for approximate string matching in a dictionary<sup>BN98</sup>.

## References

- W. Burkhard and R. Keller. Some approaches to best-match file searching, CACM, 1973<sup>[1]</sup>
- R. Baeza-Yates, W. Cunto, U. Manber, and S. Wu. Proximity matching using fixed queries trees. In M. Crochemore and D. Gusfield, editors, 5th Combinatorial Pattern Matching, LNCS 807, pages 198-212, Asilomar, CA, June 1994.
- Ricardo Baeza-Yates and Gonzalo Navarro. Fast Approximate String Matching in a Dictionary. Proc. SPIRE'98<sup>[2]</sup>

## External links

- A BK-tree implementation in Common Lisp<sup>[3]</sup> with test results and performance graphs.
- An explanation of BK-Trees and their relationship to metric spaces<sup>[4]</sup>

## References

- [1] <http://doi.acm.org/10.1145/362003.362025>
  - [2] [http://reference.kfupm.edu.sa/content/f/a/fast\\_approximate\\_string\\_matching\\_in\\_a\\_di\\_269284.pdf](http://reference.kfupm.edu.sa/content/f/a/fast_approximate_string_matching_in_a_di_269284.pdf)
  - [3] <http://cliki.net/bk-tree>
  - [4] <http://blog.notdot.net/2007/4/Damn-Cool-Algorithms-Part-1-BK-Trees>
-

# Hashes

## Hash table

Hash table		
Type	Unsorted associative array	
Invented	1953	
Time complexity in big O notation		
	Average	Worst case
Space	$O(n)^{[1]}$	$O(n)$
Search	$O(1 + n/k)$	$O(n)$
Insert	$O(1)$	$O(1)$
Delete	$O(1 + n/k)$	$O(n)$

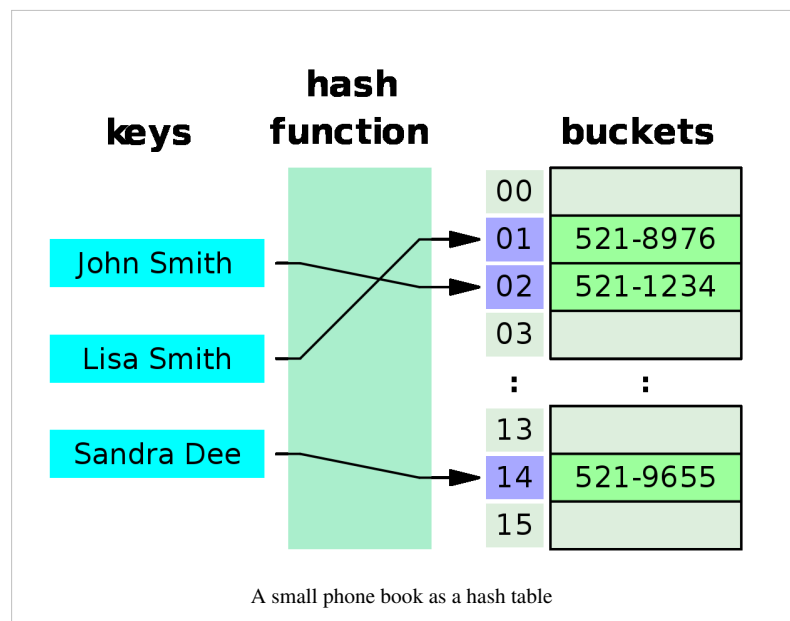
In computer science, a **hash table** or **hash map** is a data structure that uses a hash function to map identifying values, known as keys (e.g., a person's name), to their associated values (e.g., their telephone number). Thus, a hash table implements an associative array. The hash function is used to transform the key into the index (the *hash*) of an array element (the *slot* or *bucket*) where the corresponding value is to be sought.

Ideally, the hash function should map each possible key to a unique slot index, but this ideal is rarely achievable in practice (unless the hash keys are fixed; i.e. new entries are never added to the

table after it is created). Instead, most hash table designs assume that *hash collisions*—different keys that map to the same hash value—will occur and must be accommodated in some way.

In a well-dimensioned hash table, the average cost (number of instructions) for each lookup is independent of the number of elements stored in the table. Many hash table designs also allow arbitrary insertions and deletions of key-value pairs, at constant average (indeed, amortized<sup>[2]</sup>) cost per operation.<sup>[3][4]</sup>

In many situations, hash tables turn out to be more efficient than search trees or any other table lookup structure. For this reason, they are widely used in many kinds of computer software, particularly for associative arrays, database indexing, caches, and sets.



## Hash function

At the heart of the hash table algorithm is an array of items; this array is often simply called the *hash table*. Hash table algorithms calculate an index based on the data item's key and the length of the array. The index is used to find or insert the data into the array. The implementation of this calculation is the hash function,  $f$ :

```
index = f(key, arrayLength)
```

The hash function calculates an `index` into the array from the data `key` and `arrayLength` (the size of the array). For assembly language or other low-level programs, a trivial hash function can often create an index with just one or two inline machine instructions.

## Choosing a good hash function

A good hash function and implementation algorithm are essential for good hash table performance, but may be difficult to achieve.

A basic requirement is that the function should provide a uniform distribution of hash values. A non-uniform distribution increases the number of collisions, and the cost of resolving them. Uniformity is sometimes difficult to ensure by design, but may be evaluated empirically using statistical tests, e.g. a Pearson's chi-squared test for discrete uniform distributions<sup>[5]</sup> <sup>[6]</sup>

The distribution needs to be uniform only for table sizes  $s$  that occur in the application. In particular, if one uses dynamic resizing with exact doubling and halving of  $s$ , the hash function needs to be uniform only when  $s$  is a power of two. On the other hand, some hashing algorithms provide uniform hashes only when  $s$  is a prime number.<sup>[7]</sup>

For open addressing schemes, the hash function should also avoid *clustering*, the mapping of two or more keys to consecutive slots. Such clustering may cause the lookup cost to skyrocket, even if the load factor is low and collisions are infrequent. The popular multiplicative hash<sup>[3]</sup> is claimed to have particularly poor clustering behavior.<sup>[7]</sup>

Cryptographic hash functions are believed to provide good hash functions for any table size  $s$ , either by modulo reduction or by bit masking. They may also be appropriate, if there is a risk of malicious users trying to sabotage a network service by submitting requests designed to generate a large number of collisions in the server's hash tables. However, the risk of sabotage can also be avoided by cheaper methods (such as applying a secret salt to the data, or using a universal hash function).

Some authors claim that good hash functions should have the avalanche effect; that is, a single-bit change in the input key should affect, on average, half the bits in the output. Some popular hash functions do not have this property.

## Perfect hash function

If all keys are known ahead of time, a perfect hash function can be used to create a perfect hash table that has no collisions. If minimal perfect hashing is used, every location in the hash table can be used as well.

Perfect hashing allows for constant-time lookups in the worst case. This is in contrast to most chaining and open addressing methods, where the time for lookup is low on average, but may be very large (proportional to the number of entries) for some sets of keys.

## Collision resolution

Hash collisions are practically unavoidable when hashing a random subset of a large set of possible keys. For example, if 2,500 keys are hashed into a million buckets, even with a perfectly uniform random distribution, according to the birthday problem there is a 95% chance of at least two of the keys being hashed to the same slot.

Therefore, most hash table implementations have some collision resolution strategy to handle such events. Some common strategies are described below. All these methods require that the keys (or pointers to them) be stored in the table, together with the associated values.

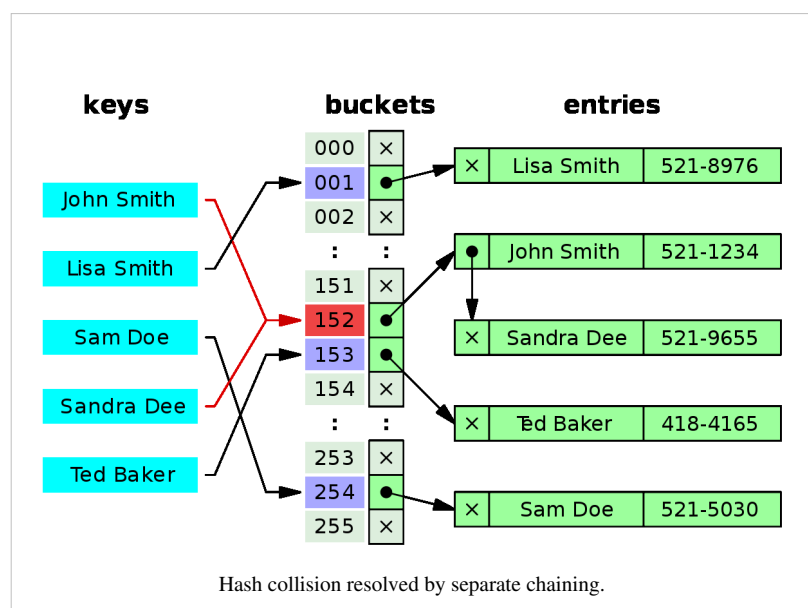
## Load factor

The performance of most collision resolution methods does not depend directly on the number  $n$  of stored entries. Instead, performance depends strongly on the table's *load factor*. *Load factor* is equal to  $n/s$ , the ratio of the number of stored entries  $n$  and the size  $s$  of the table's array of buckets. Sometimes this is referred to as the *fill factor*, as it represents the portion of the  $s$  buckets in the structure that are *filled* with one of the  $n$  stored entries. With a good hash function, the average lookup cost is nearly constant as the load factor increases from 0 to 0.7 (about 2/3 full) or so. Beyond that point, the probability of collisions and the cost of handling them increases.

A low load factor is not especially beneficial. As load factor approaches 0, the proportion of unused areas in the hash table increases, but there is not necessarily any reduction in search cost. This results in wasted memory.

## Separate chaining

In the strategy known as *separate chaining*, *direct chaining*, or simply *chaining*, each slot of the bucket array is a pointer to a linked list that contains the key-value pairs that hashed to the same location. Lookup requires scanning the list for an entry with the given key. Insertion requires adding a new entry record to either end of the list belonging to the hashed slot. Deletion requires searching the list and removing the element. (The technique is also called *open hashing* or *closed addressing*.)



Chained hash tables with linked lists are popular because they require only

basic data structures with simple algorithms, and can use simple hash functions that are unsuitable for other methods.

The cost of a table operation is that of scanning the entries of the selected bucket for the desired key. If the distribution of keys is sufficiently uniform, the *average* cost of a lookup depends only on the average number of keys per bucket—that is, on the load factor.

Chained hash tables remain effective even when the number of table entries  $n$  is much higher than the number of slots. Their performance degrades more gracefully (linearly) with the load factor. For example, a chained hash table with 1000 slots and 10,000 stored keys (load factor 10) is five to ten times slower than a 10,000-slot table (load factor 1); but still 1000 times faster than a plain sequential list, and possibly even faster than a balanced search tree.

For separate-chaining, the worst-case scenario is when all entries are inserted into the same bucket, in which case the hash table is ineffective and the cost is that of searching the bucket data structure. If the latter is a linear list, the

lookup procedure may have to scan all its entries; so the worst-case cost is proportional to the number  $n$  of entries in the table.

The bucket chains are often implemented as ordered lists, sorted by the key field; this choice approximately halves the average cost of unsuccessful lookups, compared to an unordered list. However, if some keys are much more likely to come up than others, an unordered list with move-to-front heuristic may be more effective. More sophisticated data structures, such as balanced search trees, are worth considering only if the load factor is large (about 10 or more), or if the hash distribution is likely to be very non-uniform, or if one must guarantee good performance even in a worst-case scenario. However, using a larger table and/or a better hash function may be even more effective in those cases.

Chained hash tables also inherit the disadvantages of linked lists. When storing small keys and values, the space overhead of the `next` pointer in each entry record can be significant. An additional disadvantage is that traversing a linked list has poor cache performance, making the processor cache ineffective.

### Separate chaining with list heads

Some chaining implementations store the first record of each chain in the slot array itself.<sup>[4]</sup> The purpose is to increase cache efficiency of hash table access. To save memory space, such hash tables often have about as many slots as stored entries, meaning that many slots have two or more entries.

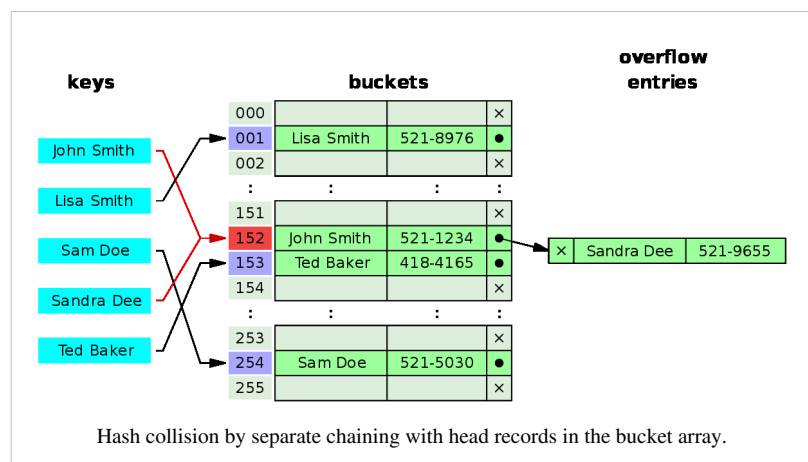
### Separate chaining with other structures

Instead of a list, one can use any other

data structure that supports the required operations. For example, by using a self-balancing tree, the theoretical worst-case time of common hash table operations (insertion, deletion, lookup) can be brought down to  $O(\log n)$  rather than  $O(n)$ . However, this approach is only worth the trouble and extra memory cost if long delays must be avoided at all costs (e.g. in a real-time application), or if one must guard against many entries hashed to the same slot (e.g. if one expects extremely non-uniform distributions, or in the case of web sites or other publicly accessible services, which are vulnerable to malicious key distributions in requests).

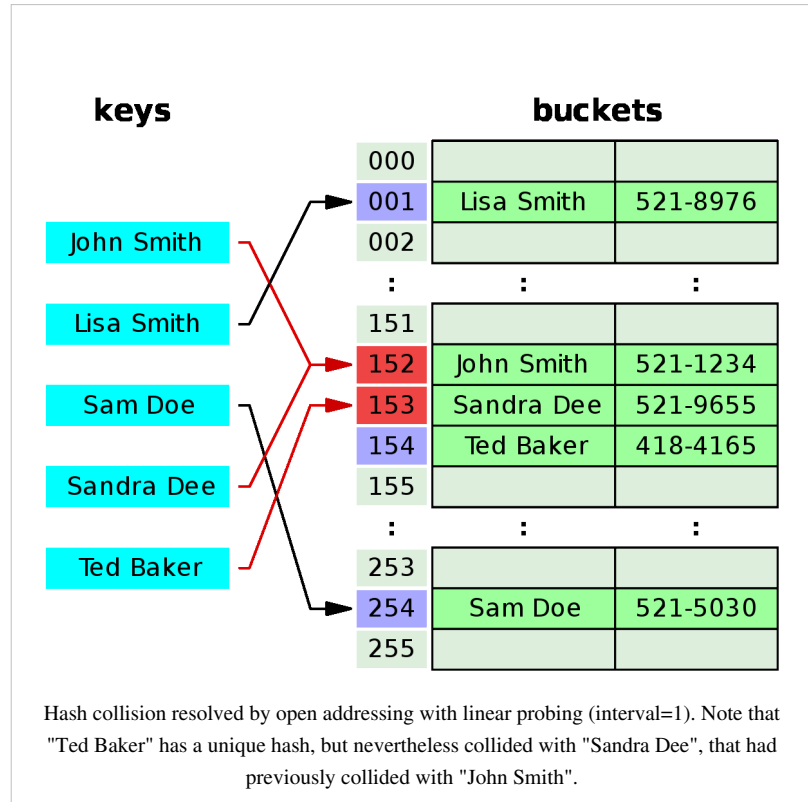
The variant called array hash table uses a dynamic array to store all the entries that hash to the same slot.<sup>[8][9][10]</sup> Each newly inserted entry gets appended to the end of the dynamic array that is assigned to the slot. The dynamic array is resized in an *exact-fit* manner, meaning it is grown only by as many bytes as needed. Alternative techniques such as growing the array by block sizes or *pages* were found to improve insertion performance, but at a cost in space. This variation makes more efficient use of CPU caching and the translation lookaside buffer (TLB), because slot entries are stored in sequential memory positions. It also dispenses with the `next` pointers that are required by linked lists, which saves space. Despite frequent array resizing, space overheads incurred by operating system such as memory fragmentation, were found to be small.

An elaboration on this approach is the so-called dynamic perfect hashing,<sup>[11]</sup> where a bucket that contains  $k$  entries is organized as a perfect hash table with  $k^2$  slots. While it uses more memory ( $n^2$  slots for  $n$  entries, in the worst case and  $n*k$  slots in the average case), this variant has guaranteed constant worst-case lookup time, and low amortized time for insertion.



## Open addressing

In another strategy, called open addressing, all entry records are stored in the bucket array itself. When a new entry has to be inserted, the buckets are examined, starting with the hashed-to slot and proceeding in some *probe sequence*, until an unoccupied slot is found. When searching for an entry, the buckets are scanned in the same sequence, until either the target record is found, or an unused array slot is found, which indicates that there is no such key in the table.<sup>[12]</sup> The name "open addressing" refers to the fact that the location ("address") of the item is not determined by its hash value. (This method is also called **closed hashing**; it should not be confused with "open hashing" or "closed addressing" that usually mean separate chaining.)



Well-known probe sequences include:

- Linear probing, in which the interval between probes is fixed (usually 1)
- Quadratic probing, in which the interval between probes is increased by adding the successive outputs of a quadratic polynomial to the starting value given by the original hash computation
- Double hashing, in which the interval between probes is computed by another hash function

A drawback of all these open addressing schemes is that the number of stored entries cannot exceed the number of slots in the bucket array. In fact, even with good hash functions, their performance dramatically degrades when the load factor grows beyond 0.7 or so. Thus a more aggressive resize scheme is needed. Separate linking works correctly with any load factor, although performance is likely to be reasonable if it is kept below 2 or so. For many applications, these restrictions mandate the use of dynamic resizing, with its attendant costs.

Open addressing schemes also put more stringent requirements on the hash function: besides distributing the keys more uniformly over the buckets, the function must also minimize the clustering of hash values that are consecutive in the probe order. Using separate chaining, the only concern is that too many objects map to the *same* hash value; whether they are adjacent or nearby is completely irrelevant.

Open addressing only saves memory if the entries are small (less than four times the size of a pointer) and the load factor is not too small. If the load factor is close to zero (that is, there are far more buckets than stored entries), open addressing is wasteful even if each entry is just two words.

Open addressing avoids the time overhead of allocating each new entry record, and can be implemented even in the absence of a memory allocator. It also avoids the extra indirection required to access the first entry of each bucket (that is, usually the only one). It also has better locality of reference, particularly with linear probing. With small record sizes, these factors can yield better performance than chaining, particularly for lookups.

Hash tables with open addressing are also easier to serialize, because they do not use pointers.

On the other hand, normal open addressing is a poor choice for large elements, because these elements fill entire CPU cache lines (negating the cache advantage), and a large amount of space is wasted on large empty table slots. If the open addressing table only stores references to elements (external storage), it uses space comparable to chaining even for large records but loses its speed advantage.

Generally speaking, open addressing is better used for hash tables with small records that can be stored within the table (internal storage) and fit in a cache line. They are particularly suitable for elements of one word or less. If the table is expected to have a high load factor, the records are large, or the data is variable-sized, chained hash tables often perform as well or better.

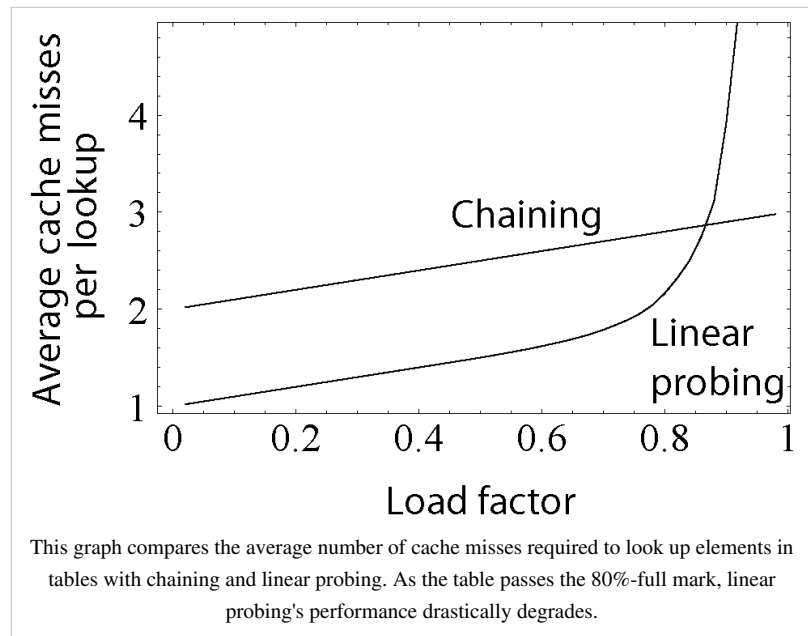
Ultimately, used sensibly, any kind of hash table algorithm is usually fast *enough*; and the percentage of a calculation spent in hash table code is low. Memory usage is rarely considered excessive. Therefore, in most cases the differences between these algorithms are marginal, and other considerations typically come into play.

## Coalesced hashing

A hybrid of chaining and open addressing, coalesced hashing links together chains of nodes within the table itself.<sup>[12]</sup> Like open addressing, it achieves space usage and (somewhat diminished) cache advantages over chaining. Like chaining, it does not exhibit clustering effects; in fact, the table can be efficiently filled to a high density. Unlike chaining, it cannot have more elements than table slots.

## Robin Hood hashing

One interesting variation on double-hashing collision resolution is Robin Hood hashing.<sup>[13]</sup> The idea is that a new key may displace a key already inserted, if its probe count is larger than that of the key at the current position. The net effect of this is that it reduces worst case search times in the table. This is similar to Knuth's ordered hash tables except that the criterion for bumping a key does not depend on a direct relationship between the keys. Since both the worst case and the variation in the number of probes is reduced dramatically, an interesting variation is to probe the table starting at the expected successful probe value and then expand from that position in both directions.<sup>[14]</sup> External Robin Hashing is an extension of this algorithm where the table is stored in an external file and each table position corresponds to a fixed-sized page or bucket with  $B$  records.<sup>[15]</sup>



## Cuckoo hashing

Another alternative open-addressing solution is cuckoo hashing, which ensures constant lookup time in the worst case, and constant amortized time for insertions and deletions. It uses two or more hash functions, which means any key/value pair could be in two or more locations. For lookup, the first hash function is used; if the key/value is not found, then the second hash function is used, and so on. If a collision happens during insertion, then the key is re-hashed with the second hash function to map it to another bucket. If all hash functions are used and there is still a collision, then the key it collided with is removed to make space for the new key, and the old key is re-hashed with one of the other hash functions, which maps it to another bucket. If that location also results in a collision, then the process repeats until there is no collision or the process traverses all the buckets, at which point the table is resized. By combining multiple hash functions with multiple cells per bucket, very high space utilisation can be achieved.

## Hopscotch hashing

Another alternative open-addressing solution is hopscotch hashing,<sup>[16]</sup> which combines the approaches of cuckoo hashing and linear probing, yet seems in general to avoid their limitations. In particular it works well even when the load factor grows beyond 0.9. The algorithm is well suited for implementing a resizable concurrent hash table.

The hopscotch hashing algorithm works by defining a neighborhood of buckets near the original hashed bucket, where a given entry is always found. Thus, search is limited to the number of entries in this neighborhood, which is logarithmic in the worst case, constant on average, and with proper alignment of the neighborhood typically requires one cache miss. When inserting an entry, one first attempts to add it to a bucket in the neighborhood. However, if all buckets in this neighborhood are occupied, the algorithm traverses buckets in sequence until an open slot (an unoccupied bucket) is found (as in linear probing). At that point, since the empty bucket is outside the neighborhood, items are repeatedly displaced in a sequence of hops. (This is similar to cuckoo hashing, but with the difference that in this case the empty slot is being moved into the neighborhood, instead of items being moved out with the hope of eventually finding an empty slot.) Each hop brings the open slot closer to the original neighborhood, without invalidating the neighborhood property of any of the buckets along the way. In the end, the open slot has been moved into the neighborhood, and the entry being inserted can be added to it.

## Dynamic resizing

To keep the load factor under a certain limit, e.g. under  $3/4$ , many table implementations expand the table when items are inserted. For example, in Java's `HashMap` class the default load factor threshold for table expansion is 0.75.

Since buckets are usually implemented on top of a dynamic array and any constant proportion for resizing greater than 1 will keep the load factor under the desired limit, the exact choice of the constant is determined by the same space-time tradeoff as for dynamic arrays.

Resizing is accompanied by a full or incremental table **rehash** whereby existing items are mapped to new bucket locations.

To limit the proportion of memory wasted due to empty buckets, some implementations also shrink the size of the table—followed by a rehash—when items are deleted. From the point of space-time tradeoffs, this operation is similar to the deallocation in dynamic arrays.



## Resizing by copying all entries

A common approach is to automatically trigger a complete resizing when the load factor exceeds some threshold  $r_{\max}$ . Then a new larger table is allocated, all the entries of the old table are removed and inserted into this new table, and the old table is returned to the free storage pool. Symmetrically, when the load factor falls below a second threshold  $r_{\min}$ , all entries are moved to a new smaller table.

If the table size increases or decreases by a fixed percentage at each expansion, the total cost of these resizings, amortized over all insert and delete operations, is still a constant, independent of the number of entries  $n$  and of the number  $m$  of operations performed.

For example, consider a table that was created with the minimum possible size and is doubled each time the load ratio exceeds some threshold. If  $m$  elements are inserted into that table, the total number of extra re-insertions that occur in all dynamic resizings of the table is at most  $m - 1$ . In other words, dynamic resizing roughly doubles the cost of each insert or delete operation.

## Incremental resizing

Some hash table implementations, notably in real-time systems, cannot pay the price of enlarging the hash table all at once, because it may interrupt time-critical operations. If one cannot avoid dynamic resizing, a solution is to perform the resizing gradually:

- During the resize, allocate the new hash table, but keep the old table unchanged.
- In each lookup or delete operation, check both tables.
- Perform insertion operations only in the new table.
- At each insertion also move  $r$  elements from the old table to the new table.
- When all elements are removed from the old table, deallocate it.

To ensure that the old table is completely copied over before the new table itself needs to be enlarged, it is necessary to increase the size of the table by a factor of at least  $(r + 1)/r$  during resizing.

## Monotonic keys

If it is known that key values will always increase (or decrease) monotonically, then a variation of consistent hashing can be achieved by keeping a list of the single most recent key value at each hash table resize operation. Upon lookup, keys that fall in the ranges defined by these list entries are directed to the appropriate hash function—and indeed hash table—both of which can be different for each range. Since it is common to grow the overall number of entries by doubling, there will only be  $O(\lg(N))$  ranges to check, and binary search time for the redirection would be  $O(\lg(\lg(N)))$ . As with consistent hashing, this approach guarantees that any key's hash, once issued, will never change, even when the hash table is later grown.

## Other solutions

Linear hashing<sup>[17]</sup> is a hash table algorithm that permits incremental hash table expansion. It is implemented using a single hash table, but with two possible look-up functions.

Another way to decrease the cost of table resizing is to choose a hash function in such a way that the hashes of most values do not change when the table is resized. This approach, called consistent hashing, is prevalent in disk-based and distributed hashes, where rehashing is prohibitively costly.

## Performance analysis

In the simplest model, the hash function is completely unspecified and the table does not resize. For the best possible choice of hash function, a table of size  $n$  with open addressing has no collisions and holds up to  $n$  elements, with a single comparison for successful lookup, and a table of size  $n$  with chaining and  $k$  keys has the minimum  $\max(0, k-n)$  collisions and  $O(1 + k/n)$  comparisons for lookup. For the worst choice of hash function, every insertion causes a collision, and hash tables degenerate to linear search, with  $\Omega(k)$  amortized comparisons per insertion and up to  $k$  comparisons for a successful lookup.

Adding rehashing to this model is straightforward. As in a dynamic array, geometric resizing by a factor of  $b$  implies that only  $k/b^i$  keys are inserted  $i$  or more times, so that the total number of insertions is bounded above by  $bk/(b-1)$ , which is  $O(k)$ . By using rehashing to maintain  $k < n$ , tables using both chaining and open addressing can have unlimited elements and perform successful lookup in a single comparison for the best choice of hash function.

In more realistic models, the hash function is a random variable over a probability distribution of hash functions, and performance is computed on average over the choice of hash function. When this distribution is uniform, the assumption is called "simple uniform hashing" and it can be shown that hashing with chaining requires  $\Theta(1 + k/n)$  comparisons on average for an unsuccessful lookup, and hashing with open addressing requires  $\Theta(1/(1 - k/n))$ .<sup>[18]</sup> Both these bounds are constant, if we maintain  $k/n < c$  using table resizing, where  $c$  is a fixed constant less than 1.

## Features

### Advantages

The main advantage of hash tables over other table data structures is speed. This advantage is more apparent when the number of entries is large. Hash tables are particularly efficient when the maximum number of entries can be predicted in advance, so that the bucket array can be allocated once with the optimum size and never resized.

If the set of key-value pairs is fixed and known ahead of time (so insertions and deletions are not allowed), one may reduce the average lookup cost by a careful choice of the hash function, bucket table size, and internal data structures. In particular, one may be able to devise a hash function that is collision-free, or even perfect (see below). In this case the keys need not be stored in the table.

### Drawbacks

Although operations on a hash table take constant time on average, the cost of a good hash function can be significantly higher than the inner loop of the lookup algorithm for a sequential list or search tree. Thus hash tables are not effective when the number of entries is very small. (However, in some cases the high cost of computing the hash function can be mitigated by saving the hash value together with the key.)

For certain string processing applications, such as spell-checking, hash tables may be less efficient than tries, finite automata, or Judy arrays. Also, if each key is represented by a small enough number of bits, then, instead of a hash table, one may use the key directly as the index into an array of values. Note that there are no collisions in this case.

The entries stored in a hash table can be enumerated efficiently (at constant cost per entry), but only in some pseudo-random order. Therefore, there is no efficient way to locate an entry whose key is *nearest* to a given key. Listing all  $n$  entries in some specific order generally requires a separate sorting step, whose cost is proportional to  $\log(n)$  per entry. In comparison, ordered search trees have lookup and insertion cost proportional to  $\log(n)$ , but allow finding the nearest key at about the same cost, and *ordered* enumeration of all entries at constant cost per entry.

If the keys are not stored (because the hash function is collision-free), there may be no easy way to enumerate the keys that are present in the table at any given moment.

Although the *average* cost per operation is constant and fairly small, the cost of a single operation may be quite high. In particular, if the hash table uses dynamic resizing, an insertion or deletion operation may occasionally take time

proportional to the number of entries. This may be a serious drawback in real-time or interactive applications.

Hash tables in general exhibit poor locality of reference—that is, the data to be accessed is distributed seemingly at random in memory. Because hash tables cause access patterns that jump around, this can trigger microprocessor cache misses that cause long delays. Compact data structures such as arrays searched with linear search may be faster, if the table is relatively small and keys are integers or other short strings. According to Moore's Law, cache sizes are growing exponentially and so what is considered "small" may be increasing. The optimal performance point varies from system to system.

Hash tables become quite inefficient when there are many collisions. While extremely uneven hash distributions are extremely unlikely to arise by chance, a malicious adversary with knowledge of the hash function may be able to supply information to a hash that creates worst-case behavior by causing excessive collisions, resulting in very poor performance, e.g. a denial of service attack<sup>[19]</sup>. In critical applications, universal hashing can be used; a data structure with better worst-case guarantees may be preferable.<sup>[20]</sup>

## Uses

### Associative arrays

Hash tables are commonly used to implement many types of in-memory tables. They are used to implement associative arrays (arrays whose indices are arbitrary strings or other complicated objects), especially in interpreted programming languages like AWK, Perl, and PHP.

When storing a new item into a multimap and a hash collision occurs, the multimap unconditionally stores both items.

When storing a new item into a typical associative array and a hash collision occurs, but the actual keys themselves are different, the associative array likewise stores both items. However, if the key of the new item exactly matches the key of an old item, the associative array typically erases the old item and overwrites it with the new item, so every item in the table has a unique key.

### Database indexing

Hash tables may also be used as disk-based data structures and database indices (such as in dbm) although B-trees are more popular in these applications.

### Caches

Hash tables can be used to implement caches, auxiliary data tables that are used to speed up the access to data that is primarily stored in slower media. In this application, hash collisions can be handled by discarding one of the two colliding entries—usually erasing the old item that is currently stored in the table and overwriting it with the new item, so every item in the table has a unique hash value.

### Sets

Besides recovering the entry that has a given key, many hash table implementations can also tell whether such an entry exists or not.

Those structures can therefore be used to implement a set data structure, which merely records whether a given key belongs to a specified set of keys. In this case, the structure can be simplified by eliminating all parts that have to do with the entry values. Hashing can be used to implement both static and dynamic sets.

---

## Object representation

Several dynamic languages, such as Perl, Python, JavaScript, and Ruby, use hash tables to implement objects. In this representation, the keys are the names of the members and methods of the object, and the values are pointers to the corresponding member or method.

## Unique data representation

Hash tables can be used by some programs to avoid creating multiple character strings with the same contents. For that purpose, all strings in use by the program are stored in a single hash table, which is checked whenever a new string has to be created. This technique was introduced in Lisp interpreters under the name hash consing, and can be used with many other kinds of data (expression trees in a symbolic algebra system, records in a database, files in a file system, binary decision diagrams, etc.)

## Implementations

### In programming languages

Many programming languages provide hash table functionality, either as built-in associative arrays or as standard library modules. In C++11, for example, the `unordered_map` class provides hash tables for keys and values of arbitrary type.

In PHP 5, the Zend 2 engine uses one of the hash functions from Daniel J. Bernstein to generate the hash values used in managing the mappings of data pointers stored in a hash table. In the PHP source code, it is labelled as `DJBX33A` (Daniel J. Bernstein, Times 33 with Addition).

Python's built-in hash table implementation, in the form of the `dict` type, as well as Perl's hash type (`%`) are highly optimized as they are used internally to implement namespaces.

In the .NET Framework, support for hash tables is provided via the non-generic `Hashtable` and generic `Dictionary` classes, which store key-value pairs, and the generic `HashSet` class, which stores only values.

### Independent packages

- SparseHash <sup>[21]</sup> (formerly Google SparseHash) An extremely memory-efficient `hash_map` implementation, with only 2 bits/entry of overhead. The SparseHash library has several C++ hash map implementations with different performance characteristics, including one that optimizes for memory use and another that optimizes for speed.
- SunriseDD <sup>[22]</sup> An open source C library for hash table storage of arbitrary data objects with lock-free lookups, built-in reference counting and guaranteed order iteration. The library can participate in external reference counting systems or use its own built-in reference counting. It comes with a variety of hash functions and allows the use of runtime supplied hash functions via callback mechanism. Source code is well documented.
- uthash <sup>[23]</sup> This is an easy-to-use hash table for C structures.

## History

The idea of hashing arose independently in different places. In January 1953, H. P. Luhn wrote an internal IBM memorandum that used hashing with chaining.<sup>[24]</sup> G. N. Amdahl, E. M. Boehme, N. Rochester, and Arthur Samuel implemented a program using hashing at about the same time. Open addressing with linear probing (relatively prime stepping) is credited to Amdahl, but Ershov (in Russia) had the same idea.<sup>[24]</sup>

## References

- [1] Thomas H. Cormen [et al.] (2009). *Introduction to Algorithms* (3rd ed.). Massachusetts Institute of Technology. pp. 253–280. ISBN 978-0-262-03384-8.
- [2] Charles E. Leiserson, *Amortized Algorithms, Table Doubling, Potential Method* ([http://videlectures.net/mit6046jf05\\_leiserson Lec13/](http://videlectures.net/mit6046jf05_leiserson Lec13/)) Lecture 13, course MIT 6.046J/18.410J Introduction to Algorithms—Fall 2005
- [3] Donald Knuth (1998). *The Art of Computer Programming*. **3: Sorting and Searching** (2nd ed.). Addison-Wesley. pp. 513–558. ISBN 0-201-89685-0.
- [4] Cormen, Thomas H.; Leiserson, Charles E.; Rivest, Ronald L.; Stein, Clifford (2001). *Introduction to Algorithms* (2nd ed.). MIT Press and McGraw-Hill. 221–252. ISBN 978-0-262-53196-2.
- [5] Karl Pearson (1900). "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling". *Philosophical Magazine, Series 5* **50** (302): pp. 157–175.
- [6] Robin Plackett (1983). "Karl Pearson and the Chi-Squared Test". *International Statistical Review (International Statistical Institute (ISI))* **51** (1): pp. 59–72.
- [7] Thomas Wang (1997), Prime Double Hash Table (<http://www.concentric.net/~Ttwang/tech/primehash.htm>). Accessed April 27, 2012
- [8] Askitis, Nikolas; Zobel, Justin (October 2005). *Cache-conscious Collision Resolution in String Hash Tables* (<http://www.springerlink.com/content/b61721172558qt03/>). **3772/2005**. pp. 91–102. doi:10.1007/11575832\_11. ISBN 978-3-540-29740-6. .
- [9] Askitis, Nikolas; Sinha, Ranjan (2010). "Engineering scalable, cache and space efficient tries for strings" (<http://www.springerlink.com/content/86574173183j6565/>). *The VLDB Journal* **17** (5): 633–660. doi:10.1007/s00778-010-0183-9. ISSN 1066-8888. .
- [10] Askitis, Nikolas (2009). *Fast and Compact Hash Tables for Integer Keys* (<http://crpit.com/confpapers/CRPITV91Askitis.pdf>). **91**. pp. 113–122. ISBN 978-1-920682-72-9. .
- [11] Erik Demaine, Jeff Lind. 6.897: Advanced Data Structures. MIT Computer Science and Artificial Intelligence Laboratory. Spring 2003. [http://courses.csail.mit.edu/6.897/spring03/scribe\\_notes/L2/lecture2.pdf](http://courses.csail.mit.edu/6.897/spring03/scribe_notes/L2/lecture2.pdf)
- [12] Tenenbaum, Aaron M.; Langsam, Yedidyah; Augenstein, Moshe J. (1990). *Data Structures Using C*. Prentice Hall. pp. 456–461, p. 472. ISBN 0-13-199746-7.
- [13] Celis, Pedro (1986). *Robin Hood hashing* (Technical report). CS-86-14.
- [14] Viola, Alfredo (October 2005). "Exact distribution of individual displacements in linear probing hashing". *Transactions on Algorithms (TALG) (ACM)* **1** (2): 214–242. doi:10.1145/1103963.1103965.
- [15] Celis, Pedro (March, 1988). *External Robin Hood Hashing* (Technical report). TR246.
- [16] Herlihy, Maurice and Shavit, Nir and Tzafrir, Moran (2008). "Hopscotch Hashing". *DISC '08: Proceedings of the 22nd international symposium on Distributed Computing*. Arcachon, France: Springer-Verlag. pp. 350–364.
- [17] Litwin, Witold (1980). "Linear hashing: A new tool for file and table addressing". *Proc. 6th Conference on Very Large Databases*. pp. 212–223.
- [18] Doug Dunham. CS 4521 Lecture Notes (<http://www.duluth.umn.edu/~ddunham/cs4521s09/notes/ch11.txt>). University of Minnesota Duluth. Theorems 11.2, 11.6. Last modified 21 April 2009.
- [19] Alexander Klink and Julian Wälde's *Efficient Denial of Service Attacks on Web Application Platforms* ([http://events.ccc.de/congress/2011/Fahrplan/attachments/2007\\_28C3\\_Effective\\_DoS\\_on\\_web\\_application\\_platforms.pdf](http://events.ccc.de/congress/2011/Fahrplan/attachments/2007_28C3_Effective_DoS_on_web_application_platforms.pdf)), December 28, 2011, 28th Chaos Communication Congress. Berlin, Germany.
- [20] Crosby and Wallach's *Denial of Service via Algorithmic Complexity Attacks* ([http://www.cs.rice.edu/~scrosby/hash/CrosbyWallach\\_UsenixSec2003.pdf](http://www.cs.rice.edu/~scrosby/hash/CrosbyWallach_UsenixSec2003.pdf)).
- [21] <http://code.google.com/p/sparsehash/>
- [22] <http://www.sunrisetel.net/software/devtools/sunrise-data-dictionary.shtml>
- [23] <http://uthash.sourceforge.net/>
- [24] Mehta, Dinesh P.; Sahni, Sartaj. *Handbook of Datastructures and Applications*. pp. 9–15. ISBN 1-58488-435-5.

## Further reading

Tamassia, Roberto; Michael T. Goodrich (2006). "Chapter Nine: Maps and Dictionaries" (in English). *Data structures and algorithms in Java : [updated for Java 5.0]* (4th ed.). Hoboken, N.J.: Wiley. pp. 369-418. ISBN 0-471-73884-0.

## External links

- A Hash Function for Hash Table Lookup (<http://www.burtleburtle.net/bob/hash/doobs.html>) by Bob Jenkins.
  - Hash Tables (<http://www.sparknotes.com/cs/searching/hashtables/summary.html>) by SparkNotes—explanation using C
  - Hash functions (<http://www.azillionmonkeys.com/qed/hash.html>) by Paul Hsieh
  - Design of Compact and Efficient Hash Tables for Java (<http://blog.griddynamics.com/2011/03/ultimate-sets-and-maps-for-java-part-i.html>)
  - Libhashish (<http://libhashish.sourceforge.net/>) hash library
  - NIST entry on hash tables (<http://www.nist.gov/dads/HTML/hashtab.html>)
  - Open addressing hash table removal algorithm from ICI programming language, *ici\_set\_unassign* in *set.c* (<http://ici.cvs.sourceforge.net/ici/ici/set.c?view=markup>) (and other occurrences, with permission).
  - A basic explanation of how the hash table works by Reliable Software (<http://www.relisoft.com/book/lang/pointer/8hash.html>)
  - Lecture on Hash Tables (<http://compgeom.cs.uiuc.edu/~jeffe/teaching/373/notes/06-hashing.pdf>)
  - Hash-tables in C (<http://task3.cc/308/hash-maps-with-linear-probing-and-separate-chaining/>)—two simple and clear examples of hash tables implementation in C with linear probing and chaining
  - Open Data Structures - Chapter 5 - Hash Tables ([http://opendatastructures.org/versions/edition-0.1e/ods-java/5\\_Hash\\_Tables.html](http://opendatastructures.org/versions/edition-0.1e/ods-java/5_Hash_Tables.html))
  - MIT's Introduction to Algorithms: Hashing 1 (<http://video.google.com/videoplay?docid=-727485696209877198&q=source:014117792397255896270&hl=en>) MIT OCW lecture Video
  - MIT's Introduction to Algorithms: Hashing 2 (<http://video.google.com/videoplay?docid=2307261494964091254&q=source:014117792397255896270&hl=en>) MIT OCW lecture Video
  - How to sort a HashMap (Java) and keep the duplicate entries (<http://www.lampos.net/sort-hashmap>)
-

# Hash function

A **hash function** is any algorithm or subroutine that maps large data sets of variable length, called *keys*, to smaller data sets of a fixed length. For example, a person's name, having a variable length, could be hashed to a single integer. The values returned by a hash function are called **hash values**, **hash codes**, **hash sums**, **checksums** or simply **hashes**.

## Descriptions

Hash functions are mostly used to accelerate table lookup or data comparison tasks such as finding items in a database, detecting duplicated or similar records in a large file, finding similar stretches in DNA sequences, and so on.

A hash function should be referentially transparent (stable), *i.e.*, if called twice on input that is "equal" (for example, strings that consist of the same sequence of characters), it should give the same result. This is a contract in many programming languages that allow the user to override equality and hash functions for an object: if two objects are equal, their hash codes must be the same. This is crucial to finding an element in a hash table quickly, because two of the same element would both hash to the same slot.

Some hash functions may map two or more keys to the same hash value, causing a collision. Such hash functions try to map the keys to the hash values as evenly as possible because collisions become more frequent as hash tables fill up. Thus, single-digit hash values are frequently restricted to 80% of the size of the table. Depending on the algorithm used, other properties may be required as well, such as double hashing and linear probing. Although the idea was conceived in the 1950s,<sup>[1]</sup> the design of good hash functions is still a topic of active research.

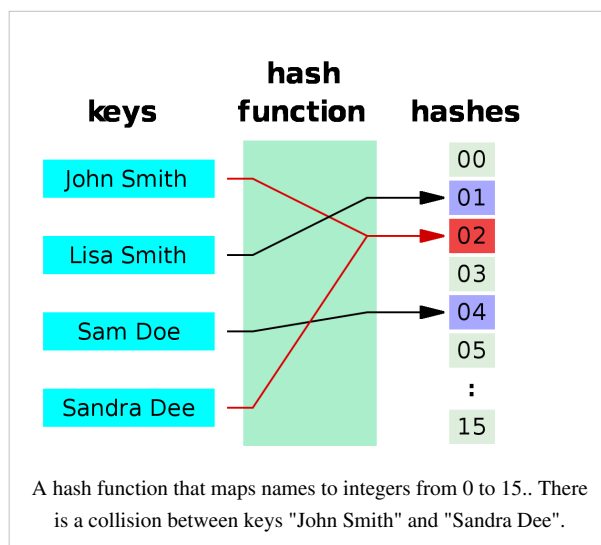
Hash functions are related to (and often confused with) checksums, check digits, fingerprints, randomization functions, error correcting codes, and cryptographic hash functions. Although these concepts overlap to some extent, each has its own uses and requirements and is designed and optimized differently. The HashKeeper database maintained by the American National Drug Intelligence Center, for instance, is more aptly described as a catalog of file fingerprints than of hash values.

## Hash tables

Hash functions are primarily used in hash tables, to quickly locate a data record (for example, a dictionary definition) given its search key (the headword). Specifically, the hash function is used to map the search key to the hash. The index gives the place where the corresponding record should be stored. Hash tables, in turn, are used to implement associative arrays and dynamic sets.

In general, a hashing function may map several different keys to the same index. Therefore, each slot of a hash table is associated with (implicitly or explicitly) a set of records, rather than a single record. For this reason, each slot of a hash table is often called a *bucket*, and hash values are also called *bucket indices*.

Thus, the hash function only hints at the record's location—it tells where one should start looking for it. Still, in a half-full table, a good hash function will typically narrow the search down to only one or two entries.



## Caches

Hash functions are also used to build caches for large data sets stored in slow media. A cache is generally simpler than a hashed search table, since any collision can be resolved by discarding or writing back the older of the two colliding items. This is also used in file comparison.

## Bloom filters

Hash functions are an essential ingredient of the Bloom filter, a compact data structure that provides an enclosing approximation to a set of them.

## Finding duplicate records

When storing records in a large unsorted file, one may use a hash function to map each record to an index into a table  $T$ , and collect in each bucket  $T[i]$  a list of the numbers of all records with the same hash value  $i$ . Once the table is complete, any two duplicate records will end up in the same bucket. The duplicates can then be found by scanning every bucket  $T[i]$  which contains two or more members, fetching those records, and comparing them. With a table of appropriate size, this method is likely to be much faster than any alternative approach (such as sorting the file and comparing all consecutive pairs).

## Finding similar records

Hash functions can also be used to locate table records whose key is similar, but not identical, to a given key; or pairs of records in a large file which have similar keys. For that purpose, one needs a hash function that maps similar keys to hash values that differ by at most  $m$ , where  $m$  is a small integer (say, 1 or 2). If one builds a table  $T$  of all record numbers, using such a hash function, then similar records will end up in the same bucket, or in nearby buckets. Then one need only check the records in each bucket  $T[i]$  against those in buckets  $T[i+k]$  where  $k$  ranges between  $-m$  and  $m$ .

This class includes the so-called acoustic fingerprint algorithms, that are used to locate similar-sounding entries in large collection of audio files. For this application, the hash function must be as insensitive as possible to data capture or transmission errors, and to "trivial" changes such as timing and volume changes, compression, etc.<sup>[2]</sup>

## Finding similar substrings

The same techniques can be used to find equal or similar stretches in a large collection of strings, such as a document repository or a genomic database. In this case, the input strings are broken into many small pieces, and a hash function is used to detect potentially equal pieces, as above.

The Rabin–Karp algorithm is a relatively fast string searching algorithm that works in  $O(n)$  time on average. It is based on the use of hashing to compare strings.

## Geometric hashing

This principle is widely used in computer graphics, computational geometry and many other disciplines, to solve many proximity problems in the plane or in three-dimensional space, such as finding closest pairs in a set of points, similar shapes in a list of shapes, similar images in an image database, and so on. In these applications, the set of all inputs is some sort of metric space, and the hashing function can be interpreted as a partition of that space into a grid of *cells*. The table is often an array with two or more indices (called a *grid file*, *grid index*, *bucket grid*, and similar names), and the hash function returns an index tuple. This special case of hashing is known as geometric hashing or *the grid method*. Geometric hashing is also used in telecommunications (usually under the name vector quantization) to encode and compress multi-dimensional signals.



## Properties

Good hash functions, in the original sense of the term, are usually required to satisfy certain properties listed below. Note that different requirements apply to the other related concepts (cryptographic hash functions, checksums, etc.).

## Determinism

A hash procedure must be deterministic—meaning that for a given input value it must always generate the same hash value. In other words, it must be a function of the data to be hashed, in the mathematical sense of the term. This requirement excludes hash functions that depend on external variable parameters, such as pseudo-random number generators or the time of day. It also excludes functions that depend on the memory address of the object being hashed, because that address may change during execution (as may happen on systems that use certain methods of garbage collection), although sometimes rehashing of the item is possible.

## Uniformity

A good hash function should map the expected inputs as evenly as possible over its output range. That is, every hash value in the output range should be generated with roughly the same probability. The reason for this last requirement is that the cost of hashing-based methods goes up sharply as the number of *collisions*—pairs of inputs that are mapped to the same hash value—increases. Basically, if some hash values are more likely to occur than others, a larger fraction of the lookup operations will have to search through a larger set of colliding table entries.

Note that this criterion only requires the value to be *uniformly distributed*, not *random* in any sense. A good randomizing function is (barring computational efficiency concerns) generally a good choice as a hash function, but the converse need not be true.

Hash tables often contain only a small subset of the valid inputs. For instance, a club membership list may contain only a hundred or so member names, out of the very large set of all possible names. In these cases, the uniformity criterion should hold for almost all typical subsets of entries that may be found in the table, not just for the global set of all possible entries.

In other words, if a typical set of  $m$  records is hashed to  $n$  table slots, the probability of a bucket receiving many more than  $m/n$  records should be vanishingly small. In particular, if  $m$  is less than  $n$ , very few buckets should have more than one or two records. (In an ideal "perfect hash function", no bucket should have more than one record; but a small number of collisions is virtually inevitable, even if  $n$  is much larger than  $m$ —see the birthday paradox).

When testing a hash function, the uniformity of the distribution of hash values can be evaluated by the chi-squared test.

## Variable range

In many applications, the range of hash values may be different for each run of the program, or may change along the same run (for instance, when a hash table needs to be expanded). In those situations, one needs a hash function which takes two parameters—the input data  $z$ , and the number  $n$  of allowed hash values.

A common solution is to compute a fixed hash function with a very large range (say, 0 to  $2^{32} - 1$ ), divide the result by  $n$ , and use the division's remainder. If  $n$  is itself a power of 2, this can be done by bit masking and bit shifting. When this approach is used, the hash function must be chosen so that the result has fairly uniform distribution between 0 and  $n - 1$ , for any value of  $n$  that may occur in the application. Depending on the function, the remainder may be uniform only for certain values of  $n$ , e.g. odd or prime numbers.

We can allow the table size  $n$  to not be a power of 2 and still not have to perform any remainder or division operation, as these computations are sometimes costly. For example, let  $n$  be significantly less than  $2^b$ . Consider a pseudo random number generator (PRNG) function  $P(\text{key})$  that is uniform on the interval  $[0, 2^b - 1]$ . A hash function uniform on the interval  $[0, n-1]$  is  $n P(\text{key})/2^b$ . We can replace the division by a (possibly faster) right bit

shift:  $nP(\text{key}) \gg b$ .

### Variable range with minimal movement (dynamic hash function)

When the hash function is used to store values in a hash table that outlives the run of the program, and the hash table needs to be expanded or shrunk, the hash table is referred to as a dynamic hash table.

A hash function that will relocate the minimum number of records when the table is resized is desirable. What is needed is a hash function  $H(z, n)$  – where  $z$  is the key being hashed and  $n$  is the number of allowed hash values – such that  $H(z, n + 1) = H(z, n)$  with probability close to  $n/(n + 1)$ .

Linear hashing and spiral storage are examples of dynamic hash functions that execute in constant time but relax the property of uniformity to achieve the minimal movement property.

Extendible hashing uses a dynamic hash function that requires space proportional to  $n$  to compute the hash function, and it becomes a function of the previous keys that have been inserted.

Several algorithms that preserve the uniformity property but require time proportional to  $n$  to compute the value of  $H(z, n)$  have been invented.

### Data normalization

In some applications, the input data may contain features that are irrelevant for comparison purposes. For example, when looking up a personal name, it may be desirable to ignore the distinction between upper and lower case letters. For such data, one must use a hash function that is compatible with the data equivalence criterion being used: that is, any two inputs that are considered equivalent must yield the same hash value. This can be accomplished by normalizing the input before hashing it, as by upper-casing all letters.

### Continuity

A hash function that is used to search for similar (as opposed to equivalent) data must be as continuous as possible; two inputs that differ by a little should be mapped to equal or nearly equal hash values.

Note that continuity is usually considered a fatal flaw for checksums, cryptographic hash functions, and other related concepts. Continuity is desirable for hash functions only in some applications, such as hash tables that use linear search.

## Hash function algorithms

For most types of hashing functions the choice of the function depends strongly on the nature of the input data, and their probability distribution in the intended application.

### Trivial hash function

If the datum to be hashed is small enough, one can use the datum itself (reinterpreted as an integer in binary notation) as the hashed value. The cost of computing this "trivial" (identity) hash function is effectively zero. This hash function is perfect, as it maps each input to a distinct hash value.

The meaning of "small enough" depends on the size of the type that is used as the hashed value. For example, in Java, the hash code is a 32-bit integer. Thus the 32-bit integer `Integer` and 32-bit floating-point `Float` objects can simply use the value directly; whereas the 64-bit integer `Long` and 64-bit floating-point `Double` cannot use this method.

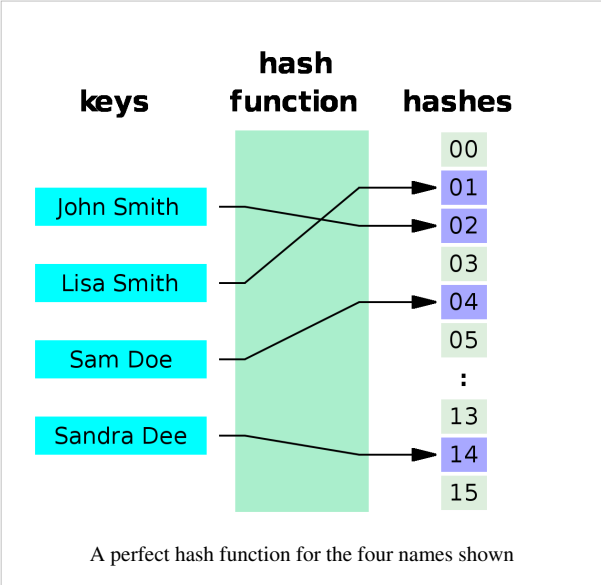
Other types of data can also use this perfect hashing scheme. For example, when mapping character strings between upper and lower case, one can use the binary encoding of each character, interpreted as an integer, to index a table that gives the alternative form of that character ("A" for "a", "8" for "8", etc.). If each character is stored in 8 bits (as

in ASCII or ISO Latin 1), the table has only  $2^8 = 256$  entries; in the case of Unicode characters, the table would have  $17 \times 2^{16} = 1114112$  entries.

The same technique can be used to map two-letter country codes like "us" or "za" to country names ( $26^2=676$  table entries), 5-digit zip codes like 13083 to city names (100000 entries), etc. Invalid data values (such as the country code "xx" or the zip code 00000) may be left undefined in the table, or mapped to some appropriate "null" value.

Perfect hashing

A hash function that is injective—that is, maps each valid input to a different hash value—is said to be **perfect**. With such a function one can directly locate the desired entry in a hash table, without any additional searching.



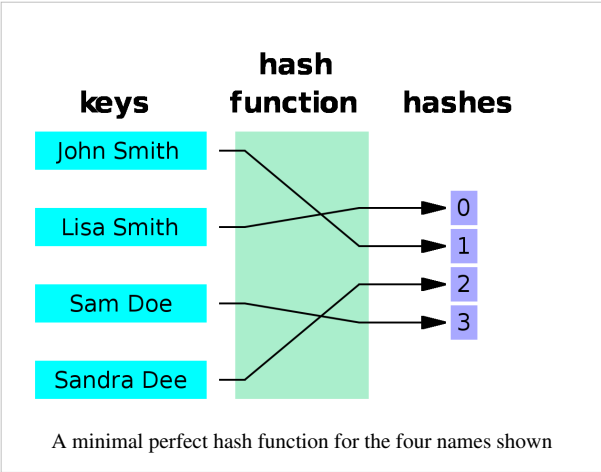
Minimal perfect hashing

A perfect hash function for  $n$  keys is said to be **minimal** if its range consists of  $n$  consecutive integers, usually from 0 to  $n-1$ . Besides providing single-step lookup, a minimal perfect hash function also yields a compact hash table, without any vacant slots. Minimal perfect hash functions are much harder to find than perfect ones with a wider range.

Hashing uniformly distributed data

If the inputs are bounded-length strings (such as telephone numbers, car license plates, invoice numbers, etc.), and each input may independently occur with uniform probability, then a hash function need only map roughly the same number of inputs to each hash value. For instance, suppose that each input is an integer  $z$  in the range 0 to  $N-1$ , and the output must be an integer  $h$  in the range 0 to  $n-1$ , where  $N$  is much larger than  $n$ . Then the hash function could be  $h = z \bmod n$  (the remainder of  $z$  divided by  $n$ ), or  $h = (z \times n) \div N$  (the value  $z$  scaled down by  $n/N$  and truncated to an integer), or many other formulas.

Warning:  $h = z \bmod n$  was used in many of the original random number generators, but was found to have a number of issues. One of which is that as  $n$  approaches  $N$ , this function becomes less and less uniform.



## Hashing data with other distributions

These simple formulas will not do if the input values are not equally likely, or are not independent. For instance, most patrons of a supermarket will live in the same geographic area, so their telephone numbers are likely to begin with the same 3 to 4 digits. In that case, if  $n$  is 10000 or so, the division formula  $(z \times n) \div N$ , which depends mainly on the leading digits, will generate a lot of collisions; whereas the remainder formula  $z \bmod n$ , which is quite sensitive to the trailing digits, may still yield a fairly even distribution.

## Hashing variable-length data

When the data values are long (or variable-length) character strings—such as personal names, web page addresses, or mail messages—their distribution is usually very uneven, with complicated dependencies. For example, text in any natural language has highly non-uniform distributions of characters, and character pairs, very characteristic of the language. For such data, it is prudent to use a hash function that depends on all characters of the string—and depends on each character in a different way.

In cryptographic hash functions, a Merkle–Damgård construction is usually used. In general, the scheme for hashing such data is to break the input into a sequence of small units (bits, bytes, words, etc.) and combine all the units  $b[1]$ ,  $b[2]$ , ...,  $b[m]$  sequentially, as follows

```
S ← S0;                                // Initialize the state.
for k in 1, 2, ..., m do                // Scan the input data units:
    S ← F(S, b[k]);                     // Combine data unit k into the state.

return G(S, n) // Extract the hash value from the state.
```

This schema is also used in many text checksum and fingerprint algorithms. The state variable  $S$  may be a 32- or 64-bit unsigned integer; in that case,  $S0$  can be 0, and  $G(S,n)$  can be just  $S \bmod n$ . The best choice of  $F$  is a complex issue and depends on the nature of the data. If the units  $b[k]$  are single bits, then  $F(S,b)$  could be, for instance

```
if highbit(S) = 0 then
    return 2 * S + b
else
```

```
return (2 * S + b) ^ P
```

Here  $highbit(S)$  denotes the most significant bit of  $S$ ; the '\*' operator denotes unsigned integer multiplication with lost overflow; '^' is the bitwise exclusive or operation applied to words; and  $P$  is a suitable fixed word.<sup>[3]</sup>

## Special-purpose hash functions

In many cases, one can design a special-purpose (heuristic) hash function that yields many fewer collisions than a good general-purpose hash function. For example, suppose that the input data are file names such as `FILE0000.CHK`, `FILE0001.CHK`, `FILE0002.CHK`, etc., with mostly sequential numbers. For such data, a function that extracts the numeric part  $k$  of the file name and returns  $k \bmod n$  would be nearly optimal. Needless to say, a function that is exceptionally good for a specific kind of data may have dismal performance on data with different distribution.

## Rolling hash

In some applications, such as substring search, one must compute a hash function  $h$  for every  $k$ -character substring of a given  $n$ -character string  $t$ ; where  $k$  is a fixed integer, and  $n$  is  $k$ . The straightforward solution, which is to extract every such substring  $s$  of  $t$  and compute  $h(s)$  separately, requires a number of operations proportional to  $k \cdot n$ . However, with the proper choice of  $h$ , one can use the technique of rolling hash to compute all those hashes with an effort proportional to  $k + n$ .

## Universal hashing

A universal hashing scheme is a randomized algorithm that selects a hashing function  $h$  among a family of such functions, in such a way that the probability of a collision of any two distinct keys is  $1/n$ , where  $n$  is the number of distinct hash values desired—independently of the two keys. Universal hashing ensures (in a probabilistic sense) that the hash function application will behave as well as if it were using a random function, for any distribution of the input data. It will however have more collisions than perfect hashing, and may require more operations than a special-purpose hash function.

## Hashing with checksum functions

One can adapt certain checksum or fingerprinting algorithms for use as hash functions. Some of those algorithms will map arbitrary long string data  $z$ , with any typical real-world distribution—no matter how non-uniform and dependent—to a 32-bit or 64-bit string, from which one can extract a hash value in 0 through  $n - 1$ .

This method may produce a sufficiently uniform distribution of hash values, as long as the hash range size  $n$  is small compared to the range of the checksum or fingerprint function. However, some checksums fare poorly in the avalanche test, which may be a concern in some applications. In particular, the popular CRC32 checksum provides only 16 bits (the higher half of the result) that are usable for hashing. Moreover, each bit of the input has a deterministic effect on each bit of the CRC32, that is one can tell without looking at the rest of the input, which bits of the output will flip if the input bit is flipped; so care must be taken to use all 32 bits when computing the hash from the checksum.<sup>[4]</sup>

## Hashing with cryptographic hash functions

Some cryptographic hash functions, such as SHA-1, have even stronger uniformity guarantees than checksums or fingerprints, and thus can provide very good general-purpose hashing functions.

In ordinary applications, this advantage may be too small to offset their much higher cost.<sup>[5]</sup> However, this method can provide uniformly distributed hashes even when the keys are chosen by a malicious agent. This feature may help protect services against denial of service attacks.

## Hashing By Nonlinear Table Lookup

Tables of random numbers (such as 256 random 32 bit integers) can provide high-quality non-linear functions to be used as hash functions or other purposes such as cryptography. The key to be hashed would be split in 8-bit (one byte) parts and each part will be used as an index for the non-linear table. The table values will be added by arithmetic or XOR addition to the hash output value. Because the table is just 1024 bytes in size, it will fit into the cache of modern microprocessors and allow for very fast execution of the hashing algorithm. As the table value is on average much longer than 8 bit, one bit of input will affect nearly all output bits. This is different to multiplicative hash functions where higher-value input bits do not affect lower-value output bits.

This algorithm has proven to be very fast and of high quality for hashing purposes (especially hashing of integer number keys).

## Efficient Hashing Of Strings

Modern microprocessors will allow for much faster processing, if 8-bit character Strings are not hashed by processing one character at a time, but by interpreting the string as an array of 32 bit or 64 bit integers and hashing/accumulating these "wide word" integer values by means of arithmetic operations (e.g. multiplication by constant and bit-shifting). The remaining characters of the string which are smaller than the word length of the CPU must be handled differently (e.g. being processed one character at a time).

This approach has proven to speed up hash code generation by a factor of five or more on modern microprocessors of a word size of 64 bit.

## Origins of the term

The term "hash" comes by way of analogy with its non-technical meaning, to "chop and mix". Indeed, typical hash functions, like the **mod** operation, "chop" the input domain into many sub-domains that get "mixed" into the output range to improve the uniformity of the key distribution.

Donald Knuth notes that Hans Peter Luhn of IBM appears to have been the first to use the concept, in a memo dated January 1953, and that Robert Morris used the term in a survey paper in CACM which elevated the term from technical jargon to formal terminology.<sup>[1]</sup>

## List of hash functions

- Bernstein hash<sup>[6]</sup>
- Fowler-Noll-Vo hash function (32, 64, 128, 256, 512, or 1024 bits)
- Jenkins hash function (32 bits)
- Pearson hashing (8 bits)
- Zobrist hashing

## References

- [1] Knuth, Donald (1973). *The Art of Computer Programming, volume 3, Sorting and Searching*. pp. 506–542.
- [2] "Robust Audio Hashing for Content Identification by Jaap Haitsma, Ton Kalker and Job Oostveen" (<http://citeseer.ist.psu.edu/rd/11787382,504088,1,0.25,Download/http://citeseer.ist.psu.edu/cache/papers/cs/25861/http:zSzzSzwww.extra.research.philips.comzSznatlabzSzdownloadzSzaudiofzpzSzcmbi01audiohashv1.0.pdf/haitsma01robust.pdf>)
- [3] Broder, A. Z. (1993). "Some applications of Rabin's fingerprinting method". *Sequences II: Methods in Communications, Security, and Computer Science*. Springer-Verlag. pp. 143–152.
- [4] Bret Mulvey, *Evaluation of CRC32 for Hash Tables* (<http://home.comcast.net/~bretm/hash/8.html>), in *Hash Functions* (<http://home.comcast.net/~bretm/hash/>). Accessed April 10, 2009.
- [5] Bret Mulvey, *Evaluation of SHA-1 for Hash Tables* (<http://home.comcast.net/~bretm/hash/9.html>), in *Hash Functions* (<http://home.comcast.net/~bretm/hash/>). Accessed April 10, 2009.
- [6] <http://www.cse.yorku.ca/~oz/hash.html>

## External links

- General purpose hash function algorithms (C/C++/Pascal/Java/Python/Ruby) (<http://www.partow.net/programming/hashfunctions/index.html>)
- Hash Functions and Block Ciphers by Bob Jenkins (<http://burtleburtle.net/bob/hash/index.html>)
- The Goulburn Hashing Function (<http://www.webcitation.org/query?url=http://www.geocities.com/drone115b/Goulburn06.pdf&date=2009-10-25+21:06:51>) (PDF) by Mayur Patel
- MIT's Introduction to Algorithms: Hashing 1 (<http://video.google.com/videoplay?docid=-727485696209877198&q=source:014117792397255896270&hl=en>) MIT OCW lecture Video
- MIT's Introduction to Algorithms: Hashing 2 (<http://video.google.com/videoplay?docid=2307261494964091254&q=source:014117792397255896270&hl=en>) MIT OCW lecture Video
- Hash Fuction Construction for Textual and Geometrical Data Retrieval ([http://herakles.zcu.cz/~skala/PUBL/PUBL\\_2010/2010\\_WSEAS-Corfu\\_Hash-final.pdf](http://herakles.zcu.cz/~skala/PUBL/PUBL_2010/2010_WSEAS-Corfu_Hash-final.pdf)) Latest Trends on Computers, Vol.2, pp. 483–489, CSCC conference, Corfu, 2010

# Open addressing

**Open addressing**, or **closed hashing**, is a method of collision resolution in hash tables. With this method a hash collision is resolved by **probing**, or searching through alternate locations in the array (the *probe sequence*) until either the target record is found, or an unused array slot is found, which indicates that there is no such key in the table.<sup>[1]</sup> Well known probe sequences include:

## Linear probing

in which the interval between probes is fixed — often at 1.

## Quadratic probing

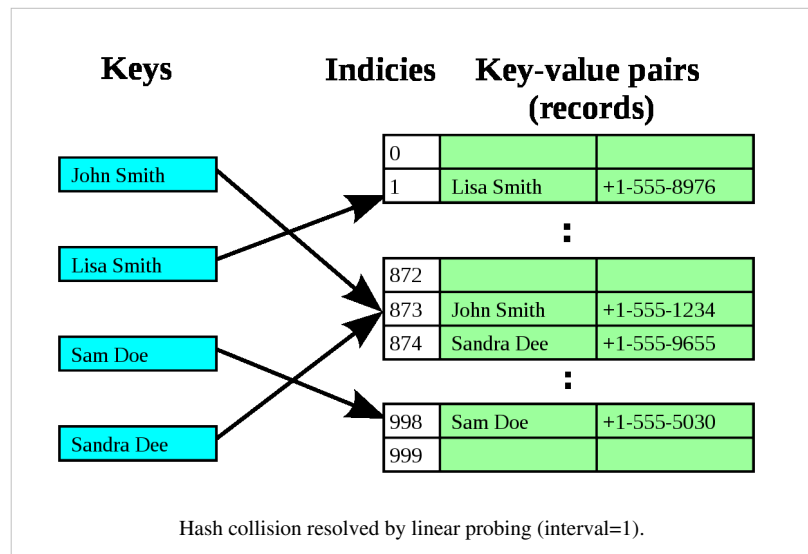
in which the interval between probes increases linearly (hence, the indices are described by a quadratic function).

## Double hashing

in which the interval between probes is fixed for each record but is computed by another hash function.

The main tradeoffs between these methods are that linear probing has the best cache performance but is most sensitive to clustering, while double hashing has poor cache performance but exhibits virtually no clustering; quadratic probing falls in-between in both areas. Double hashing can also require more computation than other forms of probing. Some open addressing methods, such as last-come-first-served hashing and cuckoo hashing move existing keys around in the array to make room for the new key. This gives better maximum search times than the methods based on probing.

A critical influence on performance of an open addressing hash table is the *load factor*; that is, the proportion of the slots in the array that are used. As the load factor increases towards 100%, the number of probes that may be required to find or insert a given key rises dramatically. Once the table becomes full, probing algorithms may even fail to terminate. Even with good hash functions, load factors are normally limited to 80%. A poor hash function can exhibit poor performance even at very low load factors by generating significant clustering. What causes hash functions to cluster is not well understood, and it is easy to unintentionally write a hash function which causes severe clustering.



## Example pseudo code

The following pseudocode is an implementation of an open addressing hash table with linear probing and single-slot stepping, a common approach that is effective if the hash function is good. Each of the **lookup**, **set** and **remove** functions use a common internal function **find\_slot** to locate the array slot that either does or should contain a given key.

```
record pair { key, value }
var pair array slot[0..num_slots-1]
```

```
function find_slot(key)
    i := hash(key) modulo num_slots
    // search until we either find the key, or find an empty slot.
    while ( (slot[i] is occupied) and ( slot[i].key ≠ key ) ) do
        i := (i + 1) modulo num_slots
    repeat
    return i
```

```
function lookup(key)
    i := find_slot(key)
    if slot[i] is occupied    // key is in table
        return slot[i].value
    else                     // key is not in table
        return not found
```

```
function set(key, value)
    i := find_slot(key)
    if slot[i] is occupied
        slot[i].value := value
    else
        if the table is almost full
            rebuild the table larger (note 1)
            i := find_slot(key)
        slot[i].key    := key
        slot[i].value := value
```

Another example showing open addressing technique. Presented function is converting each part(4) of an internet protocol address, where NOT, XOR, OR and AND are bitwise operations and << and >> are left and right logical shifts:

```
// key_1,key_2,key_3,key_4 are following 3-digit numbers - parts of ip address xxx.xxx.xxx.xxx
function ip(key parts)
    j := 1
    do
        key := (key_2 << 2)
        key := (key + (key_3 << 7))
        key := key + (j OR key_4 >> 2) * (key_4) * (j + key_1) XOR j
        key := key AND _prime_    // _prime_ is a prime number
        j := (j+1)
    while collision
    return key
```



## note 1

Rebuilding the table requires allocating a larger array and recursively using the **set** operation to insert all the elements of the old array into the new larger array. It is common to increase the array size exponentially, for example by doubling the old array size.

```
function remove(key)
    i := find_slot(key)
    if slot[i] is unoccupied
        return // key is not in the table
    j := i
    loop
        mark slot[i] as unoccupied
    r2: (note 2)
        j := (j+1) modulo num_slots
        if slot[j] is unoccupied
            exit loop
        k := hash(slot[j].key) modulo num_slots
        // k lies cyclically in ]i,j]
        // |      i.k.j |
        // |....j i.k.| or |.k..j i....|
        if ( (i<=j) ? ((i<k)&&(k<=j)) : ((i<k)|| (k<=j)) )
            goto r2;
        slot[i] := slot[j]
        i := j
```

## note 2

For all records in a cluster, there must be no vacant slots between their natural hash position and their current position (else lookups will terminate before finding the record). At this point in the pseudocode, *i* is a vacant slot that might be invalidating this property for subsequent records in the cluster. *j* is such a subsequent record. *k* is the raw hash where the record at *j* would naturally land in the hash table if there were no collisions. This test is asking if the record at *j* is invalidly positioned with respect to the required properties of a cluster now that *i* is vacant.

Another technique for removal is simply to mark the slot as deleted. However this eventually requires rebuilding the table simply to remove deleted records. The methods above provide  $O(1)$  updating and removal of existing records, with occasional rebuilding if the high water mark of the table size grows.

The  $O(1)$  remove method above is only possible in linearly probed hash tables with single-slot stepping. In the case where many records are to be deleted in one operation, marking the slots for deletion and later rebuilding may be more efficient.

## References

- [1] Tenenbaum, Aaron M.; Langsam, Yedidyah; Augenstein, Moshe J. (1990), *Data Structures Using C*, Prentice Hall, pp. 456–461, pp. 472, ISBN 0-13-199746-7

# Lazy deletion

In computer science, **lazy deletion** refers to a method of deleting elements from a hash table that uses open addressing. In this method, deletions are done by marking an element as deleted, rather than erasing it entirely. Deleted locations are treated as empty when inserting and as occupied during a search.

The problem with this scheme is that as the number of delete/insert operations increases the cost of a successful search increases. To improve this, when an element is searched and found in the table, the element is relocated to the first location marked for deletion that was probed during the search. Instead of finding an element to relocate when the deletion occurs, the relocation occurs lazily during the next search.<sup>[1][2]</sup>

## References

- [1] Celis, Pedro; Franco, John (1995), *The Analysis of Hashing with Lazy Deletions* (<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.39.9637>), Computer Science Department, Indiana University, Technical Report CS-86-14,
- [2] Celis, Pedro; Franco, John (1992), "The analysis of hashing with lazy deletions", *Information Sciences* **62**: 13, doi:10.1016/0020-0255(92)90022-Z

# Linear probing

**Linear probing** is a scheme in computer programming for resolving hash collisions of values of hash functions by sequentially searching the hash table for a free location.<sup>[1]</sup> This is accomplished using two values - one as a starting value and one as an interval between successive values in modular arithmetic. The second value, which is the same for all keys and known as the *stepsize*, is repeatedly added to the starting value until a free space is found, or the entire table is traversed. (In order to traverse the entire table the stepsize should be relatively prime to the arraysize, which is why the array size is often chosen to be a prime number.)

$$\text{newLocation} = (\text{startingValue} + \text{stepSize}) \% \text{arraySize}$$

This algorithm, which is used in open-addressed hash tables, provides good memory caching (if stepsize is equal to one), through good locality of reference, but also results in clustering, an unfortunately high probability that where there has been one collision there will be more. The performance of linear probing is also more sensitive to input distribution when compared to double hashing, where the stepsize is determined by another hash function applied to the value instead of a fixed stepsize as in linear probing.

Given an ordinary hash function  $H(x)$ , a linear probing function ( $H(x, i)$ ) would be:

$$H(x, i) = (H(x) + i) \pmod{n}.$$

Here  $H(x)$  is the starting value,  $n$  the size of the hash table, and the *stepsize* is  $i$  in this case.

## Dictionary operation in constant time

Using linear probing, dictionary operation can be implemented in constant time. In other words, insert, remove and find operations can be implemented in  $O(1)$ , as long as the load factor of the hash table is a constant strictly less than one.<sup>[2]</sup> This analysis makes the (unrealistic) assumption that the hash function is completely random, but can be extended also to 5-independent hash functions.<sup>[3]</sup> Weaker properties, such as universal hashing, are not strong enough to ensure the constant-time operation of linear probing,<sup>[4]</sup> but one practical method of hash function generation, tabulation hashing, again leads to a guaranteed constant expected time performance despite not being 5-independent.<sup>[5]</sup>

## References

- [1] Dale, Nell (2003). *C++ Plus Data Structures*. Sudbury, MA: Jones and Bartlett Computer Science. ISBN 0-7637-0481-4.
- [2] Knuth, Donald (1963), *Notes on "Open" Addressing* (<http://algo.inria.fr/AofA/Research/11-97.html>),
- [3] Pagh, Anna; Pagh, Rasmus; Ružić, Milan (2009), "Linear probing with constant independence", *SIAM Journal on Computing* **39** (3): 1107–1120, doi:10.1137/070702278, MR2538852
- [4] Pătraşcu, Mihai; Thorup, Mikkel (2010), "On the k-independence required by linear probing and minwise independence" (<http://people.csail.mit.edu/mip/papers/kwise-lb/kwise-lb.pdf>), *Automata, Languages and Programming, 37th International Colloquium, ICALP 2010, Bordeaux, France, July 6-10, 2010, Proceedings, Part I*, Lecture Notes in Computer Science, **6198**, Springer, pp. 715–726, doi:10.1007/978-3-642-14165-2\_60,
- [5] Pătraşcu, Mihai; Thorup, Mikkel (2011), "The power of simple tabulation hashing", *Proceedings of the 43rd annual ACM Symposium on Theory of Computing (STOC '11)*, pp. 1–10, arXiv:1011.5200, doi:10.1145/1993636.1993638

## External links

- How Caching Affects Hashing (<http://www.siam.org/meetings/alnex05/papers/13gheileman.pdf>) by Gregory L. Heileman and Wenbin Luo 2005.
- Open Data Structures - Section 5.2 - LinearHashTable: Linear Probing ([http://opendatastructures.org/versions/edition-0.1e/ods-java/5\\_2\\_LinearHashTable\\_Linear\\_.html](http://opendatastructures.org/versions/edition-0.1e/ods-java/5_2_LinearHashTable_Linear_.html))

# Quadratic probing

---

**Quadratic probing** is an open addressing scheme in computer programming for resolving collisions in hash tables -- when an incoming data's hash value indicates it should be stored in an already-occupied slot or bucket. Quadratic probing operates by taking the original hash index and adding successive values of an arbitrary quadratic polynomial until an open slot is found.

For a given hash value, the indices generated by linear probing are as follows:

$$H + 1, H + 2, H + 3, H + 4, \dots, H + k$$

This method results in primary clustering, and as the cluster grows larger, the search for those items hashing within the cluster becomes less efficient.

An example sequence using quadratic probing is:

$$H + 1^2, H + 2^2, H + 3^2, H + 4^2, \dots, H + k^2$$

Quadratic probing can be a more efficient algorithm in a closed hash table, since it better avoids the clustering problem that can occur with linear probing, although it is not immune. It also provides good memory caching because it preserves some locality of reference; however, linear probing has greater locality and, thus, better cache performance.

Quadratic probing is used in the Berkeley Fast File System to allocate free blocks. The allocation routine chooses a new cylinder-group when the current is nearly full using quadratic probing, because of the speed it shows in finding unused cylinder-groups.

## Quadratic Function

Let  $h(k)$  be a hash function that maps an element  $k$  to an integer in  $[0, m-1]$ , where  $m$  is the size of the table. Let the  $i^{\text{th}}$  probe position for a value  $k$  be given by the function

$$h(k, i) = (h(k) + c_1 i + c_2 i^2) \pmod{m}$$

where  $c_2 \neq 0$ . If  $c_2 = 0$ , then  $h(k, i)$  degrades to a linear probe. For a given hash table, the values of  $c_1$  and  $c_2$  remain constant.

### Examples:

- If  $h(k, i) = (h(k) + i + i^2) \pmod{m}$ , then the probe sequence will be  $h(k), h(k) + 2, h(k) + 6, \dots$
- For  $m = 2^n$ , a good choice for the constants are  $c_1 = c_2 = 1/2$ , as the values of  $h(k, i)$  for  $i$  in  $[0, m-1]$  are all distinct. This leads to a probe sequence of  $h(k), h(k) + 1, h(k) + 3, h(k) + 6, \dots$  where the values increase by 1, 2, 3, ...
- For prime  $m > 2$ , most choices of  $c_1$  and  $c_2$  will make  $h(k, i)$  distinct for  $i$  in  $[0, (m-1)/2]$ . Such choices include  $c_1 = c_2 = 1/2$ ,  $c_1 = c_2 = 1$ , and  $c_1 = 0, c_2 = 1$ . Because there are only about  $m/2$  distinct probes for a given element, it is difficult to guarantee that insertions will succeed when the load factor is  $> 1/2$ .

## Quadratic Probing Insertion

The problem, here, is to insert a key at an available key space in a given Hash Table using quadratic probing.<sup>[1]</sup>

### Algorithm to Insert key in Hash Table

```

1. Get the key k
2. Set counter j = 0
3. Compute hash function h[k] = k % SIZE
4. If hashtable[h[k]] is empty
    (4.1) Insert key k at hashtable[h[k]]
    (4.2) Stop
Else
    (4.3) The key space at hashtable[h[k]] is occupied, so we need to find the next available key space
    (4.4) Increment j
    (4.5) Compute new hash function h[k] = ( k + j * j ) % SIZE
    (4.6) Repeat Step 4 till j is more than SIZE of hash table
5. The hash table is full
6. Stop

```

### C function for Key Insertion

```

int quadratic_probing_insert(int *hashtable, int key, int *empty)
{
    /* hashtable[] is an integer hash table; empty[] is another array
    which indicates whether the key space is occupied;
    If an empty key space is found, the function returns the index
    of the bucket where the key is inserted, otherwise it
    returns (-1) if no empty key space is found */

    int j = 0, hk;
    hk = key % SIZE;

```

```

while(j < SIZE)
{
    if(empty[hk] == 1)
    {
        hashtable[hk] = key;
        empty[hk] = 0;
        return (hk);
    }
    j++;
    hk = (key + j * j) % SIZE;
}
return (-1);
}

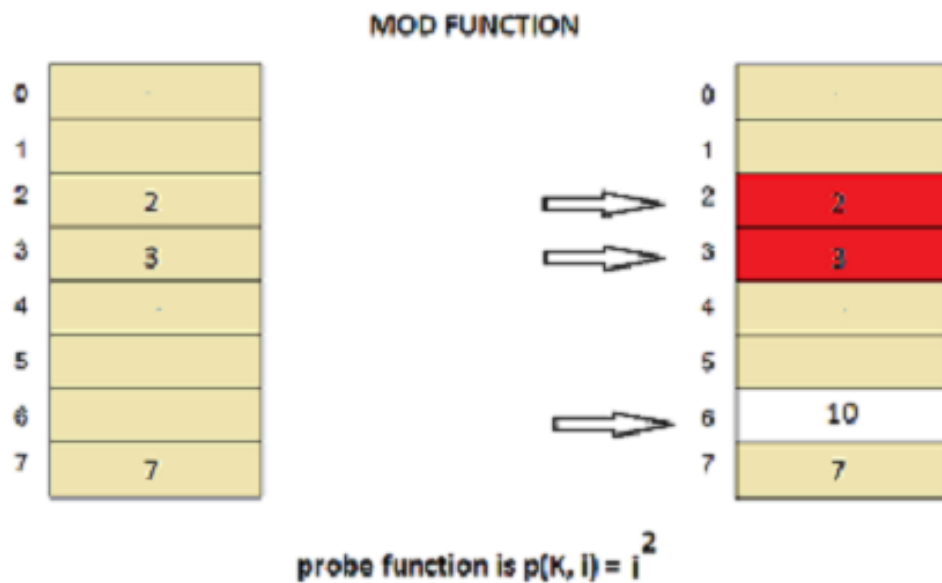
```

### Example to Insert key in Hash Table

There are two possible cases to consider:

- **Key space at position  $h[k]$  is empty :** Insert the key at the position.
- **Key space at position  $h[k]$  is occupied:** Compute the next hash function  $h[k]$ .

Consider a hash table initially containing some elements.



Suppose we want to insert a key 10 in the hash table.

$$h[k] = 10 \% 8 = 2$$

Slot 2 being occupied the hash function will search for new available key space.

$$h[k] = (k + j * j) \% SIZE$$

$$h[k] = (2 + 1 * 1) \% 8 = 3$$

Slot 3 is also occupied, so the hash function will search for next available key space.

$$h[k] = (2 + 2 * 2) \% 8 = 6$$

Slot 6 is empty, so key will be inserted here.

## Quadratic Probing Search

### Algorithm to Search Element in Hash Table

```

1. Get the key k to be searched
2. Set counter j = 0
3. Compute hash function h[k] = k % SIZE
4. If the key space at hashtable[h[k]] is occupied
    (4.1) Compare the element at hashtable[h[k]] with the key k.
    (4.2) If they are equal
        (4.2.1) The key is found at the bucket h[k]
        (4.2.2) Stop
    Else
        (4.3) The element might be placed at the next location given by the quadratic function
        (4.4) Increment j
        (4.5) Compute new hash function h[k] = ( k + j * j ) % SIZE
        (4.6) Repeat Step 4 till j is greater than SIZE of hash table
5. The key was not found in the hash table
6. Stop

```

### C function for Key Searching

```

int quadratic_probing_search(int *hashtable, int key, int *empty)
{
    /* If the key is found in the hash table, the function returns the
    index of the hashtable where the key is inserted, otherwise it
    returns (-1) if the key is not found */

    int j = 0, hk;
    hk = key % SIZE;
    while(j < SIZE)
    {
        if((empty[hk] == 0) && (hashtable[hk] == key))
            return (hk);
        j++;
        hk = (key + j * j) % SIZE;
    }
    return (-1);
}

```

### Limitations

<sup>[2]</sup> For linear probing it is a bad idea to let the hash table get nearly full, because performance is degraded as the hash table gets filled. In the case of quadratic probing, the situation is even more drastic. With the exception of the triangular number case for a power-of-two-sized hash table, there is no guarantee of finding an empty cell once the table gets more than half full, or even before the table gets half full if the table size is not prime. This is because at most half of the table can be used as alternative locations to resolve collisions. If the hash table size is  $b$  (a prime greater than 3), it can be proven that the first  $b/2$  alternative locations including the initial location  $h(k)$  are all distinct and unique. Suppose, we assume two of the alternative locations to be given by  $h(k) + x^2 \pmod{b}$

and  $h(k) + y^2 \pmod{b}$ , where  $0 \leq x, y \leq (b/2)$ . If these two locations point to the same key space, but  $x \neq y$ . Then the following have to be true,

$$\begin{aligned} h(k) + x^2 &= h(k) + y^2 \pmod{b} \\ x^2 &= y^2 \pmod{b} \\ x^2 - y^2 &= 0 \pmod{b} \\ (x - y)(x + y) &= 0 \pmod{b} \end{aligned}$$

As  $b$  (table size) is a prime greater than 3, either  $(x - y)$  or  $(x + y)$  has to be equal to zero. Since  $x$  and  $y$  are unique,  $(x - y)$  cannot be zero. Also, since  $0 \leq x, y \leq (b/2)$ ,  $(x + y)$  cannot be zero.

Thus, by contradiction, it can be said that the first  $(b/2)$  alternative locations after  $h(k)$  are unique. So an empty key space can always be found as long as at most  $(b/2)$  locations are filled, i.e., the hash table is not more than half full.

## References

- [1] Horowitz, Sahni, Anderson-Freed (2011). *Fundamentals of Data Structures in C*. University Press. ISBN 978-81-7371-605-8.
- [2] *Data Structures and Algorithm Analysis in C++*. Pearson Education. 2009. ISBN 978-81-317-1474-4.

## External links

- Tutorial/quadratic probing (<http://research.cs.vt.edu/AVresearch/hashing/quadratic.php>)

# Double hashing

**Double hashing** is a computer programming technique used in hash tables to resolve hash collisions, cases when two different values to be searched for produce the same hash key. It is a popular collision-resolution technique in open-addressed hash tables. Double hashing is implemented in many popular computer libraries.

## Classical applied data structure

Double hashing with open addressing is a classical data structure on a table  $T$ . Let  $n$  be the number of elements stored in  $T$  then  $T$ 's load factor is  $\alpha = \frac{n}{|T|}$ .

Double hashing approximates uniform open address hashing. That is, start by randomly, uniformly and independently selecting two universal hash functions  $h_1$  and  $h_2$  to build a double hashing table  $T$ .

All elements are put in  $T$  by double hashing using  $h_1$  and  $h_2$ . Given a key  $k$ , determining the  $(i + 1)$ -st hash location is computed by:

$$h(i, k) = (h_1(k) + i h_2(k)) \pmod{|T|}.$$

Let  $T$  have fixed load factor  $\alpha : 1 > \alpha > 0$ . Bradford and Katehakis<sup>[1]</sup> showed the expected number of probes for an unsuccessful search in  $T$ , still using these initially chosen hash functions, is  $\frac{1}{1 - \alpha}$  regardless of the distribution of the inputs.

Previous results include: Guibas and Szemerédi<sup>[2]</sup> showed  $\frac{1}{1 - \alpha}$  holds for unsuccessful search for load factors

$\alpha < 0.319$ . Also, Lueker and Molodowitch<sup>[3]</sup> showed this held assuming ideal randomized functions. Schmidt and Siegel<sup>[4]</sup> showed this with more realistic  $k$ -wise independent and uniform functions (for  $k = c \log n$ , and suitable constant  $c$ ).

Like linear probing, it uses one hash value as a starting point and then repeatedly steps forward an interval until the desired value is located, an empty location is reached, or the entire table has been searched; but this interval is decided using a second, independent hash function (hence the name double hashing). Unlike linear probing and quadratic probing, the interval depends on the data, so that even values mapping to the same location have different bucket sequences; this minimizes repeated collisions and the effects of clustering. In other words, given independent hash functions  $h_1$  and  $h_2$ , the  $j$ th location in the bucket sequence for value  $k$  in a hash table  $T$  is:

$$h(k, j) = (h_1(k) + j \cdot h_2(k)) \bmod |T|$$

## Disadvantages

Linear probing and, to a lesser extent, quadratic probing are able to take advantage of the data cache by accessing locations that are close together. Double hashing has larger intervals and is not able to achieve this advantage. To avoid this situation, store your data with the second key as the row, and your first key as the column. Doing this allows you to iterate on the column, thus preventing cache problems. This also prevents the need to rehash the second key.

For instance:

```
pData[hk_2][hk_1]

int hv_1 = Hash(v)
int hv_2 = Hash2(v)

int original_hash = hv_1
while (pData[hv_2][hv_1]) {
    hv_1 = hv_1 + 1
}
```

Like all other forms of open addressing, double hashing becomes linear as the hash table approaches maximum capacity. The only solution to this is to rehash to a larger size.

On top of that, it is possible for the secondary hash function to evaluate to zero. For example, if we choose  $k=5$  with the following function:

$$h_2(k) = 5 - (k \bmod 7)$$

The resulting sequence will always remain at the initial hash value. One possible solution is to change the secondary hash function to:

$$h_2(k) = (k \bmod 7) + 1$$

This ensures that the secondary hash function will always be non zero.

Essentially, Double Hashing is hashing on an already hashed key.



## Notes

- [1] P. G. Bradford and M. Katehakis *A Probabilistic Study on Combinatorial Expanders and Hashing*, SIAM Journal on Computing 2007 (37:1), 83-111. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.91.2647>
- [2] L. Guibas and E. Szemerédi: *The Analysis of Double Hashing*, Journal of Computer and System Sciences, 1978, 16, 226-274.
- [3] G. S. Lueker and M. Molodowitch: *More Analysis of Double Hashing*, Combinatorica, 1993, 13(1), 83-96.
- [4] J. P. Schmidt and A. Siegel: *Double Hashing is Computable and Randomizable with Universal Hash Functions*, manuscript.

## External links

- How Caching Affects Hashing (<http://www.siam.org/meetings/alensex05/papers/13gheileman.pdf>) by Gregory L. Heileman and Wenbin Luo 2005.
- Hash Table Animation (<http://www.cs.pitt.edu/~kirk/cs1501/animations/Hashing.html>)

# Cuckoo hashing

---

**Cuckoo hashing** is a scheme in computer programming for resolving hash collisions of values of hash functions in a table. The name derives from the behavior of some species of cuckoo, where the cuckoo chick pushes the other eggs or young out of the nest when it hatches.

## History

Cuckoo hashing was first described by Rasmus Pagh and Flemming Friche Rodler in 2001.<sup>[1]</sup>

## Theory

The basic idea is to use two hash functions instead of only one. This provides two possible locations in the hash table for each key. In one of the commonly used variants of the algorithm, the hash table is split into two smaller tables of equal size, and each hash function provides an index into one of these two tables.

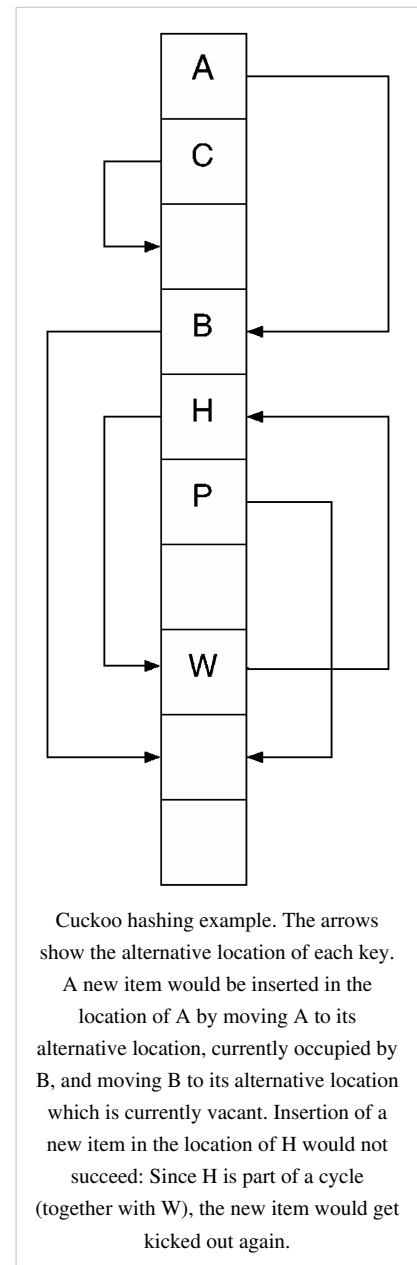
When a new key is inserted, a greedy algorithm is used: The new key is inserted in one of its two possible locations, "kicking out", that is, displacing, any key that might already reside in this location. This displaced key is then inserted in its alternative location, again kicking out any key that might reside there, until a vacant position is found, or the procedure enters an infinite loop. In the latter case, the hash table is rebuilt in-place using new hash functions:

There is no need to allocate new tables for the rehashing: We may simply run through the tables to delete and perform the usual insertion procedure on all keys found not to be at their intended position in the table.

—Pagh & Rodler, "Cuckoo Hashing"<sup>[1]</sup>

Lookup requires inspection of just two locations in the hash table, which takes constant time in the worst case (*see* Big O notation). This is in contrast to many other hash table algorithms, which may not have a constant worst-case bound on the time to do a lookup.

It can also be shown that insertions succeed in expected constant time,<sup>[1]</sup> even considering the possibility of having to rebuild the table, as long as the number of keys is kept below half of the capacity of the hash table, i.e., the load factor is below 50%. One method of proving this uses the theory of random graphs: one may form an undirected graph called the "Cuckoo Graph" that has a vertex for each hash table location, and an edge for each hashed value, with the endpoints of the edge being the two possible locations of the value. Then, the greedy insertion algorithm for adding a set of values to a cuckoo hash table succeeds if and only if the Cuckoo Graph for this set of values is a pseudoforest, a graph with at most one cycle in each of its connected components. This property is true with high probability for a random graph in which the number of edges is less than half the number of vertices.<sup>[2]</sup>



### Example

The following hashfunctions are given:

$$h(k) = k \pmod{11}$$

$$h'(k) = \left\lfloor \frac{k}{11} \right\rfloor \pmod{11}$$

<b>k</b>	<b>h(k)</b>	<b>h'(k)</b>
20	9	1
50	6	4
53	9	4
75	9	6
100	1	9
67	1	6
105	6	9
3	3	0
36	3	3
39	6	3

[illegible]

2. table for $h'(k)$											
	20	50	53	75	100	67	105	3	36	39	
0											3
1							20	20	20	20	
2											
3									36	39	
4			53	53	53	53	50	50	50	53	
5											
6				75	75	75	75	75	75	67	
7											
8											
9						100	100	100	100	105	
10											

## Cycle

If you now wish to insert the element 6, then you get into a cycle. In the last row of the table we find the same initial situation as at the beginning again.

$$h(6) = 6 \bmod 11 = 6$$

$$h'(6) = \lfloor \frac{6}{11} \rfloor \bmod 11 = 0$$

considered key	table 1		table 2	
	old value	new value	old value	new value
6	50	6	53	50
53	75	53	67	75
67	100	67	105	100
105	6	105	3	6
3	36	3	39	36
39	105	39	100	105
100	67	100	75	67
75	53	75	50	53
50	39	50	36	39
36	3	36	6	3
6	50	6	53	50

## Generalizations and applications

Generalizations of cuckoo hashing that use more than 2 alternative hash functions can be expected to utilize a larger part of the capacity of the hash table efficiently while sacrificing some lookup and insertion speed. Using just three hash functions increases the load to 91%. Another generalization of cuckoo hashing consists in using more than one key per bucket. Using just 2 keys per bucket permits a load factor above 80%.

Other algorithms that use multiple hash functions include the Bloom filter. Cuckoo hashing can be used to implement a data structure equivalent to a Bloom filter. A simplified generalization of cuckoo hashing called skewed-associative cache is used in some CPU caches.

A study by Zukowski et al.<sup>[3]</sup> has shown that cuckoo hashing is much faster than chained hashing for small, cache-resident hash tables on modern processors. Kenneth Ross<sup>[4]</sup> has shown bucketized versions of cuckoo hashing (variants that use buckets that contain more than one key) to be faster than conventional methods also for large hash tables, when space utilization is high. The performance of the bucketized cuckoo hash table was investigated further by Askitis,<sup>[5]</sup> with its performance compared against alternative hashing schemes.

A survey by Mitzenmacher<sup>[6]</sup> presents open problems related to cuckoo hashing as of 2009.

## References

- [1] Pagh, Rasmus; Rodler, Flemming Friche (2001). "Cuckoo Hashing" (<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.25.4189>). *Algorithms — ESA 2001*. Lecture Notes in Computer Science. **2161**. pp. 121–133. doi:10.1007/3-540-44676-1\_10. ISBN 978-3-540-42493-2. .
  - [2] Kutzelnigg, Reinhard (2006). "Fourth Colloquium on Mathematics and Computer Science" (<http://www.dmtcs.org/dmtcs-ojs/index.php/proceedings/article/viewFile/590/1710>). Discrete Mathematics and Theoretical Computer Science. pp. 403–406.
  - [3] Zukowski, Marcin; Heman, Sandor; Boncz, Peter (2006-06) (PDF). *Architecture-Conscious Hashing* (<http://www.cs.cmu.edu/~damon2006/pdf/zukowski06archconscioushashing.pdf>). Proceedings of the International Workshop on Data Management on New Hardware (DaMoN). . Retrieved 2008-10-16.
  - [4] Ross, Kenneth (2006-11-08) (PDF). *Efficient Hash Probes on Modern Processors* ([http://domino.research.ibm.com/library/cyberdig.nsf/papers/DF54E3545C82E8A585257222006FD9A2/\\$File/rc24100.pdf](http://domino.research.ibm.com/library/cyberdig.nsf/papers/DF54E3545C82E8A585257222006FD9A2/$File/rc24100.pdf)). IBM Research Report RC24100. RC24100. . Retrieved 2008-10-16.
  - [5] Askitis, Nikolas (2009). *Fast and Compact Hash Tables for Integer Keys* (<http://crpit.com/confpapers/CRPITV91Askitis.pdf>). **91**. pp. 113–122. ISBN 978-1-920682-72-9. .
  - [6] Mitzenmacher, Michael (2009-09-09) (PDF). *Some Open Questions Related to Cuckoo Hashing | Proceedings of ESA 2009* (<http://www.eecs.harvard.edu/~michaelm/postscripts/esa2009.pdf>). . Retrieved 2010-11-10.
- A cool and practical alternative to traditional hash tables (<http://www.ru.is/faculty/ulfar/CuckooHash.pdf>), U. Erlingsson, M. Manasse, F. Mcsherry, 2006.
  - Cuckoo Hashing for Undergraduates, 2006 (<http://www.it-c.dk/people/pagh/papers/cuckoo-undergrad.pdf>), R. Pagh, 2006.
  - Cuckoo Hashing, Theory and Practice (<http://mybiasedcoin.blogspot.com/2007/06/cuckoo-hashing-theory-and-practice-part.html>) (Part 1, Part 2 ([http://mybiasedcoin.blogspot.com/2007/06/cuckoo-hashing-theory-and-practice-part\\_15.html](http://mybiasedcoin.blogspot.com/2007/06/cuckoo-hashing-theory-and-practice-part_15.html)) and Part 3 ([http://mybiasedcoin.blogspot.com/2007/06/cuckoo-hashing-theory-and-practice-part\\_19.html](http://mybiasedcoin.blogspot.com/2007/06/cuckoo-hashing-theory-and-practice-part_19.html))), Michael Mitzenmacher, 2007.
  - Naor, Moni; Segev, Gil; Wieder, Udi (2008). "History-Independent Cuckoo Hashing" ([http://www.wisdom.weizmann.ac.il/~naor/PAPERS/cuckoo\\_hi\\_abs.html](http://www.wisdom.weizmann.ac.il/~naor/PAPERS/cuckoo_hi_abs.html)). *International Colloquium on Automata, Languages and Programming (ICALP)*. Reykjavik, Iceland. Retrieved 2008-07-21.

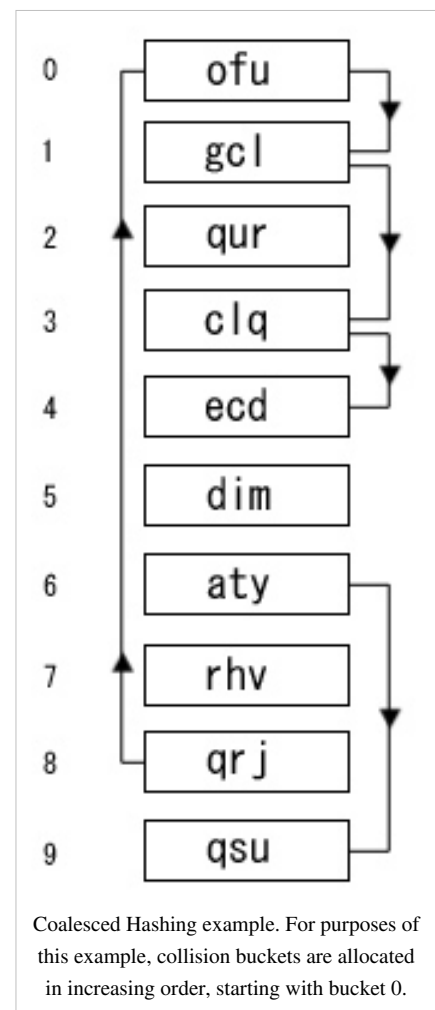
## External links

- Cuckoo hash map written in C++ (<http://sourceforge.net/projects/cuckoo-cpp/>)
- Static cuckoo hashtable generator for C/C++ (<http://www.theiling.de/projects/lookuptable.html>)
- Cuckoo hashtable written in Java (<http://lmonson.com/blog/?p=100>)
- Generic Cuckoo hashmap in Java (<http://github.com/joacima/Cuckoo-hash-map/blob/master/CuckooHashMap.java>)
- Cuckoo hash table written in Haskell (<http://hackage.haskell.org/packages/archive/hashtables/latest/doc/html/Data-HashTable-ST-Cuckoo.html>)

## Coalesced hashing

**Coalesced hashing**, also called **coalesced chaining**, is a strategy of collision resolution in a hash table that forms a hybrid of separate chaining and open addressing. In a separate chaining hash table, items that hash to the same address are placed on a list (or "chain") at that address. This technique can result in a great deal of wasted memory because the table itself must be large enough to maintain a load factor that performs well (typically twice the expected number of items), and extra memory must be used for all but the first item in a chain (unless list headers are used, in which case extra memory must be used for all items in a chain).

Given a sequence "qrj," "aty," "qur," "dim," "ofu," "gcl," "rhv," "clq," "ecd," "qsu" of randomly generated three character long strings, the following table would be generated (using Bob Jenkins' One-at-a-Time hash algorithm<sup>[1]</sup>) with a table of size 10:



(null)			
"clq"			
"qur"			
(null)			
(null)			
"dim"			
"aty"	"qsu"		
"rhv"			
"qrj"	"ofu"	"gcl"	"ecd"
(null)			
(null)			

This strategy is effective, efficient, and very easy to implement. However, sometimes the extra memory use might be prohibitive, and the most common alternative, open addressing, has uncomfortable disadvantages that decrease performance. The primary disadvantage of open addressing is primary and secondary clustering, in which searches may access long sequences of used buckets that contain items with different hash addresses; items with one hash address can thus lengthen searches for items with other hash addresses.

One solution to these issues is coalesced hashing. Coalesced hashing uses a similar technique as separate chaining, but instead of allocating new nodes for the linked list, buckets in the actual table are used. The first empty bucket in the table at the time of a collision is considered the collision bucket. When a collision occurs anywhere in the table, the item is placed in the collision bucket and a link is made between the chain and the collision bucket. It is possible for a newly inserted item to collide with items with a different hash address, such as the case in the example above when item "clq" is inserted. The chain for "clq" is said to "coalesce" with the chain of "qrj," hence the name of the algorithm. However, the extent of coalescing is minor compared with the clustering exhibited by open addressing. For example, when coalescing occurs, the length of the chain grows by only 1, whereas in open addressing, search sequences of arbitrary length may combine.

An important optimization, to reduce the effect of coalescing, is to restrict the address space of the hash function to only a subset of the table. For example, if the table has size  $M$  with buckets numbered from 0 to  $M - 1$ , we can restrict the address space so that the hash function only assigns addresses to the first  $N$  locations in the table. The remaining  $M - N$  buckets, called the *cellar*, are used exclusively for storing items that collide during insertion. No coalescing can occur until the cellar is exhausted.

The optimal choice of  $N$  relative to  $M$  depends upon the load factor (or fullness) of the table. A careful analysis shows that the value  $N = 0.86 \times M$  yields near-optimum performance for most load factors.<sup>[2]</sup> Other variants for insertion are also possible that have improved search time. Deletion algorithms have been developed that preserve randomness, and thus the average search time analysis still holds after deletions.<sup>[2]</sup>

Insertion in C:

```
/* htab is the hash table,
   N is the size of the address space of the hash function, and
   M is the size of the entire table including the cellar.
   Collision buckets are allocated in decreasing order, starting with
   bucket M-1. */

int insert ( char key[] )
{
```

```

unsigned h = hash ( key, strlen ( key ) ) % N;

if ( htab[h] == NULL ) {
    /* Make a new chain */
    htab[h] = make_node ( key, NULL );
} else {
    struct node *it;
    int cursor = M-1;

    /* Find the first empty bucket */
    while ( cursor >= 0 && htab[cursor] != NULL )
        --cursor;

    /* The table is full, terminate unsuccessfully */
    if ( cursor == -1 )
        return -1;

    htab[cursor] = make_node ( key, NULL );

    /* Find the last node in the chain and point to it */
    it = htab[h];

    while ( it->next != NULL )
        it = it->next;

    it->next = htab[cursor];
}

return 0;
}

```

One benefit of this strategy is that the search algorithm for separate chaining can be used without change in a coalesced hash table.

Lookup in C:

```

char *find ( char key[] )
{
    unsigned h = hash ( key, strlen ( key ) ) % N;

    if ( htab[h] != NULL ) {
        struct node *it;

        /* Search the chain at index h */
        for ( it = htab[h]; it != NULL; it = it->next ) {
            if ( strcmp ( key, it->data ) == 0 )
                return it->data;
        }
    }
}

```



```
return NULL;
}
```

## Performance

Coalesced chaining avoids the effects of primary and secondary clustering, and as a result can take advantage of the efficient search algorithm for separate chaining. If the chains are short, this strategy is very efficient and can be highly condensed, memory-wise. As in open addressing, deletion from a coalesced hash table is awkward and potentially expensive, and resizing the table is terribly expensive and should be done rarely, if ever.

## References

- [1] <http://burtleburtle.net/bob/>
- [2] J. S. Vitter and W.-C. Chen, *Design and Analysis of Coalesced Hashing*, Oxford University Press, New York, NY, 1987, ISBN 0-19-504182-8

# Perfect hash function

---

A **perfect hash function** for a set  $S$  is a hash function that maps distinct elements in  $S$  to a set of integers, with no collisions. A perfect hash function has many of the same applications as other hash functions, but with the advantage that no collision resolution has to be implemented. In mathematical terms, it is a total injective function.

## Properties and uses

A perfect hash function for a specific set  $S$  that can be evaluated in constant time, and with values in a small range, can be found by a randomized algorithm in a number of operations that is proportional to the size of  $S$ . The minimal size of the description of a perfect hash function depends on the range of its function values: The smaller the range, the more space is required. Any perfect hash functions suitable for use with a hash table require at least a number of bits that is proportional to the size of  $S$ .

A perfect hash function with values in a limited range can be used for efficient lookup operations, by placing keys from  $S$  (or other associated values) in a table indexed by the output of the function. Using a perfect hash function is best in situations where there is a frequently queried large set,  $S$ , which is seldom updated. Efficient solutions to performing updates are known as dynamic perfect hashing, but these methods are relatively complicated to implement. A simple alternative to perfect hashing, which also allows dynamic updates, is cuckoo hashing.

## Minimal perfect hash function

A **minimal perfect hash function** is a perfect hash function that maps  $n$  keys to  $n$  consecutive integers—usually  $[0..n-1]$  or  $[1..n]$ . A more formal way of expressing this is: Let  $j$  and  $k$  be elements of some finite set  $\mathbf{K}$ .  $F$  is a minimal perfect hash function iff  $F(j) = F(k)$  implies  $j = k$  and there exists an integer  $a$  such that the range of  $F$  is  $a..a+|\mathbf{K}|-1$ . It has been proved that a general purpose minimal perfect hash scheme requires at least 1.44 bits/key.<sup>[1]</sup> However the smallest currently use around 2.5 bits/key.

A minimal perfect hash function  $F$  is **order preserving** if keys are given in some order  $a_1, a_2, \dots$ , and for any keys  $a_j$  and  $a_k$ ,  $j < k$  implies  $F(a_j) < F(a_k)$ . Order-preserving minimal perfect hash functions require necessarily  $\Omega(n \log n)$  bits to be represented.

A minimal perfect hash function  $F$  is **monotone** if it preserves the lexicographical order of the keys. Monotone minimal perfect hash functions can be represented in very little space.

## References

- [1] Djamel Belazzougui, Fabiano C. Botelho, Martin Dietzfelbinger (2009) (PDF). *Hash, displace, and compress* (<http://cmph.sourceforge.net/papers/esa09.pdf>). Springer Berlin / Heidelberg. . Retrieved 2011-08-11.

## Further reading

- Richard J. Cichelli. *Minimal Perfect Hash Functions Made Simple*, Communications of the ACM, Vol. 23, Number 1, January 1980.
- Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7. Section 11.5: Perfect hashing, pp. 245–249.
- Fabiano C. Botelho, Rasmus Pagh and Nivio Ziviani. "Perfect Hashing for Data Management Applications" (<http://arxiv.org/pdf/cs/0702159>).
- Fabiano C. Botelho and Nivio Ziviani. "External perfect hashing for very large key sets" (<http://homepages.dcc.ufmg.br/~nivio/papers/cikm07.pdf>). 16th ACM Conference on Information and Knowledge Management (CIKM07), Lisbon, Portugal, November 2007.
- Djamel Belazzougui, Paolo Boldi, Rasmus Pagh, and Sebastiano Vigna. "Monotone minimal perfect hashing: Searching a sorted table with  $O(1)$  accesses" (<http://vigna.dsi.unimi.it/ftp/papers/MonotoneMinimalPerfectHashing.pdf>). In Proceedings of the 20th Annual ACM-SIAM Symposium On Discrete Mathematics (SODA), New York, 2009. ACM Press.
- Djamel Belazzougui, Paolo Boldi, Rasmus Pagh, and Sebastiano Vigna. "Theory and practise of monotone minimal perfect hashing" ([http://www.siam.org/proceedings/alnex/2009/alx09\\_013\\_belazzouguid.pdf](http://www.siam.org/proceedings/alnex/2009/alx09_013_belazzouguid.pdf)). In Proceedings of the Tenth Workshop on Algorithm Engineering and Experiments (ALENEX). SIAM, 2009.

## External links

- Minimal Perfect Hashing (<http://burtleburtle.net/bob/hash/perfect.html>) by Bob Jenkins
  - gperf (<http://www.gnu.org/software/gperf/>) is a Free software C and C++ perfect hash generator
  - cmph (<http://cmph.sourceforge.net/index.html>) is Free Software implementing many perfect hashing methods
  - Sux4J (<http://sux4j.dsi.unimi.it/>) is Free Software implementing perfect hashing, including monotone minimal perfect hashing in Java
  - MPHSharp (<http://www.dupuis.me/node/9>) is Free Software implementing many perfect hashing methods in C#
-

# Universal hashing

Using **universal hashing** (in a randomized algorithm or data structure) refers to selecting a hash function at random from a family of hash functions with a certain mathematical property (see definition below). This guarantees a low number of collisions in expectation, even if the data is chosen by an adversary. Many universal families are known (for hashing integers, vectors, strings), and their evaluation is often very efficient. Universal hashing has numerous uses in computer science, for example in implementations of hash tables, randomized algorithms, and cryptography.

## Introduction

Assume we want to map keys from some universe  $U$  into  $m$  bins (labelled  $[m] = \{0, \dots, m-1\}$ ). The algorithm will have to handle some data set  $S \subseteq U$  of  $|S| = n$  keys, which is not known in advance. Usually, the goal of hashing is to obtain a low number of collisions (keys from  $S$  that land in the same bin). A deterministic hash function cannot offer any guarantee in an adversarial setting if the size of  $U$  is greater than  $m^2$ , since the adversary may choose  $S$  to be precisely the preimage of a bin. This means that all data keys land in the same bin, making hashing useless. Furthermore, a deterministic hash function does not allow for *rehashing*: sometimes the input data turns out to be bad for the hash function (e.g. there are too many collisions), so one would like to change the hash function.

The solution to these problems is to pick a function randomly from a family of hash functions. A family of functions  $H = \{h : U \rightarrow [m]\}$  is called a **universal family** if,  $\forall x, y \in U, x \neq y : \Pr_{h \in H}[h(x) = h(y)] \leq \frac{1}{m}$ .

In other words, any two keys of the universe collide with probability at most  $1/m$  when the hash function  $h$  is drawn randomly from  $H$ . This is exactly the probability of collision we would expect if the hash function assigned truly random hash codes to every key. Sometimes, the definition is relaxed to allow collision probability  $O(1/m)$ . This concept was introduced by Carter and Wegman<sup>[1]</sup> in 1977, and has found numerous applications in computer science (see, for example<sup>[2]</sup>). If we have an upper bound of  $\epsilon < 1$  on the collision probability, we say that we have  $\epsilon$ -almost universality.

Many, but not all, universal families have the following stronger **uniform difference property**:

$\forall x, y \in U, x \neq y$ , when  $h$  is drawn randomly from the family  $H$ , the difference  $h(x) - h(y) \bmod m$  is uniformly distributed in  $[m]$ . Note that the definition of universality is only concerned with whether  $h(x) - h(y) = 0$ , which counts collisions. The uniform difference property is stronger.

(Similarly, a universal family can be XOR universal if  $\forall x, y \in U, x \neq y$ , the value  $h(x) \oplus h(y) \bmod m$  is uniformly distributed in  $[m]$  where  $\oplus$  is the bitwise exclusive or operation. This is only possible if  $m$  is a power of two.)

An even stronger condition is pairwise independence: we have this property when  $\forall x, y \in U, x \neq y$  we have the probability that  $x, y$  will hash to any pair of hash values  $z_1, z_2$  is as if they were perfectly random:  $P(h(x) = z_1 \wedge h(y) = z_2) = 1/m^2$ . Pairwise independence is sometimes called strong universality.

Another property is uniformity. We say that a family is uniform if all hash values are equally likely:  $P(h(x) = z) = 1/m$  for any hash value  $z$ . Universality does not imply uniformity. However, strong universality does imply uniformity.

Given a family with the uniform distance property, one can produce a pairwise independent or strongly universal hash family by adding a uniformly distributed random constant with values in  $[m]$  to the hash functions. (Similarly, if  $m$  is a power of two, we can achieve pairwise independence from an XOR universal hash family by doing an exclusive or with a uniformly distributed random constant.) Since a shift by a constant is sometimes irrelevant in applications (e.g. hash tables), a careful distinction between the uniform distance property and pairwise independent

is sometimes not made.<sup>[3]</sup>

For some applications (such as hash tables), it is important for the least significant bits of the hash values to be also universal. When a family is strongly universal, this is guaranteed: if  $H$  is a strongly universal family with  $m = 2^L$ , then the family made of the functions  $h \bmod 2^{L'}$  for all  $h \in H$  is also strongly universal for  $L' \leq L$ . Unfortunately, the same is not true of (merely) universal families. For example the family made of the identity function  $h(x) = x$  is clearly universal, but the family made of the function  $h(x) = x \bmod 2^{L'}$  fails to be universal.

## Mathematical guarantees

For any fixed set  $S$  of  $n$  keys, using a universal family guarantees the following properties.

1. For any fixed  $x$  in  $S$ , the expected number of keys in the bin  $h(x)$  is  $n/m$ . When implementing hash tables by chaining, this number is proportional to the expected running time of an operation involving the key  $x$  (for example a query, insertion or deletion).
2. The expected number of pairs of keys  $x, y$  in  $S$  with  $x \neq y$  that collide ( $h(x) = h(y)$ ) is bounded above by  $n(n-1)/2m$ , which is of order  $O(n^2/m)$ . When the number of bins,  $m$ , is  $O(n)$ , the expected number of collisions is  $O(n)$ . When hashing into  $n^2$  bins, there are no collisions at all with probability at least a half.
3. The expected number of keys in bins with at least  $t$  keys in them is bounded above by  $2n/(t - 2(n/m) + 1)$ .<sup>[4]</sup> Thus, if the capacity of each bin is capped to three times the average size ( $t = 3n/m$ ), the total number of keys in overflowing bins is at most  $O(m)$ . This only holds with a hash family whose collision probability is bounded above by  $1/m$ . If a weaker definition is used, bounding it by  $O(1/m)$ , this result is no longer true.<sup>[4]</sup>

As the above guarantees hold for any fixed set  $S$ , they hold if the data set is chosen by an adversary. However, the adversary has to make this choice before (or independent of) the algorithm's random choice of a hash function. If the adversary can observe the random choice of the algorithm, randomness serves no purpose, and the situation is the same as deterministic hashing.

The second and third guarantee are typically used in conjunction with rehashing. For instance, a randomized algorithm may be prepared to handle some  $O(n)$  number of collisions. If it observes too many collisions, it chooses another random  $h$  from the family and repeats. Universality guarantees that the number of repetitions is a geometric random variable.

## Constructions

Since any computer data can be represented as one or more machine words, one generally needs hash functions for three types of domains: machine words ("integers"); fixed-length vectors of machine words; and variable-length vectors ("strings").

### Hashing integers

This section refers to the case of hashing integers that fit in machine words; thus, operations like multiplication, addition, division, etc. are cheap machine-level instructions. Let the universe to be hashed be  $U = \{0, \dots, u-1\}$ .

The original proposal of Carter and Wegman<sup>[1]</sup> was to pick a prime  $p \geq u$  and define

$$h_{a,b}(x) = ((ax + b) \bmod p) \bmod m$$

where  $a, b$  are randomly chosen integers modulo  $p$  with  $a \neq 0$ . Technically, adding  $b$  is not needed for universality (but it does make the hash function 2-independent).

To see that  $H = \{h_{a,b}\}$  is a universal family, note that  $h(x) = h(y)$  only holds when

$$ax + b \equiv ay + b + i \cdot m \pmod{p}$$

for some integer  $i$  between 0 and  $p/m$ . If  $x \neq y$ , their difference,  $x - y$  is nonzero and has an inverse modulo  $p$ . Solving for  $a$ ,

$$a \equiv i \cdot m \cdot (x - y)^{-1} \pmod{p}.$$

There are  $p - 1$  possible choices for  $a$  (since  $a = 0$  is excluded) and, varying  $i$  in the allowed range,  $\lfloor p/m \rfloor$  possible values for the right hand side. Thus the collision probability is

$$\lfloor p/m \rfloor / (p - 1)$$

which tends to  $1/m$  for large  $p$  as required. This analysis also shows that  $b$  does not have to be randomised in order to have universality.

Another way to see  $H$  is a universal family is via the notion of statistical distance. Write the difference  $h(x) - h(y)$  as

$$h(x) - h(y) \equiv (a(x - y) \bmod p) \pmod{m}.$$

Since  $x - y$  is nonzero and  $a$  is uniformly distributed in  $\{1, \dots, p\}$ , it follows that  $a(x - y) \bmod p$  is also uniformly distributed in  $\{1, \dots, p\}$ . The distribution of  $(h(x) - h(y)) \bmod m$  is thus almost uniform, up to a difference in probability of  $\pm 1/p$  between the samples. As a result, the statistical distance to a uniform family is  $O(m/p)$ , which becomes negligible when  $p \gg m$ .

### Avoiding modular arithmetic

The state of the art for hashing integers is the **multiply-shift** scheme described by Dietzfelbinger et al. in 1997.<sup>[5]</sup> By avoiding modular arithmetic, this method is much easier to implement and also runs significantly faster in practice (usually by at least a factor of four<sup>[6]</sup>). The scheme assumes the number of bins is a power of two,  $m = 2^M$ . Let  $w$  be the number of bits in a machine word. Then the hash functions are parametrised over odd positive integers  $a < 2^w$  (that fit in a word of  $w$  bits). To evaluate  $h_a(x)$ , multiply  $x$  by  $a$  modulo  $2^w$  and then keep the high order  $M$  bits as the hash code. In mathematical notation, this is

$$h_a(x) = (a \cdot x \bmod 2^w) \operatorname{div} 2^{w-M}$$

and it can be implemented in C-like programming languages by

$$h_a(x) = (\text{unsigned}) (a * x) \gg (w - M)$$

This scheme does *not* satisfy the uniform difference property and is only  $2/m$ -almost-universal; for any  $x \neq y$ ,  $\Pr\{h_a(x) = h_a(y)\} \leq 2/m$ .

To understand the behavior of the hash function, notice that, if  $ax \bmod 2^w$  and  $ay \bmod 2^w$  have the same highest-order  $M$  bits, then  $a(x - y) \bmod 2^w$  has either all 1's or all 0's as its highest order  $M$  bits (depending on whether  $ax \bmod 2^w$  or  $ay \bmod 2^w$  is larger. Assume that the least significant set bit of  $x - y$  appears on position  $w - c$ . Since  $a$  is a random odd integer and odd integers have inverses in the ring  $\mathbb{Z}_{2^w}$ , it follows that  $a(x - y) \bmod 2^w$  will be uniformly distributed among  $w$ -bit integers with the least significant set bit on position  $w - c$ . The probability that these bits are all 0's or all 1's is therefore at most  $2/2^M = 2/m$ . On the other hand, if  $c < M$ , then higher-order  $M$  bits of  $a(x - y) \bmod 2^w$  contain both 0's and 1's, so it is certain that  $h(x) \neq h(y)$ . Finally, if  $c = M$  then bit  $w - M$  of  $a(x - y) \bmod 2^w$  is 1 and  $h_a(x) = h_a(y)$  if and only if bits  $w - 1$  through  $w - M + 1$  are also 1, which has probability  $2^{-(M-1)}$ . To obtain a truly 'universal' hash function, one can use the multiply-add-shift scheme

$$h_{a,b}(x) = ((ax + b) \bmod 2^w) \operatorname{div} 2^{w-M}$$

where  $a$  is a random odd positive integer with  $a < 2^w$  and  $b = i2^{w/2}$  where  $i$  is chosen at random from  $\{0, \dots, 2^{w/2} - 1\}$ . With these choices of  $a$  and  $b$ ,  $\Pr\{h_{a,b}(x) = h_{a,b}(y)\} \leq 1/m$  for all

$$x \not\equiv y \pmod{2^w}.^{[7]}$$

## Hashing vectors

This section is concerned with hashing a fixed-length vector of machine words. Interpret the input as a vector  $\bar{x} = (x_0, \dots, x_{k-1})$  of  $k$  machine words (integers of  $w$  bits each). If  $H$  is a universal family with the uniform difference property, the following family dating back to Carter and Wegman<sup>[1]</sup> also has the uniform difference property (and hence is universal):

$$h(\bar{x}) = \left( \sum_{i=0}^{k-1} h_i(x_i) \right) \bmod m, \text{ where each } h_i \in H \text{ is chosen independently at random.}$$

If  $m$  is a power of two, one may replace summation by exclusive or.<sup>[8]</sup>

In practice, if double-precision arithmetic is available, this is instantiated with the multiply-shift hash family of.<sup>[9]</sup> Initialize the hash function with a vector  $\bar{a} = (a_0, \dots, a_{k-1})$  of random **odd** integers on  $2w$  bits each. Then if the number of bins is  $m = 2^M$  for  $M \leq w$ :

$$h_{\bar{a}}(\bar{x}) = \left( \left( \sum_{i=0}^{k-1} x_i \cdot a_i \right) \bmod 2^{2w} \right) \text{div } 2^{2w-M}.$$

It is possible to halve the number of multiplications, which roughly translates to a two-fold speed-up in practice.<sup>[8]</sup> Initialize the hash function with a vector  $\bar{a} = (a_0, \dots, a_{k-1})$  of random **odd** integers on  $2w$  bits each. The following hash family is universal<sup>[10]</sup>:

$$h_{\bar{a}}(\bar{x}) = \left( \left( \sum_{i=0}^{\lceil k/2 \rceil} (x_{2i} + a_{2i}) \cdot (x_{2i+1} + a_{2i+1}) \right) \bmod 2^{2w} \right) \text{div } 2^{2w-M}.$$

If double-precision operations are not available, one can interpret the input as a vector of half-words ( $w/2$ -bit integers). The algorithm will then use  $\lceil k/2 \rceil$  multiplications, where  $k$  was the number of half-words in the vector.

Thus, the algorithm runs at a "rate" of one multiplication per word of input.

The same scheme can also be used for hashing integers, by interpreting their bits as vectors of bytes. In this variant, the vector technique is known as tabulation hashing and it provides a practical alternative to multiplication-based universal hashing schemes.<sup>[11]</sup>

## Hashing strings

This refers to hashing a *variable-sized* vector of machine words. If the length of the string can be bounded by a small number, it is best to use the vector solution from above (conceptually padding the vector with zeros up to the upper bound). The space required is the maximal length of the string, but the time to evaluate  $h(s)$  is just the length of  $s$  (the zero-padding can be ignored when evaluating the hash function without affecting universality<sup>[8]</sup>).

Now assume we want to hash  $\bar{x} = (x_0, \dots, x_{\ell})$ , where a good bound on  $\ell$  is not known a priori. A universal family proposed by.<sup>[9]</sup> treats the string  $x$  as the coefficients of a polynomial modulo a large prime. If  $x_i \in [u]$ , let  $p \geq \max\{u, m\}$  be a prime and define:

$$h_a(\bar{x}) = h_{\text{int}} \left( \left( \sum_{i=0}^{\ell} x_i \cdot a^i \right) \bmod p \right), \text{ where } a \in [p] \text{ is uniformly random and } h_{\text{int}} \text{ is chosen}$$

randomly from a universal family mapping integer domain  $[p] \mapsto [m]$ .

Consider two strings  $\bar{x}, \bar{y}$  and let  $\ell$  be length of the longer one; for the analysis, the shorter string is conceptually padded with zeros up to length  $\ell$ . A collision before applying  $h_{\text{int}}$  implies that  $a$  is a root of the polynomial with coefficients  $\bar{x} - \bar{y}$ . This polynomial has at most  $\ell$  roots modulo  $p$ , so the collision probability is at most  $\ell/p$ .

The probability of collision through the random  $h_{\text{int}}$  brings the total collision probability to  $\frac{1}{m} + \frac{\ell}{p}$ . Thus, if the

prime  $p$  is sufficiently large compared to the length of strings hashed, the family is very close to universal (in statistical distance).

To mitigate the computational penalty of modular arithmetic, two tricks are used in practice<sup>[8]</sup>:

1. One chooses the prime  $p$  to be close to a power of two, such as a Mersenne prime. This allows arithmetic modulo  $p$  to be implemented without division (using faster operations like addition and shifts). For instance, on modern architectures one can work with  $p = 2^{61} - 1$ , while  $x_i$ 's are 32-bit values.
2. One can apply vector hashing to blocks. For instance, one applies vector hashing to each 16-word block of the string, and applies string hashing to the  $\lceil k/16 \rceil$  results. Since the slower string hashing is applied on a substantially smaller vector, this will essentially be as fast as vector hashing.

## References

- [1] Carter, Larry; Wegman, Mark N. (1979). "Universal Classes of Hash Functions". *Journal of Computer and System Sciences* **18** (2): 143–154. doi:10.1016/0022-0000(79)90044-8. Conference version in STOC'77.
- [2] Miltersen, Peter Bro. "Universal Hashing" (<http://www.webcitation.org/5hmOaVISI>) (PDF). Archived from the original (<http://www.daimi.au.dk/~bromille/Notes/un.pdf>) on 24th June 2009. .
- [3] Motwani, Rajeev; Raghavan, Prabhakar (1995). *Randomized Algorithms*. Cambridge University Press. p. 221. ISBN 0-521-47465-5.
- [4] Baran, Ilya; Demaine, Erik D.; Pătraşcu, Mihai (2008). "Subquadratic Algorithms for 3SUM" (<http://people.csail.mit.edu/mip/papers/3sum/3sum.pdf>). *Algorithmica* **50** (4): 584–596. doi:10.1007/s00453-007-9036-3. .
- [5] Dietzfelbinger, Martin; Hagerup, Torben; Katajainen, Jyrki; Penttonen, Martti (1997). "A Reliable Randomized Algorithm for the Closest-Pair Problem" (<http://www.diku.dk/~jyrki/Paper/CP-11.4.1997.ps>) (Postscript). *Journal of Algorithms* **25** (1): 19–51. doi:10.1006/jagm.1997.0873. . Retrieved 10 February 2011.
- [6] Thorup, Mikkel. "Text-book algorithms at SODA" (<http://mybiasedcoin.blogspot.com/2009/12/text-book-algorithms-at-soda-guest-post.html>). .
- [7] Woelfel, Philipp (1999). "Efficient Strongly Universal and Optimally Universal Hashing" (<http://www.springerlink.com/content/a10p748w7pr48682/>) (PDF). LNCS. **1672**. Mathematical Foundations of Computer Science 1999. pp. 262–272. doi:10.1007/3-540-48340-3\_24. . Retrieved 17 May 2011.
- [8] Thorup, Mikkel (2009). "String hashing for linear probing" ([http://www.siam.org/proceedings/soda/2009/SODA09\\_072\\_thorupm.pdf](http://www.siam.org/proceedings/soda/2009/SODA09_072_thorupm.pdf)). *Proc. 20th ACM-SIAM Symposium on Discrete Algorithms (SODA)*. pp. 655–664. ., section 5.3
- [9] Dietzfelbinger, Martin; Gil, Joseph; Matias, Yossi; Pippenger, Nicholas (1992). "Polynomial Hash Functions Are Reliable (Extended Abstract)". *Proc. 19th International Colloquium on Automata, Languages and Programming (ICALP)*. pp. 235–246.
- [10] Black, J.; Halevi, S.; Krawczyk, H.; Krovetz, T. (1999). "UMAC: Fast and Secure Message Authentication" (<http://www.cs.ucdavis.edu/~rogaway/papers/umac-full.pdf>). *Advances in Cryptology (CRYPTO '99)*. ., Equation 1
- [11] Pătraşcu, Mihai; Thorup, Mikkel (2011). "The power of simple tabulation hashing". *Proceedings of the 43rd annual ACM Symposium on Theory of Computing (STOC '11)*. pp. 1–10. arXiv:1011.5200. doi:10.1145/1993636.1993638.

## Further reading

- Knuth, Donald Ervin (1998). *[The Art of Computer Programming], Vol. III: Sorting and Searching* (2e ed.). Reading, Mass ; London: Addison-Wesley. ISBN 0-201-89685-0. knuth.

## External links

- Open Data Structures - Section 5.1.1 - Multiplicative Hashing ([http://opendatastructures.org/versions/edition-0.1e/ods-java/5\\_1\\_ChainedHashTable\\_Hashin.html#SECTION00811000000000000000](http://opendatastructures.org/versions/edition-0.1e/ods-java/5_1_ChainedHashTable_Hashin.html#SECTION00811000000000000000))

# Linear hashing

**Linear hashing** is a dynamic hash table algorithm invented by Witold Litwin (1980),<sup>[1]</sup> and later popularized by Paul Larson. Linear hashing allows for the expansion of the hash table one slot at a time. The frequent single slot expansion can very effectively control the length of the collision chain. The cost of hash table expansion is spread out across each hash table insertion operation, as opposed to being incurred all at once.<sup>[2]</sup> Linear hashing is therefore well suited for interactive applications.

## Algorithm Details

A hash function controls the address calculation of linear hashing. In linear hashing, the address calculation is always bounded by a size that is a power of two \* N, where N is the chosen original number of buckets. The number of buckets is given by  $N * 2^{\text{Level}}$  e.g. Level 0, N; Level 1, 2N; Level 2, 4N.

```
address(level, key) = hash(key) mod (N * 2level)
```

The 'split' variable controls the read operation, and the expansion operation.

A read operation would use address(level, key) if address(level, key) is greater than or equal to the 'split' variable. Otherwise, address(level+1, key) is used. This takes into account the fact that buckets numbered less than split have been rehashed with address(level+1, key) after its contents split between two new buckets (the first bucket writing over the contents of the old single bucket prior to the split).

A linear hashing table expansion operation would consist of rehashing the entries at one slot location indicated by the 'split' variable to either of two target slot locations of address(level+1, key). This intuitively is consistent with the assertion that if  $y = x \bmod M$  and  $y' = x \bmod M * 2$ , then  $y' = y$  or  $y' = y + M$ .

The 'split' variable is incremented by 1 at the end of the expansion operation. If the 'split' variable reaches  $N * 2^{\text{level}}$ , then the 'level' variable is incremented by 1, and the 'split' variable is reset to 0.

Thus the hash buckets are expanded round robin, and seem unrelated to where buckets overflow at the time of expansion. Overflow buckets are used at the sites of bucket overflow (the normal bucket has a pointer to the overflow bucket), but these are eventually reabsorbed when the round robin comes to the bucket with the overflow bucket, and the contents of that bucket and the overflow bucket are redistributed by the future hash function  $\text{hash}(\text{key}) \bmod (N * 2^{\text{level}+1})$ .

The degenerate case, which is unlikely with a randomized hash function, is that enough entries are hashed to the same bucket so that there is enough entries to overflow more than one overflow bucket (assuming overflow bucket size = normal bucket size), before being absorbed when that bucket's turn to split comes in the round robin.

The point of the algorithm seems to be that overflow is preempted by gradually increasing the number of available buckets, and overflow buckets are eventually reabsorbed during a later split, which must eventually happen because splitting occurs round robin.

There is some flexibility in choosing how often the expansion operations are performed. One obvious choice is to perform the expansion operation each time no more slots are available at the target slot location. Another choice is to control the expansion with a programmer defined load factor.

The hash table array for linear hashing is usually implemented with a dynamic array algorithm.



## Adoption in language systems

Griswold and Townsend<sup>[3]</sup> discussed the adoption of linear hashing in the Icon language. They discussed the implementation alternatives of dynamic array algorithm used in linear hashing, and presented performance comparisons using a list of Icon benchmark applications.

## References

- [1] Litwin, Witold (1980), "Linear hashing: A new tool for file and table addressing" (<http://www.cs.cmu.edu/afs/cs.cmu.edu/user/christos/www/courses/826-resources/PAPERS+BOOK/linear-hashing.PDF>) (PDF), *Proc. 6th Conference on Very Large Databases*: 212–223,
- [2] Larson, Per-Åke (April 1988), "Dynamic Hash Tables", *Communications of the ACM* **31**: 446–457, doi:10.1145/42404.42410
- [3] Griswold, William G.; Townsend, Gregg M. (April 1993), "The Design and Implementation of Dynamic Hashing for Sets and Tables in Icon" (<http://citeseer.ist.psu.edu/griswold93design.html>), *Software - Practice and Experience* **23** (4): 351–367,

## External links

- Sorted Linear Hash Table, C++ implementation of a Linear Hashtable (<http://www.concentric.net/~Ttwang/tech/sorthash.htm>)
- TommyDS, C implementation of a Linear Hashtable (<http://tommyds.sourceforge.net/>)
- Paul E. Black, linear hashing (<http://www.nist.gov/dads/HTML/linearHashing.html>) at the NIST Dictionary of Algorithms and Data Structures.

# Extendible hashing

**Extendible hashing** is a type of hash system which treats a hash as a bit string, and uses a trie for bucket lookup.<sup>[1]</sup> Because of the hierarchical nature of the system, re-hashing is an incremental operation (done one bucket at a time, as needed). This means that time-sensitive applications are less affected by table growth than by standard full-table rehashes.

## Example

This is an example from Fagin et al. (1979).

Assume that the hash function  $h(k)$  returns a binary number. The first  $i$  bits of each string will be used as indices to figure out where they will go in the "directory" (hash table). Additionally,  $i$  is the smallest number such that the first  $i$  bits of all keys are different.

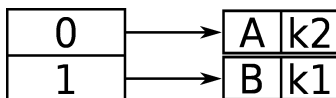
Keys to be used:

$$h(k_1) = 100100$$

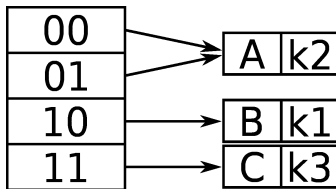
$$h(k_2) = 010110$$

$$h(k_3) = 110110$$

Let's assume that for this particular example, the bucket size is 1. The first two keys to be inserted,  $k_1$  and  $k_2$ , can be distinguished by the most significant bit, and would be inserted into the table as follows:



Now, if  $k_3$  were to be hashed to the table, it wouldn't be enough to distinguish all three keys by one bit (because  $k_3$  and  $k_1$  have 1 as their leftmost bit. Also, because the bucket size is one, the table would overflow. Because comparing the first two most significant bits would give each key a unique location, the directory size is doubled as follows:



And so now  $k_1$  and  $k_3$  have a unique location, being distinguished by the first two leftmost bits. Because  $k_2$  is in the top half of the table, both 00 and 01 point to it because there is no other key to compare to that begins with a 0.

### Further detail

$$h(k_4) = 011110$$

Now,  $k_4$  needs to be inserted, and it has the first two bits as 01..(1110), and using a 2 bit depth in the directory, this maps from 01 to Bucket A. Bucket A is full (max size 1), so it must be split; because there is more than one pointer to Bucket A, there is no need to increase the directory size.

What is needed is information about:

1. The key size that maps the directory (the global depth), and
2. The key size that has previously mapped the bucket (the local depth)

In order to distinguish the two action cases:

1. Doubling the directory when a bucket becomes full
2. Creating a new bucket, and re-distributing the entries between the old and the new bucket

Examining the initial case of an extendible hash structure, if each directory entry points to one bucket, then the local depth should be equal to the global depth.

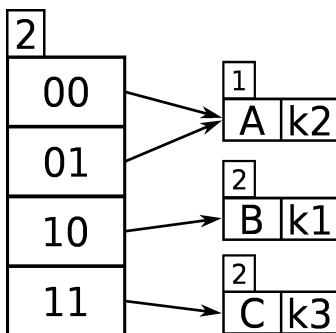
The number of directory entries is equal to  $2^{\text{global depth}}$ , and the initial number of buckets is equal to  $2^{\text{local depth}}$ .

Thus if global depth = local depth = 0, then  $2^0 = 1$ , so an initial directory of one pointer to one bucket.

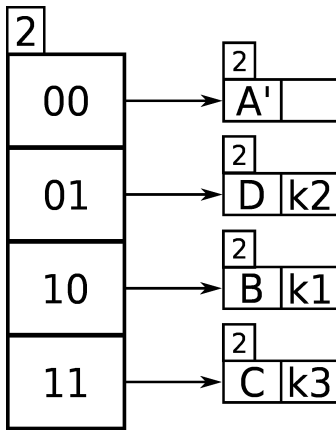
Back to the two action cases:

If the local depth is equal to the global depth, then there is only one pointer to the bucket, and there is no other directory pointers that can map to the bucket, so the directory must be doubled (case1).

If the bucket is full, if the local depth is less than the global depth, then there exists more than one pointer from the directory to the bucket, and the bucket can be split (case 2).



Key 01 points to Bucket A, and Bucket A's local depth of 1 is less than the directory's global depth of 2, which means keys hashed to Bucket A have only used a 1 bit prefix (i.e. 0), and the bucket needs to have its contents split using keys  $1 + 1 = 2$  bits in length; in general, for any local depth  $d$  where  $d$  is less than  $D$ , the global depth, then  $d$  must be incremented after a bucket split, and the new  $d$  used as the number of bits of each entry's key to redistribute the entries of the former bucket into the new buckets.



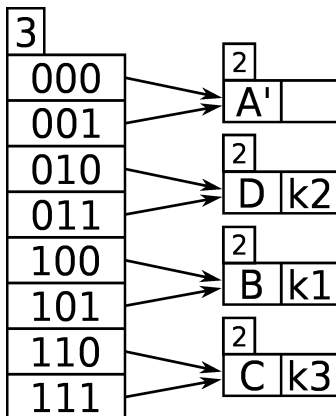
Now,  $h(k_4) = 011110$

is tried again, with 2 bits 01.., and now key 01 points to a new bucket but there is still  $k_2$  in it ( $h(k_2) = 010110$  and also begins with 01).

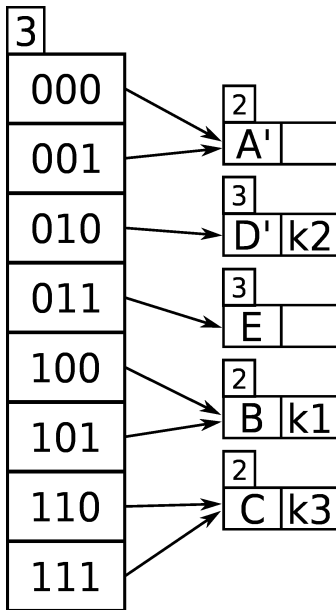
If  $k_2$  had been 000110, with key 00, there would have been no problem, because  $k_2$  would have remained in the new bucket A' and bucket D would have been empty.

(This would have been the most likely case by far when buckets are of greater size than 1 and the newly split buckets would be exceedingly unlikely to overflow, unless all the entries were all rehashed to one bucket again. But just to emphasize the role of the depth information, the example will be pursued logically to the end.)

So Bucket D needs to be split, but a check of its local depth, which is 2, is the same as the global depth, which is 2, so the directory must be split again, in order to hold keys of sufficient detail, e.g. 3 bits.



1. Bucket D needs to split due to being full.
2. As D's local depth = the global depth, the directory must double to increase bit detail of keys.
3. Global depth has incremented after directory split to 3.
4. The new entry  $k_4$  is rekeyed with global depth 3 bits and ends up in D which has local depth 2, which can now be incremented to 3 and D can be split to D' and E.
5. The contents of the split bucket D,  $k_2$ , has been re-keyed with 3 bits, and it ends up in D.
6.  $k_4$  is retried and it ends up in E which has a spare slot.



Now,  $h(k_2) = 010110$  is in D and  $h(k_4) = 011110$  is tried again, with 3 bits 011..., and it points to bucket D which already contains  $k_2$  so is full; D's local depth is 2 but now the global depth is 3 after the directory doubling, so now D can be split into bucket's D' and E, the contents of D,  $k_2$  has its  $h(k_2)$  retried with a new global depth bitmask of 3 and  $k_2$  ends up in D', then the new entry  $k_4$  is retried with  $h(k_4)$  bitmasked using the new global depth bit count of 3 and this gives 011 which now points to a new bucket E which is empty. So  $k_4$  goes in Bucket E.

## Example implementation

Below is the extendible hashing algorithm in Python, with the disc block / memory page association, caching and consistency issues removed. Note a problem exists if the depth exceeds the bit size of an integer, because then doubling of the directory or splitting of a bucket won't allow entries to be rehashed to different buckets.

The code uses the *least significant bits*, which makes it more efficient to expand the table, as the entire directory can be copied as one block (Ramakrishnan & Gehrke (2003)).

### Python example

```
PAGE_SZ = 20

class Page(object):

    def __init__(self):
        self.m = {}
        self.d = 0

    def full(self):
        return len(self.m) > PAGE_SZ

    def put(self, k, v):
        self.m[k] = v

    def get(self, k):
        return self.m.get(k)
```

```

class EH(object):

    def __init__(self):
        self.gd = 0
        p = Page()
        self.pp = [p]

    def get_page(self, k):
        h = hash(k)
        p = self.pp[ h & (( 1 << self.gd) -1)]
        return p

    def put(self, k, v):
        p = self.get_page(k)
        if p.full() and p.d == self.gd:
            self.pp *= 2
            self.gd += 1

        if p.full() and p.d < self.gd:
            p.put(k, v);
            p1 = Page()
            p2 = Page()
            for k2, v2 in p.m.items():
                h = hash(k2)
                h = h & ((1 << self.gd) -1)
                if (h >> p.d) & 1 == 1:
                    p2.put(k2, v2)
                else:
                    p1.put(k2, v2)
            for i, x in enumerate(self.pp):
                if x == p:
                    if (i >> p.d) & 1 == 1:
                        self.pp[i] = p2
                    else:
                        self.pp[i] = p1

            p2.d = p1.d = p.d + 1
        else:
            p.put(k, v)

    def get(self, k):
        p = self.get_page(k)
        return p.get(k)

if __name__ == "__main__":

```

```
eh = EH()
N = 10000
l = list(range(N))

import random
random.shuffle(l)
for x in l:
    eh.put(x, x)
print l

for i in range(N):
    print eh.get(i)
```

## Notes

- [1] Fagin, R.; Nievergelt, J.; Pippenger, N.; Strong, H. R. (September, 1979), "Extendible Hashing - A Fast Access Method for Dynamic Files", *ACM Transactions on Database Systems* **4** (3): 315–344, doi:10.1145/320083.320092

## References

- Fagin, R.; Nievergelt, J.; Pippenger, N.; Strong, H. R. (September, 1979), "Extendible Hashing - A Fast Access Method for Dynamic Files", *ACM Transactions on Database Systems* **4** (3): 315–344, doi:10.1145/320083.320092
- Ramakrishnan, R.; Gehrke, J. (2003), *Database Management Systems, 3rd Edition: Chapter 11, Hash-Based Indexing*, pp. 373–378

## External links

- Paul E. Black, Extendible hashing (<http://www.nist.gov/dads/HTML/extendibleHashing.html>) at the NIST Dictionary of Algorithms and Data Structures.
- Extendible Hashing (<http://www.isqa.unomaha.edu/haworth/isqa3300/fs009.htm>) at University of Nebraska
- Extendible Hashing notes (<http://www.csm.astate.edu/~rossa/datastruc/Extend.html>) at Arkansas State University
- Extendible hashing notes (<http://www.smckearney.com/adb/notes/lecture.extendible.hashing.pdf>)

## 2-choice hashing

---

**2-choice hashing**, also known as **2-choice chaining**, is a variant of a hash table in which keys are added by hashing with two hash functions. The key is put in the array position with the fewer (colliding) keys. Some collision resolution scheme is needed, unless keys are kept in buckets. The average-case cost of a successful search is  $O(2 + (m-1)/n)$ , where  $m$  is the number of keys and  $n$  is the size of the array. The most collisions is  $\log_2 \ln n + \theta(m/n)$  with high probability.

## Pearson hashing

---

**Pearson hashing**<sup>[1]</sup> is a hash function designed for fast execution on processors with 8-bit registers. Given an input consisting of any number of bytes, it produces as output a single byte that is strongly dependent<sup>[1]</sup> on every byte of the input. Its implementation requires only a few instructions, plus a 256-byte lookup table containing a permutation of the values 0 through 255.

This hash function is a CBC-MAC that uses an 8-bit random block cipher implemented via the permutation table. An 8-bit block cipher has negligible cryptographic security, so the Pearson hash function is not cryptographically strong; but it offers these benefits:

- It is extremely simple.
- It executes quickly on resource-limited processors.
- There is no simple class of inputs for which collisions (identical outputs) are especially likely.
- Given a small, privileged set of inputs (e.g., reserved words for a compiler), the permutation table can be adjusted so that those inputs yield distinct hash values, producing what is called a perfect hash function.

The algorithm can be described by the following pseudocode, which computes the hash of message  $C$  using the permutation table  $T$ :

```
h := 0
for each c in C loop
  index := h xor c
  h := T[index]
end loop
return h
```

## References

- [1] Pearson, Peter K. (June 1990), "Fast Hashing of Variable-Length Text Strings" (<http://portal.acm.org/citation.cfm?id=78978>), *Communications of the ACM* **33** (6): 677, doi:10.1145/78973.78978,

# Fowler–Noll–Vo hash function

---

**Fowler–Noll–Vo** is a non-cryptographic hash function created by Glenn Fowler, Landon Curt Noll, and Phong Vo.

The basis of the FNV hash algorithm was taken from an idea sent as reviewer comments to the IEEE POSIX P1003.2 committee by Glenn Fowler and Phong Vo in 1991. In a subsequent ballot round, Landon Curt Noll improved on their algorithm. Some people tried this hash and found that it worked rather well. In an EMail message to Landon, they named it the *Fowler/Noll/Vo* or FNV hash.<sup>[1]</sup>

## Overview

The current versions are FNV-1 and FNV-1a, which supply a means of creating non-zero **FNV offset basis**. FNV currently comes in 32-, 64-, 128-, 256-, 512-, and 1024-bit flavors. For pure FNV implementations, this is determined solely by the availability of **FNV primes** for the desired bit length; however, the FNV webpage discusses methods of adapting one of the above versions to a smaller length that may or may not be a power of two.<sup>[2][3]</sup>

The FNV hash algorithms and sample FNV source code<sup>[4]</sup> have been released into the public domain.<sup>[5]</sup>

FNV is not a cryptographic hash.

## The hash

One of FNV's key advantages is that it is very simple to implement. Start with an initial hash value of **FNV offset basis**. For each byte in the input, multiply *hash* by the **FNV prime**, then XOR it with the byte from the input. The alternate algorithm, FNV-1a, reverses the multiply and XOR steps.

## FNV-1 hash

The FNV-1 hash algorithm is as follows:<sup>[6]</sup>

```
hash = FNV_offset_basis
for each octet_of_data to be hashed
    hash = hash × FNV_prime
    hash = hash XOR octet_of_data
return hash
```

In the above pseudocode, all variables are unsigned integers. All variables, except for *octet\_of\_data*, have the same number of bits as the FNV hash. The variable, *octet\_of\_data*, is an 8 bit unsigned integer.

As an example, consider the 64-bit FNV-1 hash:

- All variables, except for *octet\_of\_data*, are 64-bit unsigned integers.
- The variable, *octet\_of\_data*, is an 8 bit unsigned integer.
- The **FNV\_offset\_basis** is the 64-bit **FNV offset basis** value: 14695981039346656037.
- The **FNV\_prime** is the 64-bit **FNV prime** value: 1099511628211.
- The multiply (indicated by the × symbol) returns the lower 64-bits of the product.
- The XOR is an 8-bit operation that modifies only the lower 8-bits of the hash value.
- The *hash* value returned is a 64-bit unsigned integer.

The values for **FNV prime** and **FNV offset basis** may be found in this table.<sup>[7]</sup>



## FNV-1a hash

The FNV-1a hash differs from the FNV-1 hash by only the order in which the multiply and XOR is performed: [8]

```
hash = FNV_offset_basis
for each octet_of_data to be hashed
    hash = hash XOR octet_of_data
    hash = hash × FNV_prime
return hash
```

The above pseudocode has the same assumptions that were noted for the FNV-1 pseudocode. The small change in order leads to much better avalanche characteristics. [9]

## Demonstrated weaknesses

The FNV hash has a few weaknesses identified by its authors that make it unsuitable as a cryptographic hash function. [10]

- **Speed of Computation** - As a hash designed primarily for hashtable and checksum use, FNV-1 and 1a were designed to be fast to compute. However, this same speed makes it more feasible for a computer to find hash values (and thus collisions) by brute-force.
- **Sticky State** - Being an iterative hash based primarily on multiplication and XOR, the algorithm is sensitive to the number zero. Specifically, if the hash value were to become zero at any point during calculation, and the next byte hashed was also all zeroes, the hash would not change. This makes colliding messages trivial to create given a message that results in a hash value of zero at some point in its calculation. Additional operations, such as the addition of a third constant prime on each step, can mitigate this but may have detrimental effects on avalanche effect or random distribution of hash values.
- **Diffusion** - The ideal secure hash function is one in which each byte of input has an equally-complex effect on every bit of the hash. In the FNV hash, the ones place (the rightmost bit) is always the XOR of the rightmost bit of every input byte. This can be mitigated by XOR-folding (computing a hash twice the desired length, and then XORing the bits in the "upper half" with the bits in the "lower half").

## Notes

- [1] FNV hash history (<http://www.isthe.com/chongo/tech/comp/fnv/index.html#history>)
- [2] Changing the FNV hash size - xor-folding (<http://www.isthe.com/chongo/tech/comp/fnv/index.html#xor-fold>)
- [3] Changing the FNV hash size - non-powers of 2 (<http://www.isthe.com/chongo/tech/comp/fnv/index.html#other-folding>)
- [4] <http://www.isthe.com/chongo/tech/comp/fnv/index.html#FNV-source>
- [5] FNV put into the public domain ([http://www.isthe.com/chongo/tech/comp/fnv/index.html#public\\_domain](http://www.isthe.com/chongo/tech/comp/fnv/index.html#public_domain))
- [6] The core of the FNV hash (<http://www.isthe.com/chongo/tech/comp/fnv/index.html#FNV-1>)
- [7] Parameters of the FNV-1 hash (<http://www.isthe.com/chongo/tech/comp/fnv/index.html#FNV-param>)
- [8] FNV-1a alternate algorithm (<http://www.isthe.com/chongo/tech/comp/fnv/index.html#FNV-1a>)
- [9] Avalanche (<http://murmurhash.googlepages.com/avalanche>)
- [10] The FNV Non-Cryptographic Hash Algorithm - Why is FNV Non-Cryptographic? (<http://tools.ietf.org/html/draft-eastlake-fnv-03#section-6.1>)

## External links

- Landon Curt Noll's webpage on FNV (<http://www.isthe.com/chongo/tech/comp/fnv/index.html>) (with table of base & prime parameters)
- Internet draft by Fowler, Noll, Vo, and Eastlake (<http://tools.ietf.org/html/draft-eastlake-fnv>) (2011, work in progress)

## Bitstate hashing

---

**Bitstate hashing** is a hashing method invented in 1968 by Morris.<sup>[1]</sup> It is used for state hashing, where each state (e.g. of an automaton) is represented by a number and it is passed to some hash function.

The result of the function is then taken as the index to an array of bits (a bit-field), where one looks for 1 if the state was already seen before or stores 1 by itself if not.

It usually serves as a yes–no technique without a need of storing whole state bit representation.

A shortcoming of this framework is losing precision like in other hashing techniques. Hence some tools use this technique with more than one hash function so that the bit-field gets widened by the number of used functions, each having its own row. And even after all functions return values (the indices) point to fields with contents equal to 1, the state may be uttered as visited with much higher probability.

## Use

- It is utilized in SPIN model checker for decision whether a state was already visited by nested-depth-first search searching algorithm or not. They mention savings of 98% of memory in the case of using one hash function (175 MB to 3 MB) and 92% when two hash functions are used (13 MB). The state coverage dropped to 97% in the former case.<sup>[2]</sup>

## References

[1] Morris, R. (1968). *Scatter Storage Techniques*

[2] Holzmann, G. J. (2003) Addison Wesley. *Spin Model Checker, The: Primer and Reference Manual*

# Bloom filter

A **Bloom filter**, conceived by Burton Howard Bloom in 1970,<sup>[1]</sup> is a space-efficient probabilistic data structure that is used to test whether an element is a member of a set. False positive retrieval results are possible, but false negatives are not; i.e. a query returns either "inside set (may be wrong)" or "definitely not in set". Elements can be added to the set, but not removed (though this can be addressed with a counting filter). The more elements that are added to the set, the larger the probability of false positives.

## Algorithm description

An **empty Bloom filter** is a bit array of  $m$  bits, all set to 0. There must also be  $k$  different hash functions defined, each of which maps or hashes some set element to one of the  $m$  array positions with a uniform random distribution.

To **add** an element, feed it to each of the  $k$  hash functions to get  $k$  array positions. Set the bits at all these positions to 1.

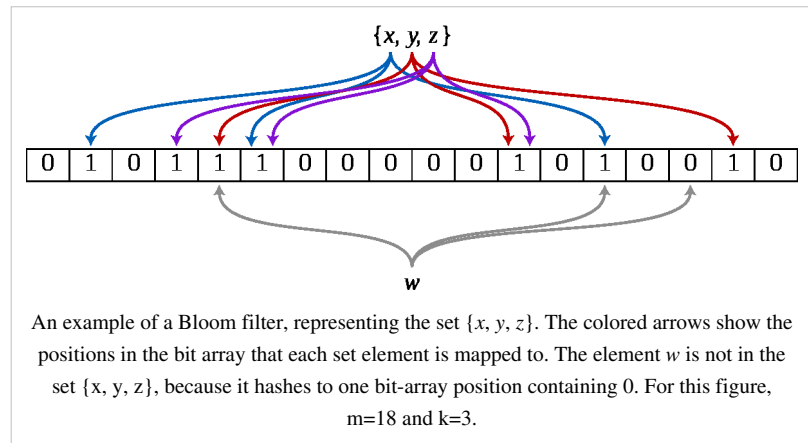
To **query** for an element (test whether

it is in the set), feed it to each of the  $k$  hash functions to get  $k$  array positions. If any of the bits at these positions are 0, the element is definitely not in the set – if it were, then all the bits would have been set to 1 when it was inserted. If all are 1, then either the element is in the set, or the bits have by chance been set to 1 during the insertion of other elements, resulting in a false positive. In a simple bloom filter, there is no way to distinguish between the two cases, but more advanced techniques can address this problem.

The requirement of designing  $k$  different independent hash functions can be prohibitive for large  $k$ . For a good hash function with a wide output, there should be little if any correlation between different bit-fields of such a hash, so this type of hash can be used to generate multiple "different" hash functions by slicing its output into multiple bit fields. Alternatively, one can pass  $k$  different initial values (such as 0, 1, ...,  $k - 1$ ) to a hash function that takes an initial value; or add (or append) these values to the key. For larger  $m$  and/or  $k$ , independence among the hash functions can be relaxed with negligible increase in false positive rate (Dillinger & Manolios (2004a), Kirsch & Mitzenmacher (2006)). Specifically, Dillinger & Manolios (2004b) show the effectiveness of deriving the  $k$  indices using enhanced double hashing or triple hashing, variants of double hashing that are effectively simple random number generators seeded with the two or three hash values.

Removing an element from this simple Bloom filter is impossible because false negatives are not permitted. An element maps to  $k$  bits, and although setting any one of those  $k$  bits to zero suffices to remove the element, it also results in removing any other elements that happen to map onto that bit. Since there is no way of determining whether any other elements have been added that affect the bits for an element to be removed, clearing any of the bits would introduce the possibility for false negatives.

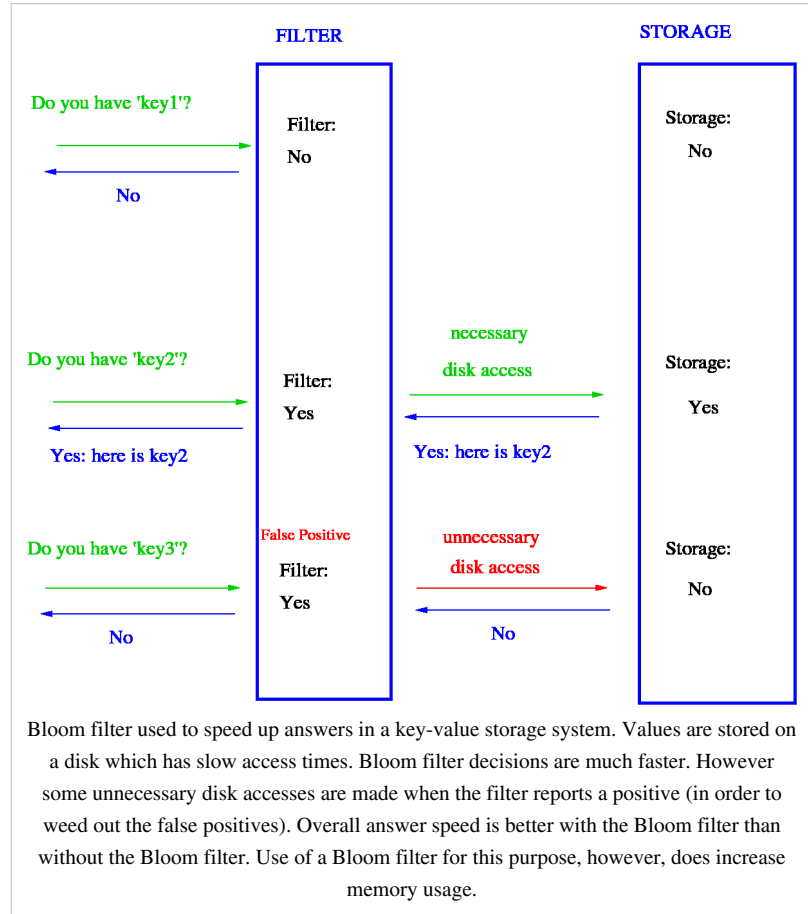
One-time removal of an element from a Bloom filter can be simulated by having a second Bloom filter that contains items that have been removed. However, false positives in the second filter become false negatives in the composite filter, which may be undesirable. In this approach re-adding a previously removed item is not possible, as one would have to remove it from the "removed" filter.



However, it is often the case that all the keys are available but are expensive to enumerate (for example, requiring many disk reads). When the false positive rate gets too high, the filter can be regenerated; this should be a relatively rare event.

## Space and time advantages

While risking false positives, Bloom filters have a strong space advantage over other data structures for representing sets, such as self-balancing binary search trees, tries, hash tables, or simple arrays or linked lists of the entries. Most of these require storing at least the data items themselves, which can require anywhere from a small number of bits, for small integers, to an arbitrary number of bits, such as for strings (tries are an exception, since they can share storage between elements with equal prefixes). Linked structures incur an additional linear space overhead for pointers. A Bloom filter with 1% error and an optimal value of  $k$ , in contrast, requires only about 9.6 bits per element — regardless of the size of the elements. This advantage comes partly from its compactness, inherited from arrays, and partly from its probabilistic nature. If a 1% false-positive rate seems too high, adding about 4.8 bits per element decreases it by ten times.



However, if the number of potential values is small and many of them can be in the set, the Bloom filter is easily surpassed by the deterministic bit array, which requires only one bit for each potential element. Note also that hash tables gain a space and time advantage if they begin ignoring collisions and store only whether each bucket contains an entry; in this case, they have effectively become Bloom filters with  $k = 1$ .

Bloom filters also have the unusual property that the time needed either to add items or to check whether an item is in the set is a fixed constant,  $O(k)$ , completely independent of the number of items already in the set. No other constant-space set data structure has this property, but the average access time of sparse hash tables can make them faster in practice than some Bloom filters. In a hardware implementation, however, the Bloom filter shines because its  $k$  lookups are independent and can be parallelized.

To understand its space efficiency, it is instructive to compare the general Bloom filter with its special case when  $k = 1$ . If  $k = 1$ , then in order to keep the false positive rate sufficiently low, a small fraction of bits should be set, which means the array must be very large and contain long runs of zeros. The information content of the array relative to its size is low. The generalized Bloom filter ( $k$  greater than 1) allows many more bits to be set while still maintaining a low false positive rate; if the parameters ( $k$  and  $m$ ) are chosen well, about half of the bits will be set, and these will be apparently random, minimizing redundancy and maximizing information content.

## Probability of false positives

Assume that a hash function selects each array position with equal probability. If  $m$  is the number of bits in the array, and  $k$  is the number of hash functions, then the probability that a certain bit is not set to 1 by a certain hash function during the insertion of an element is then

$$1 - \frac{1}{m}.$$

The probability that it is not set to 1 by any of the hash functions is

$$\left(1 - \frac{1}{m}\right)^k.$$

If we have inserted  $n$  elements, the probability that a certain bit is still 0 is

$$\left(1 - \frac{1}{m}\right)^{kn};$$

the probability that it is 1 is therefore

$$1 - \left(1 - \frac{1}{m}\right)^{kn}.$$

Now test membership of an element that is not in the set. Each of the  $k$  array positions computed by the hash functions is 1 with a probability as above. The probability of all of them being 1, which would cause the algorithm to erroneously claim that the element is in the set, is often given as

$$\left(1 - \left[1 - \frac{1}{m}\right]^{kn}\right)^k \approx \left(1 - e^{-kn/m}\right)^k.$$

This is not strictly correct as it assumes independence for the probabilities of each bit being set. However, assuming it is a close approximation we have that the probability of false positives decreases as  $m$  (the number of bits in the array) increases, and increases as  $n$  (the number of inserted elements) increases. For a given  $m$  and  $n$ , the value of  $k$  (the number of hash functions) that minimizes the probability is

$$\frac{m}{n} \ln 2 \approx 0.7 \frac{m}{n},$$

which gives the false positive probability of

$$2^{-k} \approx 0.6185^{m/n}.$$

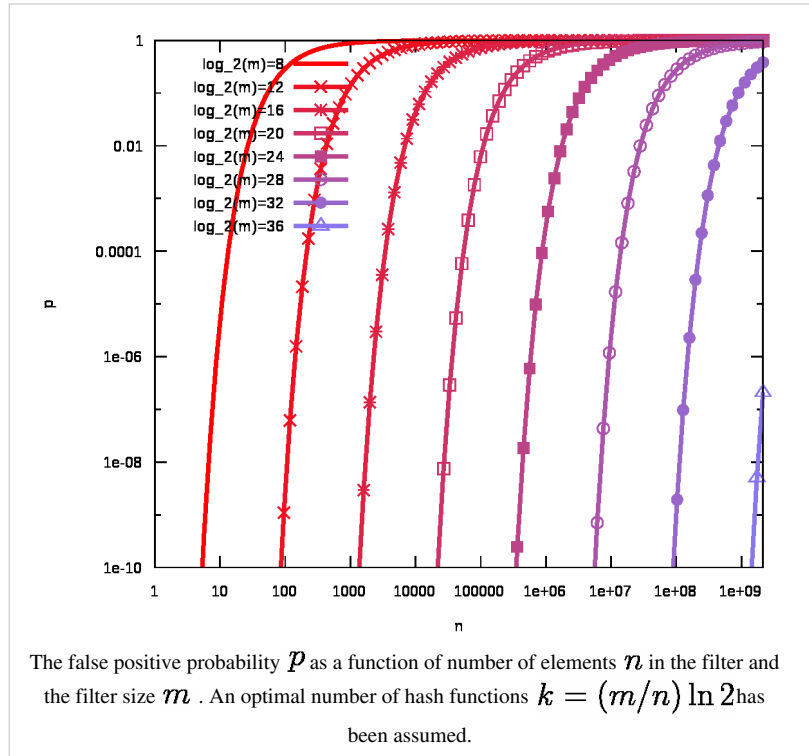
The required number of bits  $m$ , given  $n$  (the number of inserted elements) and a desired false positive probability  $p$  (and assuming the optimal value of  $k$  is used) can be computed by substituting the optimal value of  $k$  in the probability expression above:

$$p = \left(1 - e^{-(m/n \ln 2)n/m}\right)^{(m/n \ln 2)}$$

which can be simplified to:

$$\ln p = -\frac{m}{n} (\ln 2)^2.$$

This results in:



$$m = -\frac{n \ln p}{(\ln 2)^2}.$$

This means that for a given false positive probability  $p$ , the length of a Bloom filter  $m$  is proportionate to the number of elements being filtered  $n$ .<sup>[2]</sup> While the above formula is asymptotic (i.e. applicable as  $m, n \rightarrow \infty$ ), the agreement with finite values of  $m, n$  is also quite good; the false positive probability for a finite bloom filter with  $m$  bits,  $n$  elements, and  $k$  hash functions is at most

$$\left(1 - e^{-k(n+0.5)/(m-1)}\right)^k.$$

So we can use the asymptotic formula if we pay a penalty for at most half an extra element and at most one fewer bit.<sup>[3]</sup>

## Interesting properties

- Unlike sets based on hash tables, any Bloom filter can represent the entire universe of elements. In this case, all bits are 1. Another consequence of this property is that **add** never fails due to the data structure "filling up." However, the false positive rate increases steadily as elements are added until all bits in the filter are set to 1, so a negative value is never returned. At this point, the Bloom filter completely ceases to differentiate between differing inputs, and is functionally useless.
- Union and intersection of Bloom filters with the same size and set of hash functions can be implemented with bitwise OR and AND operations, respectively. The union operation on Bloom filters is lossless in the sense that the resulting Bloom filter is the same as the Bloom filter created from scratch using the union of the two sets. The intersect operation satisfies a weaker property: the false positive probability in the resulting Bloom filter is at most the false-positive probability in one of the constituent Bloom filters, but may be larger than the false positive probability in the Bloom filter created from scratch using the intersection of the two sets.
- Some kinds of superimposed code can be seen as a Bloom filter implemented with physical edge-notched cards.

## Examples

Google BigTable uses Bloom filters to reduce the disk lookups for non-existent rows or columns. Avoiding costly disk lookups considerably increases the performance of a database query operation.<sup>[4]</sup>

The Squid Web Proxy Cache uses Bloom filters for cache digests<sup>[5]</sup><sup>[6]</sup>

The Venti archival storage system uses Bloom filters to detect previously stored data.<sup>[7]</sup>

The SPIN model checker uses Bloom filters to track the reachable state space for large verification problems.<sup>[8]</sup>

The Google Chrome web browser uses Bloom filters to speed up its Safe Browsing service.<sup>[9]</sup>

## Alternatives

Classic Bloom filters use  $1.44 \log_2(1/\epsilon)$  bits of space per inserted key, where  $\epsilon$  is the false positive rate of the Bloom filter. However, the space that is strictly necessary for any data structure playing the same role as a Bloom filter is only  $\log_2(1/\epsilon)$  per key (Pagh, Pagh & Rao 2005). Hence Bloom filters use 44% more space than a hypothetical equivalent optimal data structure. The number of hash functions used to achieve a given false positive rate  $\epsilon$  is proportional to  $1/\epsilon$  which is not optimal as it has been proved that an optimal data structure would need only a constant number of hash functions independent of the false positive rate.

Stern & Dill (1996) describe a probabilistic structure based on hash tables, hash compaction, which Dillinger & Manolios (2004b) identify as significantly more accurate than a Bloom filter when each is configured optimally. Dillinger and Manolios, however, point out that the reasonable accuracy of any given Bloom filter over a wide range of numbers of additions makes it attractive for probabilistic enumeration of state spaces of unknown size. Hash

compaction is, therefore, attractive when the number of additions can be predicted accurately; however, despite being very fast in software, hash compaction is poorly suited for hardware because of worst-case linear access time.

Putze, Sanders & Singler (2007) have studied some variants of Bloom filters that are either faster or use less space than classic Bloom filters. The basic idea of the fast variant is to locate the  $k$  hash values associated with each key into one or two blocks having the same size as processor's memory cache blocks (usually 64 bytes). This will presumably improve performance by reducing the number of potential memory cache misses. The proposed variants have however the drawback of using about 32% more space than classic Bloom filters.

The space efficient variant relies on using a single hash function that generates for each key a value in the range  $[0, n/\epsilon]$  where  $\epsilon$  is the requested false positive rate. The sequence of values is then sorted and compressed using Golomb coding (or some other compression technique) to occupy a space close to  $n \log_2(1/\epsilon)$  bits. To query the Bloom filter for a given key, it will suffice to check if its corresponding value is stored in the Bloom filter. Decompressing the whole Bloom filter for each query would make this variant totally unusable. To overcome this problem the sequence of values is divided into small blocks of equal size that are compressed separately. At query time only half a block will need to be decompressed on average. Because of decompression overhead, this variant may be slower than classic Bloom filters but this may be compensated by the fact that a single hash function need to be computed.

Another alternative to classic Bloom filter is the one based on space efficient variants of cuckoo hashing. In this case once the hash table is constructed, the keys stored in the hash table are replaced with short signatures of the keys. Those signatures are strings of bits computed using a hash function applied on the keys.

## Extensions and applications

### Counting filters

Counting filters provide a way to implement a *delete* operation on a Bloom filter without recreating the filter afresh. In a counting filter the array positions (buckets) are extended from being a single bit, to an  $n$ -bit counter. In fact, regular Bloom filters can be considered as counting filters with a bucket size of one bit. Counting filters were introduced by Fan et al. (1998).

The insert operation is extended to *increment* the value of the buckets and the lookup operation checks that each of the required buckets is non-zero. The delete operation, obviously, then consists of decrementing the value of each of the respective buckets.

Arithmetic overflow of the buckets is a problem and the buckets should be sufficiently large to make this case rare. If it does occur then the increment and decrement operations must leave the bucket set to the maximum possible value in order to retain the properties of a Bloom filter.

The size of counters is usually 3 or 4 bits. Hence counting Bloom filters use 3 to 4 times more space than static Bloom filters. In theory, an optimal data structure equivalent to a counting Bloom filter should not use more space than a static Bloom filter.

Another issue with counting filters is limited scalability. Because the counting Bloom filter table cannot be expanded, the maximal number of keys to be stored simultaneously in the filter must be known in advance. Once the designed capacity of the table is exceeded the false positive rate will grow rapidly as more keys are inserted.

Bonomi et al. (2006) introduced a data structure based on d-left hashing that is functionally equivalent but uses approximately half as much space as counting Bloom filters. The scalability issue does not occur in this data structure. Once the designed capacity is exceeded, the keys could be reinserted in a new hash table of double size.

The space efficient variant by Putze, Sanders & Singler (2007) could also be used to implement counting filters by supporting insertions and deletions.

## Data synchronization

Bloom filters can be used for approximate data synchronization as in Byers et al. (2004). Counting Bloom filters can be used to approximate the number of differences between two sets and this approach is described in Agarwal & Trachtenberg (2006).

## Bloomier filters

Chazelle et al. (2004) designed a generalization of Bloom filters that could associate a value with each element that had been inserted, implementing an associative array. Like Bloom filters, these structures achieve a small space overhead by accepting a small probability of false positives. In the case of "Bloomier filters", a *false positive* is defined as returning a result when the key is not in the map. The map will never return the wrong value for a key that *is* in the map.

The simplest Bloomier filter is near-optimal and fairly simple to describe. Suppose initially that the only possible values are 0 and 1. We create a pair of Bloom filters  $A_0$  and  $B_0$  which contain, respectively, all keys mapping to 0 and all keys mapping to 1. Then, to determine which value a given key maps to, we look it up in both filters. If it is in neither, then the key is not in the map. If the key is in  $A_0$  but not  $B_0$ , then it does not map to 1, and has a high probability of mapping to 0. Conversely, if the key is in  $B_0$  but not  $A_0$ , then it does not map to 0 and has a high probability of mapping to 1.

A problem arises, however, when *both* filters claim to contain the key. We never insert a key into both, so one or both of the filters is lying (producing a false positive), but we don't know which. To determine this, we have another, smaller pair of filters  $A_1$  and  $B_1$ .  $A_1$  contains keys that map to 0 and which are false positives in  $B_0$ ;  $B_1$  contains keys that map to 1 and which are false positives in  $A_0$ . But whenever  $A_0$  and  $B_0$  both produce positives, at most one of these cases must occur, and so we simply have to determine which if any of the two filters  $A_1$  and  $B_1$  contains the key, another instance of our original problem.

It may so happen again that both filters produce a positive; we apply the same idea recursively to solve this problem. Because each pair of filters only contains keys that are in the map *and* produced false positives on all previous filter pairs, the number of keys is extremely likely to quickly drop to a very small quantity that can be easily stored in an ordinary deterministic map, such as a pair of small arrays with linear search. Moreover, the average total search time is low, because almost all queries will be resolved by the first pair, almost all remaining queries by the second pair, and so on. The total space required is in practice independent of  $n$ , and is almost entirely occupied by the first filter pair.

Now that we have the structure and a search algorithm, we also need to know how to insert new key/value pairs. The program must not attempt to insert the same key with both values. If the value is 0, insert the key into  $A_0$  and then test if the key is in  $B_0$ . If so, this is a false positive for  $B_0$ , and the key must also be inserted into  $A_1$  recursively in the same manner. If we reach the last level, we simply insert it. When the value is 1, the operation is similar but with  $A$  and  $B$  reversed.

Now that we can map a key to the value 0 or 1, how does this help us map to general values? This is simple. We create a single such Bloomier filter for each bit of the result. If the values are large, we can instead map keys to hash values that can be used to retrieve the actual values. The space required for a Bloomier filter with  $n$ -bit values is typically slightly more than the space for  $2n$  Bloom filters.

A very simple way to implement Bloomier filters is by means of minimal perfect hashing. A minimal perfect hash function  $h$  is first generated for the set of  $n$  keys. Then an array is filled with  $n$  pairs (signature,value) associated with each key at the positions given by function  $h$  when applied on each key. The signature of a key is a string of  $r$  bits computed by applying a hash function  $g$  of range  $2^r$  on the key. The value of  $r$  is chosen such that  $2^r \geq 1/\epsilon$ , where  $\epsilon$  is the requested false positive rate. To query for a given key, hash function  $h$  is first applied on the key. This will give a position into the array from which we retrieve a pair (signature,value). Then we compute the



signature of the key using function  $g$ . If the computed signature is the same as retrieved signature we return the retrieved value. The probability of false positive is  $1/2^r$ .

Another alternative to implement static bloomier and bloom filters based on matrix solving has been simultaneously proposed in Porat (2008), Dietzfelbinger & Pagh (2008) and Charles & Chellapilla (2008). The space usage of this method is optimal as it needs only  $\log_2(\epsilon)$  bits per key for a bloom filter. However time to generate the bloom or bloomier filter can be very high. The generation time can be reduced to a reasonable value at the price of a small increase in space usage.

Dynamic Bloomier filters have been studied by Mortensen, Pagh & Pătraşcu (2005). They proved that any dynamic Bloomier filter needs at least around  $\log(l)$  bits per key where  $l$  is the length of the key. A simple dynamic version of Bloomier filters can be implemented using two dynamic data structures. Let the two data structures be noted  $S1$  and  $S2$ .  $S1$  will store keys with their associated data while  $S2$  will only store signatures of keys with their associated data. Those signatures are simply hash values of keys in the range  $[0, n/\epsilon]$  where  $n$  is the maximal number of keys to be stored in the Bloomier filter and  $\epsilon$  is the requested false positive rate. To insert a key in the Bloomier filter, its hash value is first computed. Then the algorithm checks if a key with the same hash value already exists in  $S2$ . If this is not the case, the hash value is inserted in  $S2$  along with data associated with the key. If the same hash value already exists in  $S2$  then the key is inserted into  $S1$  along with its associated data. The deletion is symmetric: if the key already exists in  $S1$  it will be deleted from there, otherwise the hash value associated with the key is deleted from  $S2$ . An issue with this algorithm is on how to store efficiently  $S1$  and  $S2$ . For  $S1$  any hash algorithm can be used. To store  $S2$  golomb coding could be applied to compress signatures to use a space close to  $\log 2(1/\epsilon)$  per key.

### Compact approximators

Boldi & Vigna (2005) proposed a lattice-based generalization of Bloom filters. A **compact approximator** associates to each key an element of a lattice (the standard Bloom filters being the case of the Boolean two-element lattice). Instead of a bit array, they have an array of lattice elements. When adding a new association between a key and an element of the lattice, they maximize the current content of the  $k$  array locations associated to the key with the lattice element. When reading the value associated to a key, they minimize the values found in the  $k$  locations associated to the key. The resulting value approximates from above the original value.

### Stable Bloom filters

Deng & Rafiei (2006) proposed Stable Bloom filters as a variant of Bloom filters for streaming data. The idea is that since there is no way to store the entire history of a stream (which can be infinite), Stable Bloom filters continuously evict stale information to make room for more recent elements. Since stale information is evicted, the Stable Bloom filter introduces false negatives, which do not appear in traditional bloom filters. The authors show that a tight upper bound of false positive rates is guaranteed, and the method is superior to standard bloom filters in terms of false positive rates and time efficiency when a small space and an acceptable false positive rate are given.

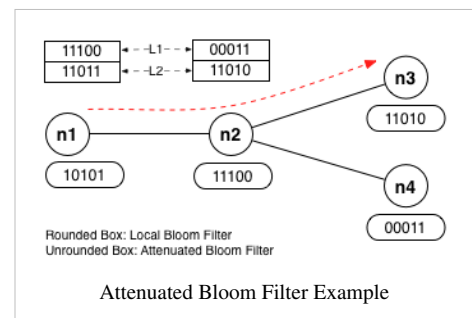
## Scalable Bloom filters

Almeida et al. (2007) proposed a variant of Bloom filters that can adapt dynamically to the number of elements stored, while assuring a minimum false positive probability. The technique is based on sequences of standard bloom filters with increasing capacity and tighter false positive probabilities, so as to ensure that a maximum false positive probability can be set beforehand, regardless of the number of elements to be inserted.

## Attenuated Bloom filters

An attenuated bloom filter of depth  $D$  can be viewed as an array of  $D$  normal bloom filters. In the context of service discovery in a network, each node stores regular and attenuated bloom filters locally. The regular or local bloom filter indicates which services are offered by the node itself. The attenuated filter of level  $i$  indicates which services can be found on nodes that are  $i$ -hops away from the current node. The  $i$ -th value is constructed by taking a union of local bloom filters for nodes  $i$ -hops away from the node.<sup>[10]</sup>

Lets take a small network shown on the graph below as an example. Say we are searching for a service A whose id hashes to bits 0,1, and 3 (pattern 11010). Let  $n1$  node to be the starting point. First, we check whether service A is offered by  $n1$  by checking its local filter. Since the patterns don't match, we check the attenuated bloom filter in order to determine which node should be the next hop. We see that  $n2$  doesn't offer service A but lies on the path to nodes that do. Hence, we move to  $n2$  and repeat the same procedure. We quickly find that  $n3$  offers the service, and hence the destination is located.<sup>[11]</sup>



By using attenuated Bloom filters consisting of multiple layers, services at more than one hop distance can be discovered while avoiding saturation of the Bloom filter by attenuating (shifting out) bits set by sources further away.<sup>[10]</sup>

## Notes

- [1] Donald Knuth. "[The Art of Computer Programming (<http://www-cs-faculty.stanford.edu/~knuth/err3.txt>)], Errata for Volume 3 (2nd ed.)". .
- [2] Starobinski, Trachtenberg & Agarwal (2003).
- [3] Goel & Gupta (2010).
- [4] (Chang et al. 2006).
- [5] <http://wiki.squid-cache.org/SquidFaq/CacheDigests>
- [6] Wessels, Duane (January 2004), "10.7 Cache Digests", *Squid: The Definitive Guide* (1st ed.), O'Reilly Media, p. 172, ISBN 0-596-00162-2, "Cache Digests are based on a technique first published by Pei Cao, called Summary Cache. The fundamental idea is to use a Bloom filter to represent the cache contents."
- [7] <http://plan9.bell-labs.com/magic/man2html/8/venti>
- [8] <http://spinroot.com/>
- [9] [http://src.chromium.org/viewvc/chrome/trunk/src/chrome/browser/safe\\_browsing/bloom\\_filter.h?view=markup](http://src.chromium.org/viewvc/chrome/trunk/src/chrome/browser/safe_browsing/bloom_filter.h?view=markup)
- [10] Koucheryavy et al. (2009)
- [11] Kubiawicz et al. (2000)

## References

- Koucheryavy, Y.; Giambene, G.; Staehle, D.; Barcelo-Arroyo, F.; Braun, T.; Siris, V. (2009), "Traffic and QoS Management in Wireless Multimedia Networks", *COST 290 Final Report* (USA): 111
- Kubiawicz, J.; Bindel, D.; Czerwinski, Y.; Geels, S.; Eaton, D.; Gummadi, R.; Rhea, S.; Weatherspoon, H. et al. (2000), "Oceanstore: An architecture for global-scale persistent storage" (<http://ftp.csd.uwo.ca/courses/CS9843b/papers/OceanStore.pdf>), *ACM SIGPLAN Notices* (USA): 190–201
- Agarwal, Sachin; Trachtenberg, Ari (2006), "Approximating the number of differences between remote sets" (<http://www.deutsche-telekom-laboratories.de/~agarwals/publications/itw2006.pdf>), *IEEE Information Theory Workshop* (Punta del Este, Uruguay): 217, doi:10.1109/ITW.2006.1633815, ISBN 1-4244-0035-X
- Ahmadi, Mahmood; Wong, Stephan (2007), "A Cache Architecture for Counting Bloom Filters" ([http://www.ieeexplore.ieee.org/xpls/abs\\_all.jsp?isnumber=4444031&arnumber=4444089&count=113&index=57](http://www.ieeexplore.ieee.org/xpls/abs_all.jsp?isnumber=4444031&arnumber=4444089&count=113&index=57)), *15th international Conference on Networks (ICON-2007)*, pp. 218, doi:10.1109/ICON.2007.4444089, ISBN 978-1-4244-1229-7
- Almeida, Paulo; Baquero, Carlos; Preguica, Nuno; Hutchison, David (2007), "Scalable Bloom Filters" (<http://gsd.di.uminho.pt/members/cbm/ps/dbloom.pdf>), *Information Processing Letters* **101** (6): 255–261, doi:10.1016/j.ipl.2006.10.007
- Byers, John W.; Considine, Jeffrey; Mitzenmacher, Michael; Rost, Stanislav (2004), "Informed content delivery across adaptive overlay networks", *IEEE/ACM Transactions on Networking* **12** (5): 767, doi:10.1109/TNET.2004.836103
- Bloom, Burton H. (1970), "Space/time trade-offs in hash coding with allowable errors", *Communications of the ACM* **13** (7): 422–426, doi:10.1145/362686.362692
- Boldi, Paolo; Vigna, Sebastiano (2005), "Mutable strings in Java: design, implementation and lightweight text-search algorithms", *Science of Computer Programming* **54** (1): 3–23, doi:10.1016/j.scico.2004.05.003
- Bonomi, Flavio; Mitzenmacher, Michael; Panigrahy, Rina; Singh, Sushil; Varghese, George (2006), "An Improved Construction for Counting Bloom Filters" (<http://theory.stanford.edu/~rinap/papers/esa2006b.pdf>), *Algorithms – ESA 2006, 14th Annual European Symposium*, Lecture Notes in Computer Science, **4168**, pp. 684–695, doi:10.1007/11841036\_61, ISBN 978-3-540-38875-3
- Broder, Andrei; Mitzenmacher, Michael (2005), "Network Applications of Bloom Filters: A Survey" (<http://www.eecs.harvard.edu/~michaelm/postscripts/im2005b.pdf>), *Internet Mathematics* **1** (4): 485–509, doi:10.1080/15427951.2004.10129096
- Chang, Fay; Dean, Jeffrey; Ghemawat, Sanjay; Hsieh, Wilson; Wallach, Deborah; Burrows, Mike; Chandra, Tushar; Fikes, Andrew et al. (2006), "Bigtable: A Distributed Storage System for Structured Data" (<http://research.google.com/archive/bigtable.html>), *Seventh Symposium on Operating System Design and Implementation*
- Charles, Denis; Chellapilla, Kumar (2008), "Bloomier Filters: A second look", *The Computing Research Repository (CoRR)*, arXiv:0807.0928
- Chazelle, Bernard; Kilian, Joe; Rubinfeld, Ronitt; Tal, Ayellet (2004), "The Bloomier filter: an efficient data structure for static support lookup tables" (<http://www.ee.technion.ac.il/~ayellet/Ps/nelson.pdf>), *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 30–39
- Cohen, Saar; Matias, Yossi (2003), "Spectral Bloom Filters" (<http://www.sigmod.org/sigmod03/eeproceedings/papers/r09p02.pdf>), *Proceedings of the 2003 ACM SIGMOD International Conference on Management of Data*, pp. 241–252, doi:10.1145/872757.872787, ISBN 158113634X
- Deng, Fan; Rafiei, Davood (2006), "Approximately Detecting Duplicates for Streaming Data using Stable Bloom Filters" (<http://webdocs.cs.ualberta.ca/~drafie/papers/DupDet06Sigmod.pdf>), *Proceedings of the ACM SIGMOD Conference*, pp. 25–36
- Dharmapurikar, Sarang; Song, Haoyu; Turner, Jonathan; Lockwood, John (2006), "Fast packet classification using Bloom filters" (<http://www.arl.wustl.edu/~sarang/ancs6819-dharmapurikar.pdf>), *Proceedings of the*

- 2006 ACM/IEEE Symposium on Architecture for Networking and Communications Systems, pp. 61–70, doi:10.1145/1185347.1185356, ISBN 1595935800
- Dietzfelbinger, Martin; Pagh, Rasmus (2008), "Succinct Data Structures for Retrieval and Approximate Membership", *The Computing Research Repository (CoRR)*, arXiv:0803.3693
  - Dillinger, Peter C.; Manolios, Panagiotis (2004a), "Fast and Accurate Bitstate Verification for SPIN" (<http://www.ccs.neu.edu/home/pete/research/spin-3spin.html>), *Proceedings of the 11th International Spin Workshop on Model Checking Software*, Springer-Verlag, Lecture Notes in Computer Science 2989
  - Dillinger, Peter C.; Manolios, Panagiotis (2004b), "Bloom Filters in Probabilistic Verification" (<http://www.ccs.neu.edu/home/pete/research/bloom-filters-verification.html>), *Proceedings of the 5th International Conference on Formal Methods in Computer-Aided Design*, Springer-Verlag, Lecture Notes in Computer Science 3312
  - Donnet, Benoit; Baynat, Bruno; Friedman, Timur (2006), "Retouched Bloom Filters: Allowing Networked Applications to Flexibly Trade Off False Positives Against False Negatives" ([http://adetti.iscte.pt/events/CONEXT06/Conext06\\_Proceedings/papers/13.html](http://adetti.iscte.pt/events/CONEXT06/Conext06_Proceedings/papers/13.html)), *CoNEXT 06 – 2nd Conference on Future Networking Technologies*
  - Eppstein, David; Goodrich, Michael T. (2007), "Space-efficient straggler identification in round-trip data streams via Newton's identities and invertible Bloom filters", *Algorithms and Data Structures, 10th International Workshop, WADS 2007*, Springer-Verlag, Lecture Notes in Computer Science 4619, pp. 637–648, arXiv:0704.3313
  - Fan, Li; Cao, Pei; Almeida, Jussara; Broder, Andrei (2000), "Summary Cache: A Scalable Wide-Area Web Cache Sharing Protocol", *IEEE/ACM Transactions on Networking* **8** (3): 281–293, doi:10.1109/90.851975. A preliminary version appeared at SIGCOMM '98.
  - Goel, Ashish; Gupta, Pankaj (2010), "Small subset queries and bloom filters using ternary associative memories, with applications", *ACM Sigmetrics 2010* **38**: 143, doi:10.1145/1811099.1811056
  - Kirsch, Adam; Mitzenmacher, Michael (2006), "Less Hashing, Same Performance: Building a Better Bloom Filter" (<http://www.eecs.harvard.edu/~kirsch/pubs/bbbf/esa06.pdf>), in Azar, Yossi; Erlebach, Thomas, *Algorithms – ESA 2006, 14th Annual European Symposium*, Lecture Notes in Computer Science, **4168**, Springer-Verlag, Lecture Notes in Computer Science 4168, pp. 456–467, doi:10.1007/11841036, ISBN 978-3-540-38875-3
  - Mortensen, Christian Worm; Pagh, Rasmus; Pătraşcu, Mihai (2005), "On dynamic range reporting in one dimension", *Proceedings of the Thirty-seventh Annual ACM Symposium on Theory of Computing*, pp. 104–111, doi:10.1145/1060590.1060606, ISBN 1581139608
  - Pagh, Anna; Pagh, Rasmus; Rao, S. Srinivasa (2005), "An optimal Bloom filter replacement" (<http://www.it-c.dk/people/pagh/papers/bloom.pdf>), *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 823–829
  - Porat, Ely (2008), "An Optimal Bloom Filter Replacement Based on Matrix Solving", *The Computing Research Repository (CoRR)*, arXiv:0804.1845
  - Putze, F.; Sanders, P.; Singler, J. (2007), "Cache-, Hash- and Space-Efficient Bloom Filters" (<http://algo2.iti.uni-karlsruhe.de/singler/publications/cacheefficientbloomfilters-wea2007.pdf>), in Demetrescu, Camil, *Experimental Algorithms, 6th International Workshop, WEA 2007*, Lecture Notes in Computer Science, **4525**, Springer-Verlag, Lecture Notes in Computer Science 4525, pp. 108–121, doi:10.1007/978-3-540-72845-0, ISBN 978-3-540-72844-3
  - Sethumadhavan, Simha; Desikan, Rajagopalan; Burger, Doug; Moore, Charles R.; Keckler, Stephen W. (2003), "Scalable hardware memory disambiguation for high ILP processors" (<http://www.cs.utexas.edu/users/simha/publications/lsq.pdf>), *36th Annual IEEE/ACM International Symposium on Microarchitecture, 2003, MICRO-36*, pp. 399–410, doi:10.1109/MICRO.2003.1253244, ISBN 0-7695-2043-X
  - Shanmugasundaram, Kulesh; Brönnimann, Hervé; Memon, Nasir (2004), "Payload attribution via hierarchical Bloom filters", *Proceedings of the 11th ACM Conference on Computer and Communications Security*, pp. 31–41,

doi:10.1145/1030083.1030089, ISBN 1581139616

- Starobinski, David; Trachtenberg, Ari; Agarwal, Sachin (2003), "Efficient PDA Synchronization", *IEEE Transactions on Mobile Computing* **2** (1): 40, doi:10.1109/TMC.2003.1195150
- Stern, Ulrich; Dill, David L. (1996), "A New Scheme for Memory-Efficient Probabilistic Verification" (<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.47.4101>), *Proceedings of Formal Description Techniques for Distributed Systems and Communication Protocols, and Protocol Specification, Testing, and Verification: IFIP TC6/WG6.1 Joint International Conference*, Chapman & Hall, IFIP Conference Proceedings, pp. 333–348

## External links

- Table of false-positive rates for different configurations (<http://www.cs.wisc.edu/~cao/papers/summary-cache/node8.html>) from a University of Wisconsin–Madison website
- Interactive Processing demonstration (<http://tr.ashcan.org/2008/12/bloomers.html>) from ashcan.org
- "More Optimal Bloom Filters," Ely Porat (Nov/2007) Google TechTalk video (<http://www.youtube.com/watch?v=947gWqwkhu0>) on YouTube
- "Using Bloom Filters" ([http://www.perl.com/pub/2004/04/08/bloom\\_filters.html](http://www.perl.com/pub/2004/04/08/bloom_filters.html)) Detailed Bloom Filter explanation using Perl

## Implementations

- Implementation in C ([http://en.literateprograms.org/Bloom\\_filter\\_\(C\)](http://en.literateprograms.org/Bloom_filter_(C))) from literateprograms.org
- Implementation in C++ and Object Pascal (<http://www.partow.net/programming/hashfunctions/index.html>) from partow.net
- Implementation in C# (<http://codeplex.com/bloomfilter>) from codeplex.com
- Implementation in Erlang (<http://sites.google.com/site/scalablebloomfilters/>) from sites.google.com
- Implementation in Haskell (<http://hackage.haskell.org/cgi-bin/hackage-scripts/package/bloomfilter>) from hackage.org
- Implementation in Java (<https://github.com/MagnusS/Java-BloomFilter>) from MagnusS on github.com
- Implementation in Javascript (<http://la.ma.la/misc/js/bloomfilter/>) from la.ma.la
- Implementation in Lisp (<http://lemonodor.com/archives/000881.html>) from lemonodor.com
- Implementation in Perl (<http://search.cpan.org/dist/Bloom-Filter/>) from cpan.org
- Implementation in PHP (<http://code.google.com/p/php-bloom-filter/>) from code.google.com
- Implementation in Python, Scalable Bloom Filter (<http://pypi.python.org/pypi/pybloom/1.0.2>) from pypi.python.org
- Implementation in Ruby (<http://www.rubyinside.com/bloom-filters-a-powerful-tool-599.html>) from rubyinside.com
- Implementation in Scala (<http://www.codecommit.com/blog/scala/bloom-filters-in-scala>) from codecommit.com
- Implementation in Tcl (<http://www.kocjan.org/tclmentor/61-bloom-filters-in-tcl.html>) from kocjan.org

# Locality preserving hashing

---

In computer science, a **locality preserving hashing** is a hash function  $f$  that maps a point or points in a multidimensional coordinate space to a scalar value, such that if we have three points  $A$ ,  $B$  and  $C$  such that

$$|A - B| < |B - C| \Rightarrow |f(A) - f(B)| < |f(B) - f(C)|.$$

In other words, these are hash functions where the relative distance between the input values is preserved in the relative distance between of the output hash values; input values that are closer to each other will produce output hash values that are closer to each other.

This is in contrast to cryptographic hash functions and checksums, which are designed to have maximum output difference between adjacent inputs.

Locality preserving hashes are related to space-filling curves and locality sensitive hashing.

## External links

- Indyk, Piotr; Motwani, Rajeev; Raghavan, Prabhakar; Vempala, Santosh (1997). "Locality-preserving hashing in multidimensional spaces". *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing*. STOC '97. pp. 618–625. doi:10.1145/258533.258656. ISBN 0-89791-888-6. CiteSeerX: 10.1.1.50.4927 <sup>[1]</sup>.
- Chin, Andrew (1994). "Locality-preserving hash functions for general purpose parallel computation" <sup>[2]</sup>. *Algorithmica* **12** (2–3): 170–181. doi:10.1007/BF01185209.

## References

[1] <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.50.4927>

[2] <http://www.unclaw.com/chin/scholarship/hashfunctions.pdf>

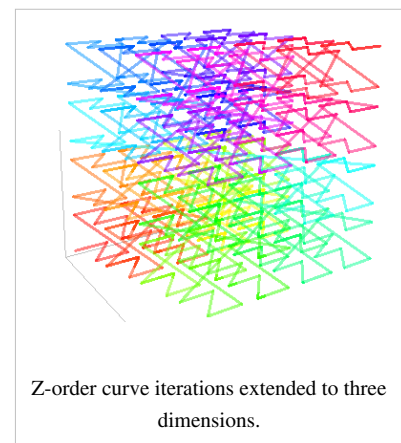
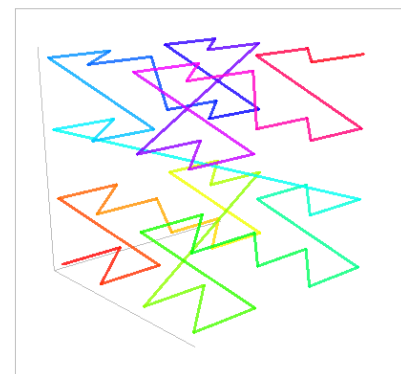
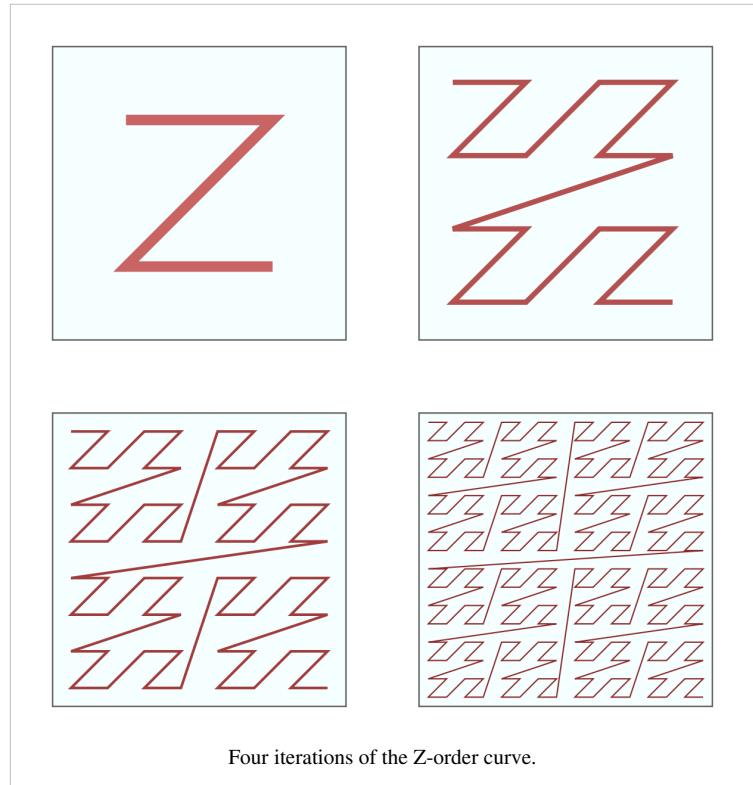
---

# Morton number

In mathematical analysis and computer science, **Z-order**, **Morton order**, or **Morton code** is a function which maps multidimensional data to one dimension while preserving locality of the data points. It was introduced in 1966 by G. M. Morton.<sup>[1]</sup> The z-value of a point in multidimensions is simply calculated by interleaving the binary representations of its coordinate values. Once the data are sorted into this ordering, any one-dimensional data structure can be used such as binary search trees, B-trees, skip lists or (with low significant bits truncated) hash tables. The resulting ordering can equivalently be described as the order one would get from a depth-first traversal of a quadtree; because of its close connection with quadtrees, the Z-ordering can be used to efficiently construct quadtrees and related higher dimensional data structures.<sup>[2]</sup>

## Coordinate values

The figure below shows the Z-values for the two dimensional case with integer coordinates  $0 \leq x \leq 7$ ,  $0 \leq y \leq 7$  (shown both in decimal and binary). Interleaving the binary coordinate values yields binary z-values as shown. Connecting the z-values in their numerical order produces the recursively Z-shaped curve.



	x:	0	1	2	3	4	5	6	7
		000	001	010	011	100	101	110	111
y: 0		000000	000001	000100	000101	010000	010001	010100	010101
1		000010	000011	000110	000111	010010	010011	010110	010111
2		001000	001001	001100	001101	011000	011001	011100	011101
3		001010	001011	001110	001111	011010	011011	011110	011111
4		100000	100001	100100	100101	110000	110001	110100	110101
5		100010	100011	100110	100111	110010	110011	110110	110111
6		101000	101001	101100	101101	111000	111001	111100	111101
7		101010	101011	101110	101111	111010	111011	111110	111111

## Efficiently building quadtrees

As mentioned, the Z-ordering can be used to efficiently build a quadtree for a set of points. The basic idea is to sort the input set according to Z-order. Once sorted, the points can either be stored in a binary search tree and used directly, which is called a linear quadtree,<sup>[3]</sup> or they can be used to build a pointer based quadtree.

The input points are usually scaled in each dimension to be positive integers, either as a fixed point representation over the unit range [0, 1] or corresponding to the machine word size. Both representations are equivalent and allow for the highest order non-zero bit to be found in constant time. Each square in the quadtree has a side length which is a power of two, and corner coordinates which are multiples of the side length. Given any two points, the *derived square* for the two points is the smallest square covering both points. The interleaving of bits from the x and y components of each point is called the *shuffle* of x and y, and can be extended to higher dimensions.<sup>[2]</sup>

Points can be sorted according to their shuffle without explicitly interleaving the bits. To do this, for each dimension, the most significant bit of the exclusive or of the coordinates of the two points for that dimension is examined. The dimension for which the most significant bit is largest is then used to compare the two points to determine their shuffle order.

The exclusive or operation masks off the higher order bits for which the two coordinates are identical. Since the shuffle interleaves bits from higher order to lower order, identifying the coordinate with the largest most significant bit, identifies the first bit in the shuffle order which differs, and that coordinate can be used to compare the two points.<sup>[4]</sup> This is shown in the following Python code:

```
def cmp_zorder(a, b):
    j = 0
    k = 0
```



```

x = 0
for k in range(dim):
    y = a[k] ^ b[k]
    if less_msb(x, y):
        j = k
        x = y
return a[j] - b[j]

```

One way to determine whether the most significant smaller is to compare the floor of the base-2 logarithm of each point. It turns out the following operation is equivalent, and only requires exclusive or operations<sup>[4]</sup>:

```

def less_msb(x, y):
    return x < y and x < (x ^ y)

```

It is also possible to compare floating point numbers using the same technique. The *less\_msb* function is modified to first compare the exponents. Only when they are equal is the standard *less\_msb* function used on the mantissas.<sup>[5]</sup>

Once the points are in sorted order, two properties make it easy to build a quadtree: The first is that the points contained in a square of the quadtree form a contiguous interval in the sorted order. The second is that if more than one child of a square contains an input point, the square is the *derived square* for two adjacent points in the sorted order.

For each adjacent pair of points, the derived square is computed and its side length determined. For each derived square, the interval containing it is bounded by the first larger square to the right and to the left in sorted order.<sup>[2]</sup> Each such interval corresponds to a square in the quadtree. The result of this is a compressed quadtree, where only nodes containing input points or two or more children are present. A non-compressed quadtree can be built by restoring the missing nodes, if desired.

Rather than building a pointer based quadtree, the points can be maintained in sorted order in a data structure such as a binary search tree. This allows points to be added and deleted in  $O(\log n)$  time. Two quadtrees can be merged by merging the two sorted sets of points, and removing duplicates. Point location can be done by searching for the points preceding and following the query point in the sorted order. If the quadtree is compressed, the predecessor node found may be an arbitrary leaf inside the compressed node of interest. In this case, it is necessary to find the predecessor of the least common ancestor of the query point and the leaf found.<sup>[6]</sup>

## Use with one-dimensional data structures for range searching

Although preserving locality well, for efficient range searches an algorithm is necessary for calculating, from a point encountered in the data structure, the next Z-value which is in the multidimensional search range:

x=	0	1	2	3	4	5	6	7
y= 0	0	1	4	5	16	17	<b>20</b>	<del>21</del>
y= 1	2	3	6	7	18	<b>19</b>	22	23
y= 2	8	9	12	13	24	25	28	29
y= 3	10	11	14	15	26	27	30	31
y= 4	<del>32</del>	<del>33</del>	<b>36</b>	<b>37</b>	48	49	52	53
y= 5	<del>34</del>	<del>35</del>	38	39	50	51	54	55
y= 6	<del>40</del>	<del>41</del>	<del>44</del>	<del>45</del>	56	57	60	61
y= 7	<del>42</del>	<del>43</del>	46	47	58	59	62	63

In this example, the range being queried ( $x = 2, \dots, 3, y = 2, \dots, 6$ ) is indicated by the dotted rectangle. Its highest Z-value (MAX) is 45. In this example, the value  $F = 19$  is encountered when searching a data structure in increasing Z-value direction, so we would have to search in the interval between  $F$  and MAX (hatched area). To speed up the search, one would calculate the next Z-value which is in the search range, called BIGMIN (36 in the example) and only search in the interval between BIGMIN and MAX (bold values), thus skipping most of the hatched area. Searching in decreasing direction is analogous with LITMAX which is the highest Z-value in the query range lower than  $F$ . The BIGMIN problem has first been stated and its solution shown in Tropf and Herzog.<sup>[7]</sup> This solution is also used in UB-trees ("GetNextZ-address"). As the approach does not depend on the one dimensional data structure chosen, there is still free choice of structuring the data, so well known methods such as balanced trees can be used to cope with dynamic data (in contrast for example to R-trees where special considerations are necessary). Similarly, this independence makes it easier to incorporate the method into existing databases.

Applying the method hierarchically (according to the data structure at hand), optionally in both increasing and decreasing direction, yields highly efficient multidimensional range search which is important in both commercial and technical applications, e.g. as a procedure underlying nearest neighbour searches. Z-order is one of the few multidimensional access methods that has found its way into commercial database systems (Oracle database 1995,<sup>[8]</sup> Transbase 2000<sup>[9]</sup>).

As long ago as 1966, G.M.Morton proposed Z-order for file sequencing of a static two dimensional geographical database. Areal data units are contained in one or a few quadratic frames represented by their sizes and lower right corner Z-values, the sizes complying with the Z-order hierarchy at the corner position. With high probability, changing to an adjacent frame is done with one or a few relatively small scanning steps.

## Related structures

As an alternative, the Hilbert curve has been suggested as it has a better order-preserving behaviour, but here the calculations are much more complicated, leading to significant processor overhead. BIGMIN source code for both Z-curve and Hilbert-curve were described in a patent by H. Tropic.<sup>[10]</sup>

For a recent overview on multidimensional data processing, including e.g. nearest neighbour searches, see Hanan Samet's textbook.<sup>[11]</sup>

## Applications in linear algebra

The Strassen algorithm for matrix multiplication is based on splitting the matrices in four blocks, and then recursively each of these blocks in four smaller blocks, until the blocks are single elements (or more practically: until reaching matrices so small that the trivial algorithm is faster). Arranging the matrix elements in Z-order then improves locality, and has the additional advantage (compared to row- or column-major ordering) that the subroutine for multiplying two blocks does not need to know the total size of the matrix, but only the size of the blocks and their location in memory. Effective use of Strassen multiplication with Z-order has been demonstrated, see Valsalam and Skjellum's 2002 paper<sup>[12]</sup>.

## References

- [1] Morton, G. M. (1966), *A computer Oriented Geodetic Data Base; and a New Technique in File Sequencing*, Technical Report, Ottawa, Canada: IBM Ltd.
- [2] Bern, M.; Eppstein, D.; Teng, S.-H. (1999), "Parallel construction of quadrees and quality triangulations", *Int. J. Comp. Geom. & Appl.* **9** (6): 517–532, doi:10.1142/S0218195999000303.
- [3] Gargantini, I. (1982), "An effective way to represent quadrees", *Communications of the ACM* **25** (12): 905–910.
- [4] Chan, T. (2002), "Closest-point problems simplified on the RAM" ([http://www.cs.uwaterloo.ca/~tmchan/ram\\_soda.ps.gz](http://www.cs.uwaterloo.ca/~tmchan/ram_soda.ps.gz)), *ACM-SIAM Symposium on Discrete Algorithms*, .
- [5] Connor, M.; Kumar, P (2009), "Fast construction of k-nearest neighbour graphs for point clouds" ([http://compgeom.com/~piyush/papers/tvcg\\_stann.pdf](http://compgeom.com/~piyush/papers/tvcg_stann.pdf)), *IEEE Transactions on Visualization and Computer Graphics*,
- [6] Har-Peled, S. (2010), *Data structures for geometric approximation* ([http://www.madalgo.au.dk/img/SS2010/Course Material/Data-Structures for Geometric Approximation by Sariel Har-Peled.pdf](http://www.madalgo.au.dk/img/SS2010/Course%20Material/Data-Structures%20for%20Geometric%20Approximation%20by%20Sariel%20Har-Peled.pdf)),
- [7] Tropic, H.; Herzog, H. (1981), "Multidimensional Range Search in Dynamically Balanced Trees" (<http://www.vision-tools.com/h-tropic/multidimensionalrangequery.pdf>), *Angewandte Informatik* **2**: 71–77, .
- [8] Gaede, Volker; Guenther, Oliver (1998), "Multidimensional access methods" (<http://www-static.cc.gatech.edu/computing/Database/readinggroup/articles/p170-gaede.pdf>), *ACM Computing Surveys* **30** (2): 170–231, doi:10.1145/280277.280279, .
- [9] Ramsak, Frank; Markl, Volker; Fenk, Robert; Zirkel, Martin; Elhardt, Klaus; Bayer, Rudolf (2000), "Integrating the UB-tree into a Database System Kernel" (<http://www.mistral.in.tum.de/results/publications/RMF+00.pdf>), *Int. Conf. on Very Large Databases (VLDB)*, pp. 263–272, .
- [10] US 7321890 (<http://worldwide.espacenet.com/textdoc?DB=EPODOC&IDX=US7321890>), Tropic, H., "Database system and method for organizing data elements according to a Hilbert curve", issued January 22, 2008.
- [11] Samet, H. (2006), *Foundations of Multidimensional and Metric Data Structures*, San Francisco: Morgan-Kaufmann.
- [12] Vinod Valsalam, Anthony Skjellum: A framework for high-performance matrix multiplication based on hierarchical abstractions, algorithms and optimized low-level kernels. *Concurrency and Computation: Practice and Experience* 14(10): 805-839 (2002)

## External links

- STANN: A library for approximate nearest neighbor search, using Z-order curve (<http://www.compgeom.com/~stann>)
- Methods for programming bit interleaving (<http://graphics.stanford.edu/~seander/bithacks.html#InterleaveTableObvious>), Sean Eron Anderson, Stanford University

# Zobrist hashing

**Zobrist hashing** (also referred to as **Zobrist keys** or **Zobrist signatures**<sup>[1]</sup>) is a hash function construction used in computer programs that play abstract board games, such as chess and Go, to implement transposition tables, a special kind of hash table that is indexed by a board position and used to avoid analyzing the same position more than once. Zobrist hashing is named for its inventor, Albert Lindsey Zobrist.<sup>[2]</sup>

Zobrist hashing starts by randomly generating bitstrings for each possible element of a board game. Given a certain board position, it breaks up the board into independent components, finds out what state each component is in, and combines the bitstrings representing those elements together using bitwise XOR. If the bitstrings are long enough, different board positions will almost certainly hash to different values; however longer bitstrings require proportionally more computer resources to manipulate. Many game engines store only the hash values in the transposition table, omitting the position information itself entirely to reduce memory usage, and assuming that hash collisions will not occur, or will not greatly influence the results of the table if they do.

Zobrist hashing is the first known instance of tabulation hashing. The result is a 3-wise independent hash family. In particular, it is strongly universal.

As an example, in chess, each of the 64 squares can at any time be empty, or contain one of the 6 game pieces, which are either black or white. That is, each square can be in one of  $1 + 6 \times 2 = 13$  possible states at any time. Thus one needs to generate at most  $13 \times 64 = 832$  random bitstrings. Given a position, one obtains its Zobrist hash by finding out which pieces are on which squares, and combining the relevant bitstrings together.

The position of a board can be updated simply by XORing out the bitstring(s) for states which have changed, and XORing in the bitstrings for the new states. For instance, if a pawn on a chessboard square is replaced by a rook from another square, the resulting position would be produced by XORing the existing hash with the bitstrings for:

```
'pawn at this square'      (XORing out the pawn at this square)
'rook at this square'      (XORing in the rook at this square)
'rook at source square'    (XORing out the rook at the source square)
'nothing at source square' (XORing in nothing at the source square).
```

This makes Zobrist hashing very efficient for traversing a game tree.

In computer go, this technique is also used for superko detection.

## References

- [1] Zobrist keys: a means of enabling position comparison. (<http://web.archive.org/web/20070822204038/http://www.seanet.com/~brucemo/topics/zobrist.htm>)
- [2] Albert Lindsey Zobrist, *A New Hashing Method with Application for Game Playing* (<https://www.cs.wisc.edu/techreports/1970/TR88.pdf>), Tech. Rep. 88, Computer Sciences Department, University of Wisconsin, Madison, Wisconsin, (1969).

# Rolling hash

A **rolling hash** is a hash function where the input is hashed in a window that moves through the input.

A few hash functions allow a rolling hash to be computed very quickly -- the new hash value is rapidly calculated given only the old hash value, the old value removed from the window, and the new value added to the window -- similar to the way a moving average function can be computed much more quickly than other low-pass filters.

One of the main applications is the Rabin-Karp string search algorithm, which uses the rolling hash described below.

Another popular application is rsync program which uses a checksum based on Mark Adler's adler-32 as its rolling hash.

At best, rolling hash values are pairwise independent<sup>[1]</sup> or strongly universal. They cannot be 3-wise independent, for example.

## Rabin-Karp rolling hash

The Rabin-Karp string search algorithm is normally used with a very simple rolling hash function that only uses multiplications and additions:

$H = c_1a^{k-1} + c_2a^{k-2} + c_3a^{k-3} + \dots + c_ka^0$  where  $a$  is a constant and  $c_1, \dots, c_k$  are the input characters.

In order to avoid manipulating huge  $H$  values, all math is done modulo  $n$ . The choice of  $a$  and  $n$  is critical to get good hashing; see linear congruential generator for more discussion.

Removing and adding chars simply involves adding or subtracting the first or last term. Shifting all chars by one position to the left requires multiplying the entire sum  $H$  by  $a$ . Shifting all chars by one position to the right requires dividing the entire sum  $H$  by  $a$ . Note that in modulo arithmetic,  $a$  can be chosen to have a multiplicative inverse  $a^{-1}$  by which  $H$  can be multiplied to get the result of the division without actually performing a division.

## Cyclic polynomial

Hashing by cyclic polynomial<sup>[2]</sup>—sometimes called Buzhash<sup>[3][4]</sup>—is also simple, but it has the benefit of avoiding multiplications, using barrel shifts instead. It is a form of tabulation hashing: it presumes that there is some hash function  $h$  from characters to integers in the interval  $[0, 2^L)$ . This hash function might be simply an array or a hash table mapping characters to random integers. Let the function  $s$  be a cyclic binary rotation (or barrel shift): it rotates the bits by 1 to the left, pushing the latest bit in the first position. E.g.,  $s(10011) = 00111$ . Let  $\oplus$  be the bit-wise exclusive or. The hash values are defined as

$$H = s^{k-1}(h(c_1)) \oplus s^{k-2}(h(c_2)) \oplus \dots \oplus s(h(c_{k-1})) \oplus h(c_k)$$

where the multiplications by powers of two can be implemented by binary shifts. The result is a number in  $[0, 2^L)$ .

Computing the hash values in a rolling fashion is done as follows. Let  $H$  be the previous hash value. Rotate  $H$  once:  $H \leftarrow s(H)$ . If  $c_1$  is the character to be removed, rotate it  $k$  times:  $s^k(h(c_1))$ . Then simply set

$$H \leftarrow s(H) \oplus s^k(h(c_1)) \oplus h(c_{k+1})$$

where  $c_{k+1}$  is the new character.

Hashing by cyclic polynomials is strongly universal or pairwise independent: simply keep the first  $L - k + 1$  bits. That is, take the result  $H$  and dismiss any  $k - 1$  consecutive bits.<sup>[1]</sup> In practice, this can be achieved by an integer division  $H \rightarrow H \div 2^{k-1}$ .

## Computational complexity

All rolling hash functions are linear in the number of characters, but their complexity with respect to the length of the window ( $k$ ) varies. Rabin-Karp rolling hash requires the multiplications of two  $k$ -bit numbers, integer multiplication is in  $O(k \log k 2^{O(\log^* k)})$ <sup>[5]</sup>. Hashing ngrams by cyclic polynomials can be done in linear time.<sup>[1]</sup>

## Software

- ngramhashing<sup>[6]</sup> is a Free software C++ implementation of several rolling hash functions
- rollinghashjava<sup>[7]</sup> is an Apache licensed Java implementation of rolling hash functions

## Footnotes

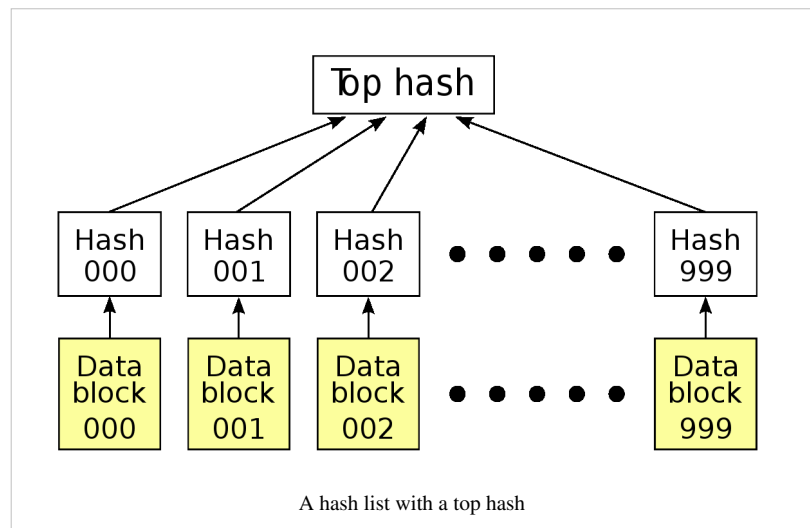
- [1] Daniel Lemire, Owen Kaser: Recursive n-gram hashing is pairwise independent, at best, Computer Speech & Language 24 (4), pages 698-710, 2010. arXiv:0705.4676
- [2] Jonathan D. Cohen, Recursive Hashing Functions for n-Grams (<http://www.cparity.com/projects/AcmClassification/samples/256168.pdf>), ACM Trans. Inf. Syst. 15 (3), 1997
- [3] Buzhash Java implementation (<https://www.se.auckland.ac.nz/courses/SOFTENG250/archive/2006/assignments/Hashing/BuzHash.java>)
- [4] Buzhash C implementation (<https://yara-project.googlecode.com/svn/trunk/libyara/hash.c>)
- [5] M. Fürer, Faster integer multiplication, in: STOC '07, 2007, pp. 57–66.
- [6] <http://code.google.com/p/ngramhashing/>
- [7] <http://code.google.com/p/rollinghashjava/>

# Hash list

In computer science, a **hash list** is typically a list of hashes of the data blocks in a file or set of files. Lists of hashes are used for many different purposes, such as fast table lookup (hash tables) and distributed databases (distributed hash tables). This article covers hash lists that are used to guarantee data integrity.

A hash list is an extension of the old concept of hashing an item (for instance, a file). A hash list is usually sufficient for most needs, but a more advanced form of the concept is a hash tree.

Hash lists can be used to protect any kind of data stored, handled and transferred in and between computers. An important use of hash lists is to make sure that data blocks received from other peers in a peer-to-peer network are received undamaged and unaltered, and to check that the other peers do not "lie" and send fake blocks.



Usually a cryptographic hash function such as SHA-1 is used for the hashing. If the hash list only needs to protect against unintentional damage less secure checksums such as CRCs can be used.

Hash lists are better than a simple hash of the entire file since, in the case of a data block being damaged, this is noticed, and only the damaged block needs to be redownloaded. With only a hash of the file, many undamaged blocks would have to be redownloaded, and the file reconstructed and tested until the correct hash of the entire file is

obtained. Hash lists also protect against nodes that try to sabotage by sending fake blocks, since in such a case the damaged block can be acquired from some other source.

## Root hash

Often, an additional hash of the hash list itself (a *top hash*, also called *root hash* or *master hash*) is used. Before downloading a file on a p2p network, in most cases the top hash is acquired from a trusted source, for instance a friend or a web site that is known to have good recommendations of files to download. When the top hash is available, the hash list can be received from any non-trusted source, like any peer in the p2p network. Then the received hash list is checked against the trusted top hash, and if the hash list is damaged or fake, another hash list from another source will be tried until the program finds one that matches the top hash.

In some systems (like for example BitTorrent), instead of a top hash the whole hash list is available on a web site in a small file. Such a "torrent file" contains a description, file names, a hash list and some additional data.

## Hash tree

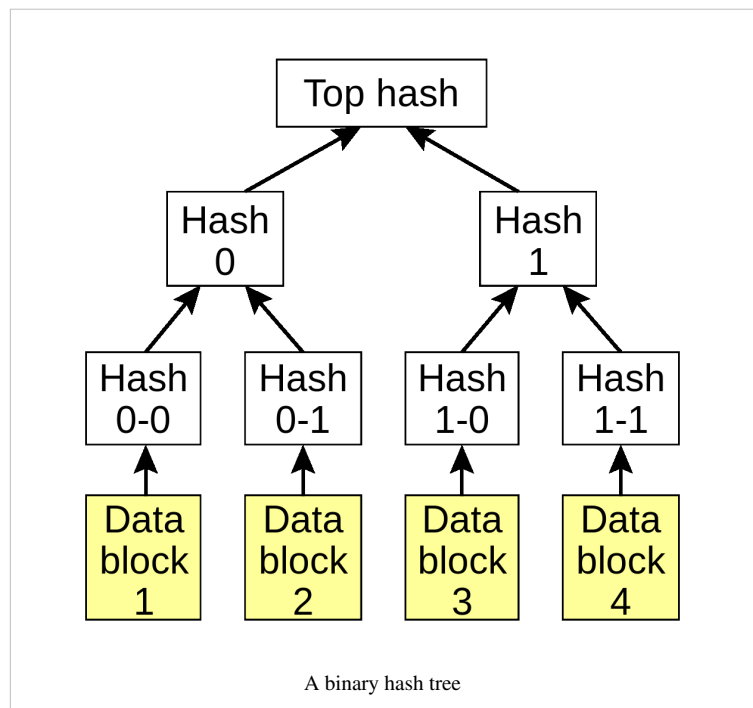
In cryptography and computer science **hash trees** or **Merkle trees** are a type of data structure that contains a tree of summary information about a larger piece of data – for instance a file – used to verify its contents. The concept is named after Ralph Merkle. Hash trees are a combination of hash lists and hash chaining, which in turn are extensions of hashing. Hash trees in which the underlying hash function is Tiger are often called **Tiger trees** or **Tiger tree hashes**.

### Uses

Hash trees can be used to verify any kind of data stored, handled and transferred in and between computers. Currently the main use of hash trees is to make sure that data blocks

received from other peers in a peer-to-peer network are received undamaged and unaltered, and even to check that the other peers do not lie and send fake blocks. Suggestions have been made to use hash trees in trusted computing systems. Sun Microsystems has used Hash Trees in the ZFS filesystem.<sup>[1]</sup> Hash Trees are used in Google Wave protocol,<sup>[2]</sup> Git distributed revision control system, the Tahoe-LAFS backup system, the Bitcoin peer-to-peer network, and a number of NoSQL systems like Apache Cassandra & Riak<sup>[3]</sup>

Hash trees were invented in 1979 by Ralph Merkle.<sup>[4]</sup> The original purpose was to make it possible to efficiently handle many Lamport one-time signatures. Lamport signatures are believed to still be secure in the event that quantum computers become reality. Unfortunately each Lamport key can only be used to sign a single message. But combined with hash trees they can be used for many messages and then become a fairly efficient digital signature scheme.



## How hash trees work

A hash tree is a tree of hashes in which the leaves are hashes of data blocks in, for instance, a file or set of files. Nodes further up in the tree are the hashes of their respective children. For example, in the picture *hash 0* is the result of hashing *hash 0-0* and then *hash 0-1*. That is,  $hash\ 0 = hash(hash\ 0-0 \parallel hash\ 0-1)$  where  $\parallel$  denotes concatenation.

Most hash tree implementations are binary (two child nodes under each node) but they can just as well use many more child nodes under each node.

Usually, a cryptographic hash function such as SHA-1, Whirlpool, or Tiger is used for the hashing. If the hash tree only needs to protect against unintentional damage, much less secure checksums such as CRCs can be used.

In the top of a hash tree there is a *top hash* (or *root hash* or *master hash*). Before downloading a file on a p2p network, in most cases the top hash is acquired from a trusted source, for instance a friend or a web site that is known to have good recommendations of files to download. When the top hash is available, the hash tree can be received from any non-trusted source, like any peer in the p2p network. Then, the received hash tree is checked against the trusted top hash, and if the hash tree is damaged or fake, another hash tree from another source will be tried until the program finds one that matches the top hash.

The main difference from a hash list is that one branch of the hash tree can be downloaded at a time and the integrity of each branch can be checked immediately, even though the whole tree is not available yet. For example, in the picture the integrity of *data block 2* can be verified immediately if the tree already contains *hash 0-0* and *hash 1* by hashing the data block and iteratively combining the result with *hash 0-0* and then *hash 1* and finally comparing the result with the *top hash*. Similarly, the integrity of *data block 3* can be verified if the tree already has *hash 1-1* and *hash 0*. This can be an advantage since it is efficient to split files up in very small data blocks so that only small blocks have to be redownloaded if they get damaged. If the hashed file is very big, such a hash tree or hash list becomes fairly big. But if it is a tree, one small branch can be downloaded quickly, the integrity of the branch can be checked, and then the downloading of data blocks can start.

There are several additional tricks, benefits and details regarding hash trees. See the references and external links below for more in-depth information.

## Tiger tree hash

The Tiger tree hash is a widely used form of hash tree. It uses a binary hash tree (two child nodes under each node), usually has a data block size of 1024-bytes and uses the cryptographically secure Tiger hash.

Tiger tree hashes are used in the Gnutella, Gnutella2, and Direct Connect P2P file sharing protocols and in file sharing applications such as Phex, BearShare, LimeWire, Shareaza, DC++<sup>[5]</sup> and Valknut.

## References

- Merkle tree patent 4,309,569 <sup>[6]</sup> – Explains both the hash tree structure and the use of it to handle many one-time signatures.
- Tree Hash EXchange format (THEX) <sup>[7]</sup> – A detailed description of Tiger trees.
- Efficient Use of Merkle Trees <sup>[8]</sup> – RSA labs explanation of the original purpose of Merkle trees: To handle many Lamport one-time signatures.

[1] Jeff Bonwick's Blog *ZFS End-to-End Data Integrity* ([https://blogs.oracle.com/bonwick/entry/zfs\\_end\\_to\\_end\\_data](https://blogs.oracle.com/bonwick/entry/zfs_end_to_end_data))

[2] Google Wave Federation Protocol *Wave Protocol Verification Paper* (<http://www.waveprotocol.org/protocol/whitepapers/wave-protocol-verification>)

[3] "When a replica is down for an extended period of time, or the machine storing hinted handoffs for an unavailable replica goes down as well, replicas must synchronize from one-another. In this case, Cassandra and Riak implement a Dynamo-inspired process called anti-entropy. In anti-entropy, replicas exchange Merkle trees to identify parts of their replicated key ranges which are out of sync. A Merkle tree is a hierarchical hash verification: if the hash over the entire keyspace is not the same between two replicas, they will exchange hashes of smaller and smaller portions of the replicated keyspace until the out-of-sync keys are identified. This approach reduces unnecessary data transfer



between replicas which contain mostly similar data." <http://www.aosabook.org/en/nosql.html>

- [4] R. C. Merkle, *A digital signature based on a conventional encryption function*, Crypto '87
- [5] "DC++'s feature list" (<http://dcplusplus.sourceforge.net/features.html>)
- [6] <http://www.google.com/patents?vid=4309569>
- [7] <http://web.archive.org/web/20080316033726/http://www.open-content.net/specs/draft-jchapweske-thex-02.html>
- [8] <http://www.rsasecurity.com/rsalabs/node.asp?id=2003>

## External links

- <http://www.codeproject.com/cs/algorithms/thexcs.asp> – Tiger Tree Hash (TTH) source code in C# – by Gil Schmidt
- <http://sourceforge.net/projects/tigertree/> – Tiger Tree Hash (TTH) implementations in C and Java
- RHash (<http://rhash.sourceforge.net/>), an open source command-line tool, which can calculate TTH and magnet links with TTH.
- "Series of mini-lectures about cryptographic hash functions" (<http://www.guardtime.com/educational-series-on-hashes/>); includes application in time-stamping and provable security; by A. Buldas, 2011.

## Prefix hash tree

---

A **prefix hash tree** (PHT) is a distributed data structure that enables more sophisticated queries over a distributed hash table (DHT). The prefix hash tree uses the lookup interface of a DHT to construct a trie-based data structure that is both efficient (updates are doubly logarithmic in the size of the domain being indexed), and resilient (the failure of any given node in a prefix hash tree does not affect the availability of data stored at other nodes).

## External links

- <http://berkeley.intel-research.net/sylvia/pht.pdf> - *Prefix Hash Tree: An Indexing Data Structure over Distributed Hash Tables*
  - <http://pier.cs.berkeley.edu> - *PHT was developed as part of work on the PIER project.*
-

# Hash trie

---

In computer science, **hash trie** can refer to:

- A space-efficient implementation of a sparse trie, in which the descendants of each node may be interleaved in memory. (The name is suggested by a similarity to a closed hash table.)<sup>[1]</sup>
- An ordinary trie used to store hash values, for example, in an implementation of a hash tree.
- A data structure which "combines features of hash tables and LC-tries in order to perform efficient lookups and updates"<sup>[2]</sup>

## References

- [1] Liang, Frank (June 1983), *Word hy-phen-a-tion by com-pu-ter* (<http://www.tug.org/docs/liang/liang-thesis.pdf>), Frank M. Liang, Ph.D. thesis, Stanford University., , retrieved 2010-03-28
- [2] Thomas, Roshan; Mark, Brian; Johnson, Tommy; Croall, James (2004), *High-speed Legitimacy-based DDoS Packet Filtering with Network Processors: A Case Study and Implementation on the Intel IXP1200* (<http://napl.gmu.edu/pubs/BookContrib/ThomasMarkJC-NPW04.pdf>), , retrieved 2009-05-03

# Hash array mapped trie

---

A **hash array mapped trie**<sup>[1]</sup> (HAMT) is an implementation of an associative array that combines the characteristics of a hash table and an array mapped trie.<sup>[2]</sup>

## Operation

A HAMT is an array mapped trie where the keys are first hashed in order to ensure an even distribution of keys and to ensure a constant key length.

In a typical implementation of an array mapped trie, each node may branch to up to 32 other nodes. However, as allocating space for 32 pointers for each node would be expensive, each node instead contains a bitmap which is 32 bits long where each bit indicates the presence of a path. This is followed by an array of pointers equal in length to the Hamming weight of the bitmap.

## Advantages of HAMTs

The hash array mapped trie achieves almost hash table-like speed, despite being a functional, persistent data structure.

## Problems with HAMTs

Implementation of a HAMT involves the use of the population count function, which counts the number of ones in the binary representation of a number. This operation is available in many instruction set architectures (where it is sometimes called "CTPOP"), but it is only available in some high-level languages. Although population count can be implemented in software in O(1) time using a series of shift and add instructions, doing so may perform the operation an order of magnitude slower.

## Implementations

The programming languages Clojure<sup>[3]</sup> and Scala use a persistent variant of hash array mapped tries for their native hash map type. The Haskell library `unordered-containers`<sup>[4]</sup> uses the same to implement persistent map and set data types.<sup>[5]</sup> The Rubinius<sup>[6]</sup> implementation of Ruby includes a HAMT, mostly written in Ruby but with 3<sup>[7]</sup> primitives. The concurrent lock-free version<sup>[8]</sup> of the hash trie called Ctrie is a mutable thread-safe implementation which ensures progress. The data-structure has been proven to be correct<sup>[9]</sup> - Ctrie operations have been shown to have the atomicity, linearizability and lock-freedom properties.

## References

- [1] Bagwell, P. (2001) Ideal Hash Trees (<http://lampwww.epfl.ch/papers/idealhashtrees.pdf>). Technical Report, 2001.
- [2] Bagwell, P. (2000) Fast And Space Efficient Trie Searches (<http://lampwww.epfl.ch/papers/triesearches.pdf.gz>). Technical Report, 2000.
- [3] Java source file of Clojure's hash map type. (<https://github.com/clojure/clojure/blob/master/src/jvm/clojure/lang/PersistentHashMap.java>)
- [4] <http://hackage.haskell.org/package/unordered-containers>
- [5] Johan Tibell, A. Announcing unordered-containers 0.2 (<http://blog.johantibell.com/2012/03/announcing-unordered-containers-02.html>)
- [6] Ruby source file of Rubinius's HAMT ([https://github.com/rubinius/rubinius/blob/master/kernel/common/hash\\_hamt.rb](https://github.com/rubinius/rubinius/blob/master/kernel/common/hash_hamt.rb))
- [7] (<https://github.com/rubinius/rubinius/blob/master/vm/builtin/system.cpp#L1626-1655>)
- [8] Prokopec, A. Implementation of Concurrent Hash Tries on GitHub (<https://github.com/axel22/Ctries>)
- [9] Prokopec, A. et al. (2011) Cache-Aware Lock-Free Concurrent Hash Tries (<http://infoscience.epfl.ch/record/166908/files/ctries-techreport.pdf>). Technical Report, 2011.

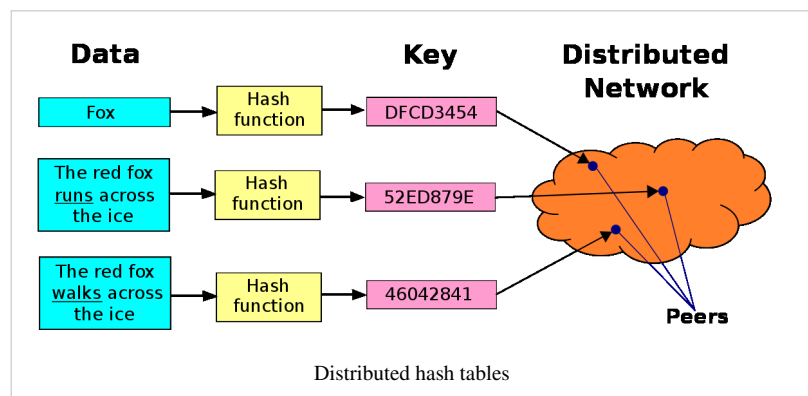
## Distributed hash table

A **distributed hash table (DHT)** is a class of a decentralized distributed system that provides a lookup service similar to a hash table; (*key, value*) pairs are stored in a DHT, and any participating node can efficiently retrieve the value associated with a given key. Responsibility for maintaining the mapping from keys to values is distributed among the nodes, in such a way that a change in the set of participants causes a minimal amount of disruption. This allows a DHT to scale to extremely large numbers of nodes and to handle continual node arrivals, departures, and failures.

DHTs form an infrastructure that can be used to build more complex services, such as anycast, cooperative Web caching, distributed file systems, domain name services, instant messaging, multicast, and also peer-to-peer file sharing and content distribution systems. Notable distributed networks that use DHTs include BitTorrent's distributed tracker, the Coral Content Distribution Network, the Kad network, the Storm botnet, and the YaCy search engine.

## History

DT research was originally motivated, in part, by peer-to-peer systems such as Freenet, gnutella, and Napster, which took advantage of resources distributed across the Internet to provide a single useful application. In particular, they took advantage of increased bandwidth and hard disk capacity to provide a file-sharing service.



These systems differed in how they *found* the data their peers contained:

- Napster, the first large-scale P2P content delivery system to exist, had a central index server: each node, upon joining, would send a list of locally held files to the server, which would perform searches and refer the querier to the nodes that held the results. This central component left the system vulnerable to attacks and lawsuits.
- Gnutella and similar networks moved to a flooding query model—in essence, each search would result in a message being broadcast to every other machine in the network. While avoiding a single point of failure, this method was significantly less efficient than Napster.
- Finally, Freenet is fully distributed, but employs a heuristic key-based routing in which each file is associated with a key, and files with similar keys tend to cluster on a similar set of nodes. Queries are likely to be routed through the network to such a cluster without needing to visit many peers.<sup>[1]</sup> However, Freenet does not guarantee that data will be found.

Distributed hash tables use a more structured key-based routing in order to attain both the decentralization of Freenet and gnutella, and the efficiency and guaranteed results of Napster. One drawback is that, like Freenet, DHTs only directly support exact-match search, rather than keyword search, although Freenet's routing algorithm can be generalized to any key type where a closeness operation can be defined.<sup>[2]</sup>

In 2001, four systems—CAN, Chord,<sup>[3]</sup> Pastry, and Tapestry—ignited DHTs as a popular research topic, and this area of research remains active. Outside academia, DHT technology has been adopted as a component of BitTorrent and in the Coral Content Distribution Network.

## Properties

DHTs characteristically emphasize the following properties:

- **Autonomy and Decentralization:** the nodes collectively form the system without any central coordination.
- **Fault tolerance:** the system should be reliable (in some sense) even with nodes continuously joining, leaving, and failing.
- **Scalability:** the system should function efficiently even with thousands or millions of nodes.

A key technique used to achieve these goals is that any one node needs to coordinate with only a few other nodes in the system — most commonly,  $O(\log n)$  of the  $n$  participants (see below) — so that only a limited amount of work needs to be done for each change in membership.

Some DHT designs seek to be secure against malicious participants<sup>[4]</sup> and to allow participants to remain anonymous, though this is less common than in many other peer-to-peer (especially file sharing) systems; see anonymous P2P.

Finally, DHTs must deal with more traditional distributed systems issues such as load balancing, data integrity, and performance (in particular, ensuring that operations such as routing and data storage or retrieval complete quickly).

## Structure

The structure of a DHT can be decomposed into several main components.<sup>[5][6]</sup> The foundation is an abstract **keyspace**, such as the set of 160-bit strings. A **keyspace partitioning** scheme splits ownership of this keyspace among the participating nodes. An **overlay network** then connects the nodes, allowing them to find the owner of any given key in the keyspace.

Once these components are in place, a typical use of the DHT for storage and retrieval might proceed as follows. Suppose the keyspace is the set of 160-bit strings. To store a file with given *filename* and *data* in the DHT, the SHA-1 hash of *filename* is generated, producing a 160-bit key  $k$ , and a message *put*( $k$ , *data*) is sent to any node participating in the DHT. The message is forwarded from node to node through the overlay network until it reaches the single node responsible for key  $k$  as specified by the keyspace partitioning. That node then stores the key and the data. Any other client can then retrieve the contents of the file by again hashing *filename* to produce  $k$  and asking any DHT node to find the data associated with  $k$  with a message *get*( $k$ ). The message will again

be routed through the overlay to the node responsible for  $k$ , which will reply with the stored *data*.

The keyspace partitioning and overlay network components are described below with the goal of capturing the principal ideas common to most DHTs; many designs differ in the details.

## Keyspace partitioning

Most DHTs use some variant of consistent hashing to map keys to nodes. This technique employs a function  $\delta(k_1, k_2)$  that defines an abstract notion of the *distance* between the keys  $k_1$  and  $k_2$ , which is unrelated to geographical distance or network latency. Each node is assigned a single key called its *identifier* (ID). A node with ID  $i_x$  owns all the keys  $k_m$  for which  $i_x$  is the closest ID, measured according to  $\delta(k_m, i_x)$ .

**Example.** The Chord DHT treats keys as points on a circle, and  $\delta(k_1, k_2)$  is the distance traveling clockwise around the circle from  $k_1$  to  $k_2$ . Thus, the circular keyspace is split into contiguous segments whose endpoints are the node identifiers. If  $i_1$  and  $i_2$  are two adjacent IDs, with a shorter clockwise distance from  $i_1$  to  $i_2$ , then the node with ID  $i_2$  owns all the keys that fall between  $i_1$  and  $i_2$ .

Consistent hashing has the essential property that removal or addition of one node changes only the set of keys owned by the nodes with adjacent IDs, and leaves all other nodes unaffected. Contrast this with a traditional hash table in which addition or removal of one bucket causes nearly the entire keyspace to be remapped. Since any change in ownership typically corresponds to bandwidth-intensive movement of objects stored in the DHT from one node to another, minimizing such reorganization is required to efficiently support high rates of churn (node arrival and failure).

Locality-preserving hashing ensures that similar keys are assigned to similar objects. This can enable a more efficient execution of range queries. Self-Chord<sup>[7]</sup> decouples object keys from peer IDs and sorts keys along the ring with a statistical approach based on the swarm intelligence paradigm. Sorting ensures that similar keys are stored by neighbour nodes and that discovery procedures, including range queries, can be performed in logarithmic time.

## Overlay network

Each node maintains a set of links to other nodes (its *neighbors* or routing table). Together, these links form the overlay network. A node picks its neighbors according to a certain structure, called the network's topology.

All DHT topologies share some variant of the most essential property: for any key  $k$ , each node either has a node ID that owns  $k$  or has a link to a node whose node ID is *closer* to  $k$ , in terms of the keyspace distance defined above. It is then easy to route a message to the owner of any key  $k$  using the following greedy algorithm (that is not necessarily globally optimal): at each step, forward the message to the neighbor whose ID is closest to  $k$ . When there is no such neighbor, then we must have arrived at the closest node, which is the owner of  $k$  as defined above. This style of routing is sometimes called key-based routing.

Beyond basic routing correctness, two important constraints on the topology are to guarantee that the maximum number of hops in any route (route length) is low, so that requests complete quickly; and that the maximum number of neighbors of any node (maximum node degree) is low, so that maintenance overhead is not excessive. Of course, having shorter routes requires higher maximum degree. Some common choices for maximum degree and route length are as follows, where  $n$  is the number of nodes in the DHT, using Big O notation:

Degree	Route length	Note
$O(1)$	$O(n)$	
$O(\log n)$	$O(\log n / \log(\log n))$	
$O(\log n)$	$O(\log n)$	most common, but not optimal (degree/route length)
$O(1)$	$O(\log n)$	
$O(\sqrt{n})$	$O(1)$	

The most common choice,  $O(\log n)$  degree/route length, is not optimal in terms of degree/route length tradeoff, as such topologies typically allow more flexibility in choice of neighbors. Many DHTs use that flexibility to pick neighbors that are close in terms of latency in the physical underlying network.

Maximum route length is closely related to diameter: the maximum number of hops in any shortest path between nodes. Clearly, the network's worst case route length is at least as large as its diameter, so DHTs are limited by the degree/diameter tradeoff<sup>[8]</sup> that is fundamental in graph theory. Route length can be greater than diameter, since the greedy routing algorithm may not find shortest paths.<sup>[9]</sup>

### Algorithms for overlay networks

Aside from routing, there exist many algorithms that exploit the structure of the overlay network for sending a message to all nodes, or a subset of nodes, in a DHT.<sup>[10]</sup> These algorithms are used by applications to do overlay multicast, range queries, or to collect statistics. Two systems that are based on this approach are Structella,<sup>[11]</sup> which implements flooding and random walks on a Pastry overlay, and DQ-DHT,<sup>[12]</sup> which implements a dynamic querying search algorithm over a Chord network.

### DHT implementations

Most notable differences encountered in practical instances of DHT implementations include at least the following:

- The address space is a parameter of DHT. Several real world DHTs use 128-bit or 160-bit key space
- Some real-world DHTs use hash functions other than SHA-1.
- In the real world the key  $k$  could be a hash of a file's *content* rather than a hash of a file's *name* to provide content-addressable storage, so that renaming of the file does not prevent users from finding it.
- Some DHTs may also publish objects of different types. For example, key  $k$  could be the node  $ID$  and associated data could describe how to contact this node. This allows publication-of-presence information and often used in IM applications, etc. In the simplest case,  $ID$  is just a random number that is directly used as key  $k$  (so in a 160-bit DHT  $ID$  will be a 160-bit number, usually randomly chosen). In some DHTs, publishing of nodes IDs is also used to optimize DHT operations.
- Redundancy can be added to improve reliability. The  $(k, data)$  key pair can be stored in more than one node corresponding to the key. Usually, rather than selecting just one node, real world DHT algorithms select  $i$  suitable nodes, with  $i$  being an implementation-specific parameter of the DHT. In some DHT designs, nodes agree to handle a certain key space range, the size of which may be chosen dynamically, rather than hard-coded.
- Some advanced DHTs like Kademlia perform iterative lookups through the DHT first in order to select a set of suitable nodes and send  $put(k, data)$  messages only to those nodes, thus drastically reducing useless traffic, since published messages are only sent to nodes that seem suitable for storing the key  $k$ ; and iterative lookups cover just a small set of nodes rather than the entire DHT, reducing useless forwarding. In such DHTs, forwarding of  $put(k, data)$  messages may only occur as part of a self-healing algorithm: if a target node receives a  $put(k, data)$  message, but believes that  $k$  is out of its handled range and a closer node (in terms of DHT key space) is known, the message is forwarded to that node. Otherwise, data are indexed locally. This leads to a somewhat self-balancing DHT behavior. Of course, such an algorithm requires nodes to publish their presence

data in the DHT so the iterative lookups can be performed.

## Examples

### DHT protocols and implementations

- Apache Cassandra
- BitTorrent DHT (based on Kademlia as provided by Khashmir<sup>[13]</sup>)
- CAN (Content Addressable Network)
- Chord
- Kademlia
- Pastry
- P-Grid
- Tapestry
- TomP2P

### Applications employing DHTs

- BTDig: BitTorrent DHT search engine
- CloudSNAP: a decentralized web application deployment platform
- Codeen: Web caching
- Coral Content Distribution Network
- Dijer: Freenet-like distribution network
- FAROO: Peer-to-peer Web search engine
- Freenet: A censorship-resistant anonymous network
- GUNet: Freenet-like distribution network including a DHT implementation
- JXTA: Opensource P2P platform
- maidsafe: C++ implementation of Kademlia, with NAT traversal and crypto libraries. On its home page listed as "Available as a technology licence and a software solution written in cross platform C++."<sup>[14]</sup>
- Oracle Coherence: An In Memory Data Grid built on a Java DHT implementation
- Retroshare: a Friend-to-friend network<sup>[15]</sup>
- WebSphere eXtreme Scale: proprietary DHT implementation by IBM,<sup>[16]</sup> used for object caching
- YaCy: distributed search engine
- i2p-Bote: Decentralised email service which stores encrypted messages in a DHT.
- Cjdns: Routing engine for mesh-based networks

## References

- [1] *Searching in a Small World Chapters 1 & 2* (<https://freenetproject.org/papers/lic.pdf>), , retrieved 2012-01-10
- [2] "Section 5.2.2" (<https://freenetproject.org/papers/ddisr.pdf>), *A Distributed Decentralized Information Storage and Retrieval System*, , retrieved 2012-01-10
- [3] Hari Balakrishnan, M. Frans Kaashoek, David Karger, Robert Morris, and Ion Stoica. Looking up data in P2P systems (<http://www.cs.berkeley.edu/~istoica/papers/2003/cacm03.pdf>). In *Communications of the ACM*, February 2003.
- [4] Guido Urdaneta, Guillaume Pierre and Maarten van Steen. A Survey of DHT Security Techniques ([http://www.globule.org/publi/SDST\\_acmcs2009.html](http://www.globule.org/publi/SDST_acmcs2009.html)). *ACM Computing Surveys* 43(2), January 2011.
- [5] Moni Naor and Udi Wieder. Novel Architectures for P2P Applications: the Continuous-Discrete Approach (<http://www.wisdom.weizmann.ac.il/~naor/PAPERS/dh.pdf>). *Proc. SPAA*, 2003.
- [6] Gurmeet Singh Manku. Dipsea: A Modular Distributed Hash Table (<http://www-db.stanford.edu/~manku/phd/index.html>). Ph. D. Thesis (Stanford University), August 2004.
- [7] Agostino Forestiero, Emilio Leonardi, Carlo Mastroianni and Michela Meo. Self-Chord: a Bio-Inspired P2P Framework for Self-Organizing Distributed Systems (<http://dx.doi.org/10.1109/TNET.2010.2046745>). *IEEE/ACM Transactions on Networking*, 2010.
- [8] *The (Degree,Diameter) Problem for Graphs* ([http://maite71.upc.es/grup\\_de\\_grafs/table\\_g.html](http://maite71.upc.es/grup_de_grafs/table_g.html)), Maite71.upc.es, , retrieved 2012-01-10

- [9] Gurmeet Singh Manku, Moni Naor, and Udi Wieder. Know thy Neighbor's Neighbor: the Power of Lookahead in Randomized P2P Networks (<http://citeseer.ist.psu.edu/naor04know.html>). Proc. STOC, 2004.
- [10] Ali Ghodsi. Distributed k-ary System: Algorithms for Distributed Hash Tables (<http://www.sics.se/~ali/thesis/>). KTH-Royal Institute of Technology, 2006.
- [11] Miguel Castro, Manuel Costa, and Antony Rowstron. Should we build Gnutella on a structured overlay? (<http://dx.doi.org/10.1145/972374.972397>). Computer Communication Review, 2004.
- [12] Domenico Talia and Paolo Trunfio. Enabling Dynamic Querying over Distributed Hash Tables (<http://dx.doi.org/10.1016/j.jpdc.2010.08.012>). Journal of Parallel and Distributed Computing, 2010.
- [13] Tribler wiki (<http://www.tribler.org/trac/wiki/Khashmir>) retrieved January 2010.
- [14] *maidsafe-dht* (<http://code.google.com/p/maidsafe-dht/>), Code.google.com, , retrieved 2012-01-10
- [15] Retroshare FAQ ([http://retroshare.sourceforge.net/wiki/index.php/Frequently\\_Asked\\_Questions#4-1\\_How\\_does\\_RetroShare\\_know\\_my\\_friend.27s\\_IP\\_address\\_and\\_port.3F\\_Why\\_don.27t\\_I\\_need\\_a\\_static\\_IP\\_address.3F\\_What\\_is\\_DHT\\_for.3F](http://retroshare.sourceforge.net/wiki/index.php/Frequently_Asked_Questions#4-1_How_does_RetroShare_know_my_friend.27s_IP_address_and_port.3F_Why_don.27t_I_need_a_static_IP_address.3F_What_is_DHT_for.3F)) retrieved December 2011
- [16] Billy Newport, IBM Distinguished Engineer (<http://www.devwebsphere.com/devwebsphere/2010/01/implementing-global-indexes-on-websphere-extreme-scale.html>) retrieved October 2010.

## External links

- Distributed Hash Tables, Part 1 (<http://linuxjournal.com/article/6797>) by Brandon Wiley.
- Distributed Hash Tables links (<http://deim.urv.cat/~cpairot/dhts.html>) Carles Pairot's Page on DHT and P2P research
- kademlia.scs.cs.nyu.edu ([http://web.archive.org/web/\\*/http://kademlia.scs.cs.nyu.edu/](http://web.archive.org/web/*/http://kademlia.scs.cs.nyu.edu/)) Archive.org snapshots of kademlia.scs.cs.nyu.edu
- Hazelcast (<http://code.google.com/p/hazelcast>) on Google Code: open source DHT implementation
- scale4j (<http://code.google.com/p/scale4j>) on Google Code: highly scalable domain oriented data-distributed platform for java
- Eng-Keong Lua; Crowcroft, Jon; Pias, Marcelo; Sharma, Ravi; Lim, Steve. *IEEE Survey on overlay network schemes*. CiteSeerX: 10.1.1.111.4197 (<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.111.4197>): covering unstructured and structured decentralized overlay networks including DHTs (Chord, Pastry, Tapestry and others).



# Consistent hashing

---

**Consistent hashing** is a special kind of hashing. By using consistent hashing, only  $K/n$  keys need to be remapped on average, where  $K$  is the number of keys, and  $n$  is the number of slots. In contrast, in most traditional hash tables, a change in the number of array slots causes nearly all keys to be remapped. Consistent hashing could play an increasingly important role as internet use increases and as distributed systems grow more prevalent.

## History

Originally devised by Karger *et al.* at MIT for use in distributed caching, the idea has now been expanded to other areas also. An academic paper from 1997 introduced the term "consistent hashing" as a way of distributing requests among a changing population of Web servers. Each slot is then represented by a node in a distributed system. The addition (joins) and removal (leaves/failures) of nodes only requires  $K/n$  items to be re-shuffled when the number of slots/nodes change.<sup>[1]</sup>

Consistent hashing has also been used to reduce the impact of partial system failures in large Web applications as to allow for robust caches without incurring the system wide fallout of a failure.<sup>[2]</sup>

The consistent hashing concept also applies to the design of distributed hash tables (DHTs). DHTs use consistent hashing to partition a keyspace among a distributed set of nodes, and additionally provide an overlay network that connects nodes such that the node responsible for any key can be efficiently located.

## Need for consistent hashing

While running collections of caching machines some limitations are experienced. A common way of load balancing  $n$  cache machines is to put object  $o$  in cache machine number  $\text{hash}(o) \bmod n$ . But this will not work if a cache machine is added or removed because  $n$  changes and every object is hashed to a new location. This can be disastrous since the originating content servers are flooded with requests from the cache machines. Hence consistent hashing is needed to avoid swamping of servers.

Consistent hashing maps objects to the same cache machine, as far as possible. It means when a cache machine is added, it takes its share of objects from all the other cache machines and when it is removed, its objects are shared between the remaining machines.

The main idea behind the consistent hashing algorithm is to associate each cache with one or more hash value intervals where the interval boundaries are determined by calculating the hash of each cache identifier. (The hash function used to define the intervals does not have to be the same function used to hash the cached values. Only the range of the two functions need match.) If the cache is removed its interval is taken over by a cache with an adjacent interval. All the remaining caches are unchanged.

## Technique

Consistent hashing is based on mapping each object to a point on the edge of a circle (or equivalently, mapping each object to a real angle). The system maps each available machine (or other storage bucket) to many pseudo-randomly distributed points on the edge of the same circle.

To find where an object should be placed, the system finds the location of that object's key on the edge of the circle; then walks around the circle until falling into the first bucket it encounters (or equivalently, the first available bucket with a higher angle). The result is that each bucket contains all the resources located between its point and the next bucket point.

If a bucket becomes unavailable (for example because the computer it resides on is not reachable), then the angles it maps to will be removed. Requests for resources that would have mapped to each of those points now map to the

next highest point. Since each bucket is associated with many pseudo-randomly distributed points, the resources that were held by that bucket will now map to many different buckets. The items that mapped to the lost bucket must be redistributed among the remaining ones, but values mapping to other buckets will still do so and do not need to be moved.

A similar process occurs when a bucket is added. By adding a bucket point, we make any resources between that and the next smaller angle map to the new bucket. These resources will no longer be associated with the previous bucket, and any value previously stored there will not be found by the selection method described above.

The portion of the keys associated with each bucket can be altered by altering the number of angles that bucket maps to.

## Monotonic keys

If it is known that key values will always increase monotonically, an alternative approach to consistent hashing is possible.

## Properties

David Karger et al. list several properties of consistent hashing that make it useful for distributed caching protocols on the Internet: <sup>[1]</sup>

- "spread"
- "load"
- "smoothness"
- "balance"
- "monotonic"

## References

- [1] Karger, D.; Lehman, E.; Leighton, T.; Panigrahy, R.; Levine, M.; Lewin, D. (1997). "Consistent Hashing and Random Trees: Distributed Caching Protocols for Relieving Hot Spots on the World Wide Web" (<http://portal.acm.org/citation.cfm?id=258660>). *Proceedings of the Twenty-ninth Annual ACM Symposium on Theory of Computing*. ACM Press New York, NY, USA. pp. 654–663. doi:10.1145/258533.258660. .
- [2] Karger, D.; Sherman, A.; Berkheimer, A.; Bogstad, B.; Dhanidina, R.; Iwamoto, K.; Kim, B.; Matkins, L.; Yerushalmi, Y. (1999). "Web Caching with Consistent Hashing" (<http://www8.org/w8-papers/2a-webserver/caching/paper2.html>). *Computer Networks* **31** (11): 1203–1213. doi:10.1016/S1389-1286(99)00055-9. .

## External links

- Consistent hashing implementation in C++ (<http://www.martinbroadhurst.com/Consistent-Hash-Ring.html>)
  - Consistent hashing implementation in Erlang ([https://github.com/basho/riak\\_core/blob/master/src/chash.erl](https://github.com/basho/riak_core/blob/master/src/chash.erl))
  - Consistent hashing implementation in Go (<https://github.com/stathat/consistent>)
  - Consistent hashing implementation in C# (<http://code.google.com/p/consistent-hash/>)
  - Consistent hashing implementation in Java ([http://weblogs.java.net/blog/tomwhite/archive/2007/11/consistent\\_hash.html](http://weblogs.java.net/blog/tomwhite/archive/2007/11/consistent_hash.html))
  - Consistent hashing implementation in Python (<http://amix.dk/blog/post/19367>)
  - Understanding Consistent hashing (<http://www.spiteful.com/2008/03/17/programmers-toolbox-part-3-consistent-hashing/>)
  - Consistent hashing implementation in php (<https://github.com/pda/flexihash>)
-

# Stable hashing

**Stable hashing** is a tool used to implement randomized load balancing and distributed lookup in peer-to-peer computer systems.

## Koorde

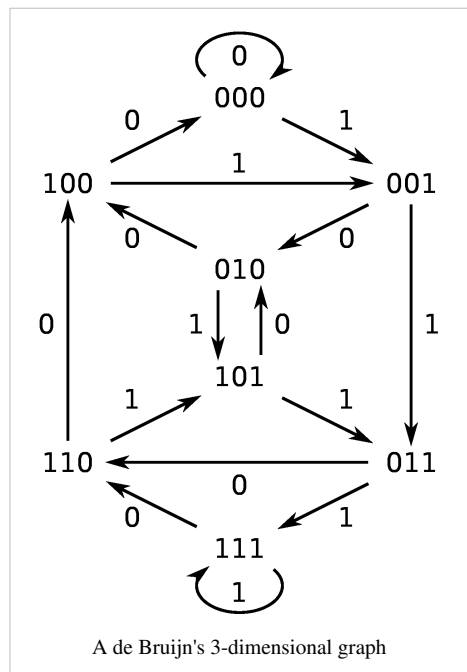
In peer-to-peer networks, **Koorde** is a Distributed hash table (DHT) system based on the Chord DHT and the De Bruijn graph (De Bruijn sequence). Inheriting the simplicity of Chord, Koorde meets  $O(\log n)$  hops per node (where  $n$  is the number of nodes in the DHT), and  $O(\log n / \log \log n)$  hops per lookup request with  $O(\log n)$  neighbors per node.

The Chord concept is based on a wide range of identifiers (e.g.  $2^{160}$ ) in a structure of a ring where an identifier can stand for both node and data. Node-successor is responsible for the whole range of IDs between itself and its predecessor.

### De Bruijn's graphs

Koorde is based on Chord but also on De Bruijn graph (De Bruijn sequence). In a  $d$ -dimensional de Bruijn graph, there are  $2^d$  nodes, each of which has a unique  $d$ -bit ID. The node with ID  $i$  is connected to nodes  $2i$  modulo  $2^d$  and  $2i+1$  modulo  $2^d$ . Thanks to this property, the routing algorithm can route to any destination in  $d$  hops by successively „shifting in” the bits of the destination ID but only if the dimensions of the distance between modulo  $1d$  and  $3d$  are equal.

Routing a message from node  $m$  to node  $k$  is accomplished by taking the number  $m$  and shifting in the bits of  $k$  one at a time until the number has been replaced by  $k$ . Each shift corresponds to a routing hop to the next intermediate address; the hop is valid because each node's neighbors are the two possible outcomes of shifting a 0 or 1 onto its own address. Because of the structure of de Bruijn graphs, when the last bit of  $k$  has been shifted, the query will be at node  $k$ . Node  $k$  responds whether key  $k$  exists.



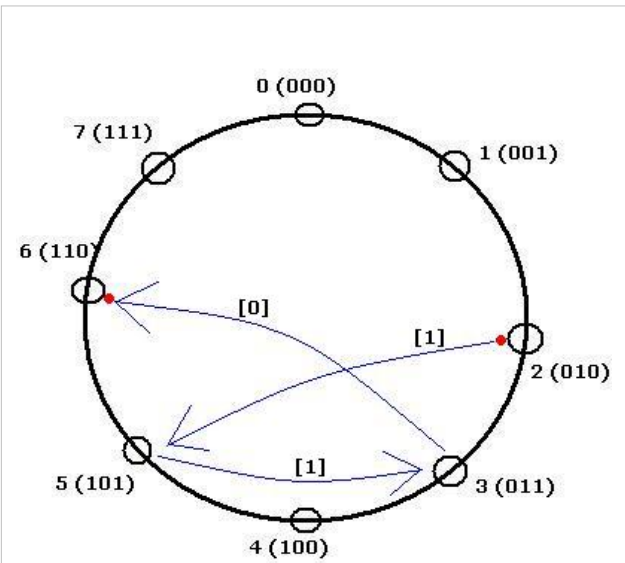
## Routing example

For example, when a message needs to be routed from node “2” (which is “010”) to “6” (which is “110”), the steps are following:

Step 1) Node #2 routes the message to Node #5 (using its connection to  $2i+1 \bmod 8$ ), shifts the bits left and puts “1” as the youngest bit (right side).

Step 2) Node #5 routes the message to Node #3 (using its connection to  $2i+1 \bmod 8$ ), shifts the bits left and puts “1” as the youngest bit (right side).

Step 3) Node #3 routes the message to Node #6 (using its connection to  $2i \bmod 8$ ), shifts the bits left and puts “0” as the youngest bit (right side).



Example of the way Koorde routes from Node2 to Node6 using a 3-dimensional, binary graph.

## Non-constant degree Koorde

The  $d$ -dimensional de Bruijn can be generalized to base  $k$ , in which case node  $i$  is connected to nodes  $k * i + j$  modulo  $k^d$ ,  $0 \leq j < k$ . The diameter is reduced to  $\Theta(\log_k n)$ . Koorde node  $i$  maintains pointers to  $k$  consecutive nodes beginning at the predecessor of  $k * i$  modulo  $k^d$ . Each de Bruijn routing step can be emulated with an expected constant number of messages, so routing uses  $O(\log_k n)$  expected hops. For  $k = \Theta(\log n)$ , we get  $\Theta(\log n)$  degree and  $\Theta(\log n / \log \log n)$  diameter.

```
function n.lookup(k, shift, i)
{
  if k ∈ (n, s] return (s);
  else if i ∈ (n, s] return
    (p.lookup(k, shift << 1, i ⊕ topBit(shift)));
  else return (s.lookup(k, shift, i));
}
```

The Koorde lookup algorithm at node  $n$ .  $k$  is the key.  $i$  is the imaginary De Bruijn node.  $p$  is the reference to the predecessor of  $2n$ .  $s$  is the reference to the successor of  $n$ .

Koorde lookup algorithm.

## References

- "Internet Algorithms" by Greg Plaxton, Fall 2003: [1]
- "Koorde: A simple degree-optimal distributed hash table" by M. Frans Kaashoek and David R. Karger: [2]
- Chord and Koorde descriptions: [3]

## References

- [1] <http://web.archive.org/web/20040929211835/http://www.cs.utexas.edu/users/plaxton/c/395t/slides/DynamicTopologies-2.pdf>
  - [2] <http://iptps03.cs.berkeley.edu/final-papers/koorde.ps>
  - [3] [http://www.cs.jhu.edu/~scheideler/courses/600.348\\_F03/lecture\\_10.pdf](http://www.cs.jhu.edu/~scheideler/courses/600.348_F03/lecture_10.pdf)
-

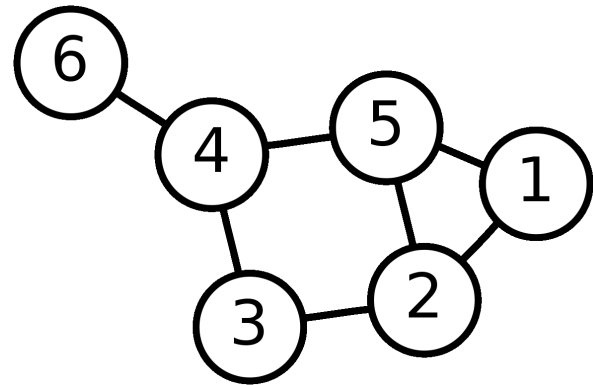
# Graphs

## Graph

In computer science, a **graph** is an abstract data type that is meant to implement the graph and hypergraph concepts from mathematics.

A graph data structure consists of a finite (and possibly mutable) set of ordered pairs, called **edges** or **arcs**, of certain entities called **nodes** or **vertices**. As in mathematics, an edge  $(x,y)$  is said to **point** or **go from**  $x$  **to**  $y$ . The nodes may be part of the graph structure, or may be external entities represented by integer indices or references.

A graph data structure may also associate to each edge some **edge value**, such as a symbolic label or a numeric attribute (cost, capacity, length, etc.).



A labeled graph of 6 vertices and 7 edges.

## Algorithms

Graph algorithms are a significant field of interest within computer science. Typical higher-level operations associated with graphs are: finding a path between two nodes, like depth-first search and breadth-first search and finding the shortest path from one node to another, like Dijkstra's algorithm. A solution to finding the shortest path from each node to every other node also exists in the form of the Floyd–Warshall algorithm.

A directed graph can be seen as a flow network, where each edge has a capacity and each edge receives a flow. The Ford–Fulkerson algorithm is used to find out the maximum flow from a source to a sink in a graph.

## Operations

The basic operations provided by a graph data structure  $G$  usually include:

- `adjacent( $G, x, y$ )`: tests whether there is an edge from node  $x$  to node  $y$ .
- `neighbors( $G, x$ )`: lists all nodes  $y$  such that there is an edge from  $x$  to  $y$ .
- `add( $G, x, y$ )`: adds to  $G$  the edge from  $x$  to  $y$ , if it is not there.
- `delete( $G, x, y$ )`: removes the edge from  $x$  to  $y$ , if it is there.
- `get_node_value( $G, x$ )`: returns the value associated with the node  $x$ .
- `set_node_value( $G, x, a$ )`: sets the value associated with the node  $x$  to  $a$ .

Structures that associate values to the edges usually also provide:

- `get_edge_value( $G, x, y$ )`: returns the value associated to the edge  $(x,y)$ .
- `set_edge_value( $G, x, y, v$ )`: sets the value associated to the edge  $(x,y)$  to  $v$ .

## Representations

Different data structures for the representation of graphs are used in practice:

### Adjacency list

Vertices are stored as records or objects, and every vertex stores a list of adjacent vertices. This data structure allows the storage of additional data on the vertices.

### Incidence list

Vertices and edges are stored as records or objects. Each vertex stores its incident edges, and each edge stores its incident vertices. This data structure allows the storage of additional data on vertices and edges.

### Adjacency matrix

A two-dimensional matrix, in which the rows represent source vertices and columns represent destination vertices. Data on edges and vertices must be stored externally. Only the cost for one edge can be stored between each pair of vertices.

### Incidence matrix

A two-dimensional Boolean matrix, in which the rows represent the vertices and columns represent the edges. The entries indicate whether the vertex at a row is incident to the edge at a column.

The following table gives the time complexity cost of performing various operations on graphs, for each of these representations. In the matrix representations, the entries encode the cost of following an edge. The cost of edges that are not present are assumed to be  $\infty$ .

	Adjacency list	Incidence list	Adjacency matrix	Incidence matrix
Storage	$O( V  +  E )$	$O( V  +  E )$	$O( V ^2)$	$O( V  \cdot  E )$
Add vertex	$O(1)$	$O(1)$	$O( V ^2)$	$O( V  \cdot  E )$
Add edge	$O(1)$	$O(1)$	$O(1)$	$O( V  \cdot  E )$
Remove vertex	$O( E )$	$O( E )$	$O( V ^2)$	$O( V  \cdot  E )$
Remove edge	$O( E )$	$O( E )$	$O(1)$	$O( V  \cdot  E )$
Query: are vertices $u, v$ adjacent? (Assuming that the storage positions for $u, v$ are known)	$O( V )$	$O( E )$	$O(1)$	$O( E )$
Remarks	When removing edges or vertices, need to find all vertices or edges		Slow to add or remove vertices, because matrix must be resized/copied	Slow to add or remove vertices and edges, because matrix must be resized/copied

Adjacency lists are generally preferred because they efficiently represent sparse graphs. An adjacency matrix is preferred if the graph is dense, that is the number of edges  $|E|$  is close to the number of vertices squared,  $|V|^2$ , or if one must be able to quickly look up if there is an edge connecting two vertices.<sup>[1]</sup>

Types

- Skip graphs

References

[1] Cormen, Thomas H.; Leiserson, Charles E.; Rivest, Ronald L.; Stein, Clifford (2001). Introduction to Algorithms (2nd ed.). MIT Press and McGraw–Hill. ISBN 0-262-53196-8.

External links

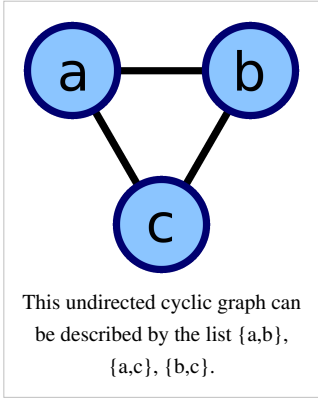
- Boost Graph Library: a powerful C++ graph library (<http://www.boost.org/libs/graph>)
- PulseView Graphing: an efficient Data-To-Graph program (<http://www.quarktet.com/PulseView.html>)

Adjacency list

In graph theory, an **adjacency list** is the representation of all edges or arcs in a graph as a list.

If the graph is undirected, every entry is a set (or multiset) of two nodes containing the two ends of the corresponding edge; if it is directed, every entry is a tuple of two nodes, one denoting the source node and the other denoting the destination node of the corresponding arc.

Typically, adjacency lists are unordered.



Application in computer science

The graph pictured above has this adjacency list representation:		
a	adjacent to	b,c
b	adjacent to	a,c
c	adjacent to	a,b

In computer science, an adjacency list is a data structure for representing graphs. In an adjacency list representation, we keep, for each vertex in the graph, a list of all other vertices which it has an edge to (that vertex's "adjacency list"). For instance, the representation suggested by van Rossum, in which a hash table is used to associate each vertex with an array of adjacent vertices, can be seen as an example of this type of representation. Another example is the representation in Cormen et al. in which an array indexed by vertex numbers points to a singly linked list of the neighbors of each vertex.

One difficulty with the adjacency list structure is that it has no obvious place to store data associated with the edges of a graph, such as the lengths or costs of the edges. To remedy this, some texts, such as that of Goodrich and Tamassia, advocate a more object oriented variant of the adjacency list structure, sometimes called an incidence list, which stores for each vertex a list of objects representing the edges incident to that vertex. To complete the structure, each edge must point back to the two vertices forming its endpoints. The extra edge objects in this version of the adjacency list cause it to use more memory than the version in which adjacent vertices are listed directly, but these



extra edges are also a convenient location to store additional information about each edge (e.g. their length).

## Trade-offs

The main alternative to the adjacency list is the adjacency matrix. For a graph with a sparse adjacency matrix an adjacency list representation of the graph occupies less space, because it does not use any space to represent edges that are *not* present. Using a naive array implementation of adjacency lists on a 32-bit computer, an adjacency list for an undirected graph requires about  $8e$  bytes of storage, where  $e$  is the number of edges: each edge gives rise to entries in the two adjacency lists and uses four bytes in each.

On the other hand, because each entry in an adjacency matrix requires only one bit, they can be represented in a very compact way, occupying only  $n^2/8$  bytes of contiguous space, where  $n$  is the number of vertices. Besides just avoiding wasted space, this compactness encourages locality of reference.

Noting that a graph can have at most  $n^2$  edges (allowing loops) we can let  $d = e/n^2$  denote the *density* of the graph. Then, if  $8e > n^2/8$ , the adjacency list representation occupies more space, which is true when  $d > 1/64$ . Thus a graph must be sparse for an adjacency list representation to be more memory efficient than an adjacency matrix. However, this analysis is valid only when the representation is intended to store the connectivity structure of the graph without any numerical information about its edges.

Besides the space trade-off, the different data structures also facilitate different operations. It is easy to find all vertices adjacent to a given vertex in an adjacency list representation; you simply read its adjacency list. With an adjacency matrix you must instead scan over an entire row, taking  $O(n)$  time. If you, instead, want to perform a neighbor test on two vertices (i.e., determine if they have an edge between them), an adjacency matrix provides this at once. However, this neighbor test in an adjacency list requires time proportional to the number of edges associated with the two vertices.

## References

- Joe Celko (2004). *Trees and Hierarchies in SQL for Smarties*. Morgan Kaufmann. excerpt from Chapter 2: "Adjacency List Model" <sup>[1]</sup>. ISBN 1-55860-920-2.
  - Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein (2001). *Introduction to Algorithms, Second Edition*. MIT Press and McGraw-Hill. pp. 527–529 of section 22.1: Representations of graphs. ISBN 0-262-03293-7.
  - David Eppstein (1996). "ICS 161 Lecture Notes: Graph Algorithms" <sup>[2]</sup>.
  - Michael T. Goodrich and Roberto Tamassia (2002). *Algorithm Design: Foundations, Analysis, and Internet Examples*. John Wiley & Sons. ISBN 0-471-38365-1.
  - Guido van Rossum (1998). "Python Patterns — Implementing Graphs" <sup>[3]</sup>.
-

External links

- The Boost Graph Library implements an efficient adjacency list <sup>[4]</sup>
- Open Data Structures - Section 12.2 - AdjacencyList: A Graph as a Collection of Lists <sup>[5]</sup>

References

[1] <http://www.SQLSummit.com/AdjacencyList.htm>  
[2] <http://www.ics.uci.edu/~eppstein/161/960201.html>  
[3] <http://www.python.org/doc/essays/graphs/>  
[4] [http://www.boost.org/doc/libs/1\\_43\\_0/libs/graph/doc/index.html](http://www.boost.org/doc/libs/1_43_0/libs/graph/doc/index.html)  
[5] [http://opendatastructures.org/versions/edition-0.1e/ods-java/12\\_2\\_AdjacencyLists\\_Graph\\_a.html](http://opendatastructures.org/versions/edition-0.1e/ods-java/12_2_AdjacencyLists_Graph_a.html)

Adjacency matrix

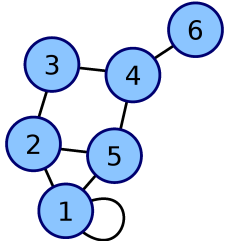
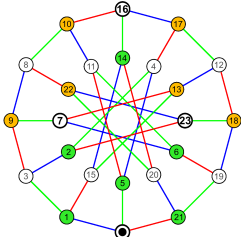
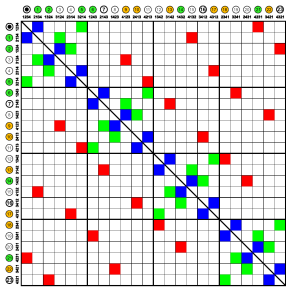
In mathematics and computer science, an **adjacency matrix** is a means of representing which vertices (or nodes) of a graph are adjacent to which other vertices. Another matrix representation for a graph is the incidence matrix.

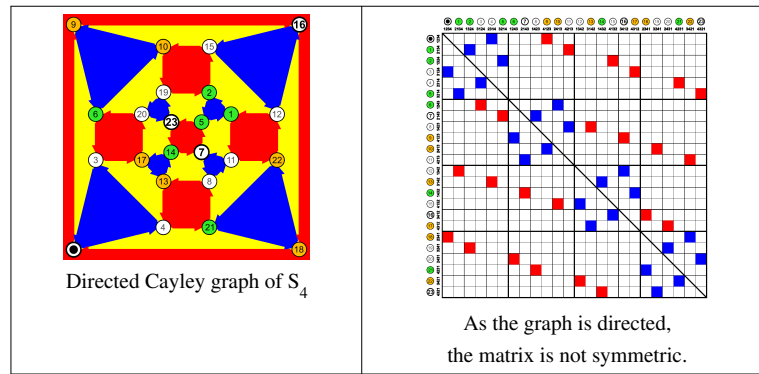
Specifically, the adjacency matrix of a finite graph **G** on *n* vertices is the *n* × *n* matrix where the non-diagonal entry *a*<sub>*ij*</sub> is the number of edges from vertex *i* to vertex *j*, and the diagonal entry *a*<sub>*ii*</sub>, depending on the convention, is either once or twice the number of edges (loops) from vertex *i* to itself. Undirected graphs often use the latter convention of counting loops twice, whereas directed graphs typically use the former convention. There exists a unique adjacency matrix for each isomorphism class of graphs (up to permuting rows and columns), and it is not the adjacency matrix of any other isomorphism class of graphs. In the special case of a finite simple graph, the adjacency matrix is a (0,1)-matrix with zeros on its diagonal. If the graph is undirected, the adjacency matrix is symmetric.

The relationship between a graph and the eigenvalues and eigenvectors of its adjacency matrix is studied in spectral graph theory.

Examples

The convention followed here is that an adjacent edge counts 1 in the matrix for an undirected graph.

Labeled graph	Adjacency matrix
	$\begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$ <p>Coordinates are 1-6.</p>
 <p>The Nauru graph</p>	 <p>Coordinates are 0-23. White fields are zeros, colored fields are ones.</p>



- The adjacency matrix of a complete graph is all 1's except for 0's on the diagonal.
- The adjacency matrix of an empty graph is a zero matrix.

## Adjacency matrix of a bipartite graph

The adjacency matrix  $A$  of a bipartite graph whose parts have  $r$  and  $s$  vertices has the form

$$A = \begin{pmatrix} O & B \\ B^T & O \end{pmatrix},$$

where  $B$  is an  $r \times s$  matrix and  $O$  is an all-zero matrix. Clearly, the matrix  $B$  uniquely represents the bipartite graphs. It is sometimes called the biadjacency matrix. Formally, let  $G = (U, V, E)$  be a bipartite graph with parts  $U = u_1, \dots, u_r$  and  $V = v_1, \dots, v_s$ . The **biadjacency matrix** is the  $r \times s$  0-1 matrix  $B$  in which  $b_{i,j} = 1$  iff  $(u_i, v_j) \in E$ .

If  $G$  is a bipartite multigraph or weighted graph then the elements  $b_{i,j}$  are taken to be the number of edges between the vertices or the weight of the edge  $(u_i, v_j)$ , respectively.

## Properties

The adjacency matrix of an undirected simple graph is symmetric, and therefore has a complete set of real eigenvalues and an orthogonal eigenvector basis. The set of eigenvalues of a graph is the **spectrum** of the graph.

Suppose two directed or undirected graphs  $G_1$  and  $G_2$  with adjacency matrices  $A_1$  and  $A_2$  are given.  $G_1$  and  $G_2$  are isomorphic if and only if there exists a permutation matrix  $P$  such that

$$PA_1P^{-1} = A_2.$$

In particular,  $A_1$  and  $A_2$  are similar and therefore have the same minimal polynomial, characteristic polynomial, eigenvalues, determinant and trace. These can therefore serve as isomorphism invariants of graphs. However, two graphs may possess the same set of eigenvalues but not be isomorphic. <sup>[1]</sup>

If  $A$  is the adjacency matrix of the directed or undirected graph  $G$ , then the matrix  $A^n$  (i.e., the matrix product of  $n$  copies of  $A$ ) has an interesting interpretation: the entry in row  $i$  and column  $j$  gives the number of (directed or undirected) walks of length  $n$  from vertex  $i$  to vertex  $j$ . This implies, for example, that the number of triangles in an undirected graph  $G$  is exactly the trace of  $A^3$  divided by 6.

The main diagonal of every adjacency matrix corresponding to a graph without loops has all zero entries. Note that here 'loops' means, for example  $A \rightarrow A$ , not 'cycles' such as  $A \rightarrow B \rightarrow A$ .

For  $(d)$ -regular graphs,  $d$  is also an eigenvalue of  $A$  for the vector  $v = (1, \dots, 1)$ , and  $G$  is connected if and only if the multiplicity of  $d$  is 1. It can be shown that  $-d$  is also an eigenvalue of  $A$  if  $G$  is a connected bipartite graph. The above are results of Perron–Frobenius theorem.

## Variations

An **( $a, b, c$ )-adjacency matrix**  $A$  of a simple graph has  $A_{ij} = a$  if  $ij$  is an edge,  $b$  if it is not, and  $c$  on the diagonal. The Seidel adjacency matrix is a **( $-1, 1, 0$ )-adjacency matrix**. This matrix is used in studying strongly regular graphs and two-graphs.<sup>[2]</sup>

The **distance matrix** has in position  $(i, j)$  the distance between vertices  $v_i$  and  $v_j$ . The distance is the length of a shortest path connecting the vertices. Unless lengths of edges are explicitly provided, the length of a path is the number of edges in it. The distance matrix resembles a high power of the adjacency matrix, but instead of telling only whether or not two vertices are connected (i.e., the connection matrix, which contains boolean values), it gives the exact distance between them.

## Data structures

For use as a data structure, the main alternative to the adjacency matrix is the adjacency list. Because each entry in the adjacency matrix requires only one bit, it can be represented in a very compact way, occupying only  $n^2/8$  bytes of contiguous space, where  $n$  is the number of vertices. Besides avoiding wasted space, this compactness encourages locality of reference.

However, if the graph is sparse, adjacency lists require less storage space, because they do not waste any space to represent edges that are *not* present. Using a naïve array implementation on a 32-bit computer, an adjacency list for an undirected graph requires about  $8e$  bytes of storage, where  $e$  is the number of edges.

Noting that a simple graph can have at most  $n^2$  edges, allowing loops, we can let  $d = e/n^2$  denote the *density* of the graph. Then,  $8e > n^2/8$ , or the adjacency list representation occupies more space precisely when  $d > 1/64$ .

Thus a graph must be sparse indeed to justify an adjacency list representation.

Besides the space tradeoff, the different data structures also facilitate different operations. Finding all vertices adjacent to a given vertex in an adjacency list is as simple as reading the list. With an adjacency matrix, an entire row must instead be scanned, which takes  $O(n)$  time. Whether there is an edge between two given vertices can be determined at once with an adjacency matrix, while requiring time proportional to the minimum degree of the two vertices with the adjacency list.

## References

- [1] Godsil, Chris; Royle, Gordon *Algebraic Graph Theory*, Springer (2001), ISBN 0-387-95241-1, p.164
- [2] Seidel, J. J. (1968). "Strongly Regular Graphs with  $(-1, 1, 0)$  Adjacency Matrix Having Eigenvalue 3". *Lin. Alg. Appl.* **1** (2): 281–298. doi:10.1016/0024-3795(68)90008-6.

## Further reading

- Cormen, Thomas H.; Leiserson, Charles E.; Rivest, Ronald L.; Stein, Clifford (2001). "Section 22.1: Representations of graphs". *Introduction to Algorithms* (Second ed.). MIT Press and McGraw-Hill. pp. 527–531. ISBN 0-262-03293-7.
- Godsil, Chris; Royle, Gordon (2001). *Algebraic Graph Theory*. New York: Springer. ISBN 0-387-95241-1.

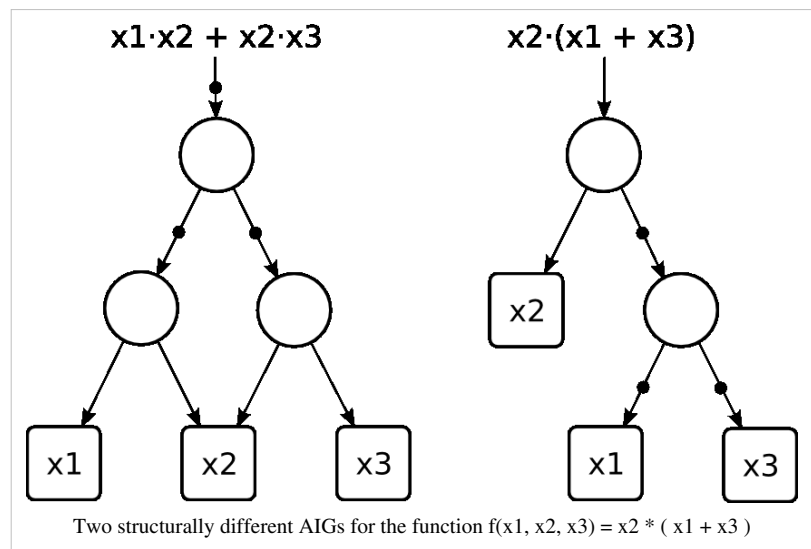
## External links

- Fluffsack (<http://www.x2d.org/java/projects/fluffsack.jnlp>) — an educational Java web start game demonstrating the relationship between adjacency matrices and graphs.
- Open Data Structures - Section 12.1 - AdjacencyMatrix: Representing a Graph by a Matrix ([http://opendatastructures.org/versions/edition-0.1e/ods-java/12\\_1\\_AdjacencyMatrix\\_Repres.html](http://opendatastructures.org/versions/edition-0.1e/ods-java/12_1_AdjacencyMatrix_Repres.html))
- McKay, Brendan. "Description of graph6 and sparse6 encodings" (<http://cs.anu.edu.au/~bdm/data/formats.txt>).
- Café math : Adjacency Matrices of Graphs (<http://cafemath.kegtux.org/mathblog/article.php?page=GoodWillHunting.php>) : Application of the adjacency matrices to the computation generating series of walks.

## And-inverter graph

An **and-inverter graph (AIG)** is a directed, acyclic graph that represents a structural implementation of the logical functionality of a circuit or network. An AIG consists of two-input nodes representing logical conjunction, terminal nodes labeled with variable names, and edges optionally containing markers indicating logical negation. This representation of a logic function is rarely structurally efficient for large circuits, but is an efficient representation for manipulation of boolean functions. Typically, the abstract graph is represented as a data structure in software.

Conversion from the network of logic gates to AIGs is fast and scalable. It only requires that every gate be expressed in terms of AND gates and inverters. This conversion does not lead to unpredictable increase in memory use and runtime. This makes the AIG an efficient representation in comparison with either the binary decision diagram (BDD) or the "sum-of-product" ( $\Sigma\Pi$ ) form, that is, the canonical form in Boolean algebra known as the disjunctive normal form (DNF). The BDD and DNF may also be viewed as circuits, but they involve



formal constraints that deprive them of scalability. For example,  $\Sigma\Pi$ s are circuits with at most two levels while BDDs are canonical, that is, they require that input variables be evaluated in the same order on all paths.

Circuits composed of simple gates, including AIGs, are an "ancient" research topic. The interest in AIGs started in the late 1950s<sup>[1]</sup> and continued in the 1970s when various local transformations have been developed. These transformations were implemented in several logic synthesis and verification systems, such as Darringer et al.<sup>[2]</sup> and Smith et al.<sup>[3]</sup>, which reduce circuits to improve area and delay during synthesis, or to speed up formal equivalence checking. Several important techniques were discovered early at IBM, such as combining and reusing multi-input logic expressions and subexpressions, now known as structural hashing.

Recently there has been a renewed interest in AIGs as a functional representation for a variety of tasks in synthesis and verification. That is because representations popular in the 1990s (such as BDDs) have reached their limits of scalability in many of their applications. Another important development was the recent emergence of much more efficient boolean satisfiability (SAT) solvers. When coupled with AIGs as the circuit representation, they lead to

remarkable speedups in solving a wide variety of boolean problems.

AIGs found successful use in diverse EDA applications. A well-tuned combination of AIGs and boolean satisfiability made an impact on formal verification, including both model checking and equivalence checking.<sup>[4]</sup> Another recent work shows that efficient circuit compression techniques can be developed using AIGs.<sup>[5]</sup> There is a growing understanding that logic and physical synthesis problems can be solved using AIGs simulation and boolean satisfiability compute functional properties (such as symmetries<sup>[6]</sup>) and node flexibilities (such as don't-cares, resubstitutions, and SPFDs<sup>[7]</sup>). This work shows that AIGs are a promising *unifying* representation, which can bridge logic synthesis, technology mapping, physical synthesis, and formal verification. This is, to a large extent, due to the simple and uniform structure of AIGs, which allow rewriting, simulation, mapping, placement, and verification to share the same data structure.

In addition to combinational logic, AIGs have also been applied to sequential logic and sequential transformations. Specifically, the method of structural hashing was extended to work for AIGs with memory elements (such as D-type flip-flops with an initial state, which, in general, can be unknown) resulting in a data structure that is specifically tailored for applications related to retiming.<sup>[8]</sup>

Ongoing research includes implementing a modern logic synthesis system completely based on AIGs. The prototype called ABC<sup>[9]</sup> features an AIG package, several AIG-based synthesis and equivalence-checking techniques, as well as an experimental implementation of sequential synthesis. One such technique combines technology mapping and retiming in a single optimization step. These optimizations can be implemented using networks composed of arbitrary gates, but the use of AIGs makes them more scalable and easier to implement.

## Implementations

- Logic Synthesis and Verification System ABC<sup>[9]</sup>
- A set of utilities for AIGs AIGER<sup>[10]</sup>
- OpenAccess Gear<sup>[11]</sup>

## References

- [1] L. Hellerman (June 1963). "A catalog of three-variable Or-Inverter and And-Inverter logical circuits". *IEEE Trans. Electron. Comput.* **EC-12** (3): 198–223. doi:10.1109/PGEC.1963.263531.
- [2] A. Darringer, W. H. Joyner, Jr., C. L. Berman, L. Trevillyan (1981). "Logic synthesis through local transformations". *IBM J. of Research and Development* **25** (4): 272–280. doi:10.1147/rd.254.0272.
- [3] G. L. Smith, R. J. Bahnsen, H. Halliwell (1982). "Boolean comparison of hardware and flowcharts". *IBM J. of Research and Development* **26** (1): 106–116. doi:10.1147/rd.261.0106.
- [4] A. Kuehlmann, V. Paruthi, F. Krohm, and M. K. Ganai (2002). "Robust boolean reasoning for equivalence checking and functional property verification". *IEEE Trans. CAD* **21** (12): 1377–1394.
- [5] P. Bjesse and A. Boralv. "DAG-aware circuit compression for formal verification". *Proc. ICCAD '04*. pp. 42–49.
- [6] K.-H. Chang, I. L. Markov, V. Bertacco. "Post-placement rewiring and rebuffering by exhaustive search for functional symmetries". *Proc. ICCAD '05* pages=56–63.
- [7] A. Mishchenko, J. S. Zhang, S. Sinha, J. R. Burch, R. Brayton, and M. Chrzanoska-Jeske (May 2006). "Using simulation and satisfiability to compute flexibilities in Boolean networks". *IEEE Trans. CAD* **25** (5): 743–755.
- [8] J. Baumgartner and A. Kuehlmann. "Min-area retiming on flexible circuit structures". *Proc. ICCAD'01*. pp. 176–182.
- [9] <http://www.eecs.berkeley.edu/~alanmi/abc/>
- [10] <http://fmv.jku.at/aiger/index.html>
- [11] [http://www.si2.org/openeda.si2.org/help/group\\_ld.php?group=73](http://www.si2.org/openeda.si2.org/help/group_ld.php?group=73)

# Binary decision diagram

In the field of computer science, a **binary decision diagram (BDD)** or **branching program**, like a negation normal form (NNF) or a propositional directed acyclic graph (PDAG), is a data structure that is used to represent a Boolean function. On a more abstract level, BDDs can be considered as a compressed representation of sets or relations. Unlike other compressed representations, operations are performed directly on the compressed representation, i.e. without decompression.

## Definition

A Boolean function can be represented as a rooted, directed, acyclic graph, which consists of several decision nodes and terminal nodes. There are two types of terminal nodes called 0-terminal and 1-terminal. Each decision node  $N$  is labeled by Boolean variable  $V_N$  and has two child nodes called low child and high child. The edge from node  $V_N$  to a low (or high) child represents an assignment of  $V_N$  to 0 (resp. 1). Such a **BDD** is called 'ordered' if different variables appear in the same order on all paths from the root. A BDD is said to be 'reduced' if the following two rules have been applied to its graph:

- Merge any isomorphic subgraphs.
- Eliminate any node whose two children are isomorphic.

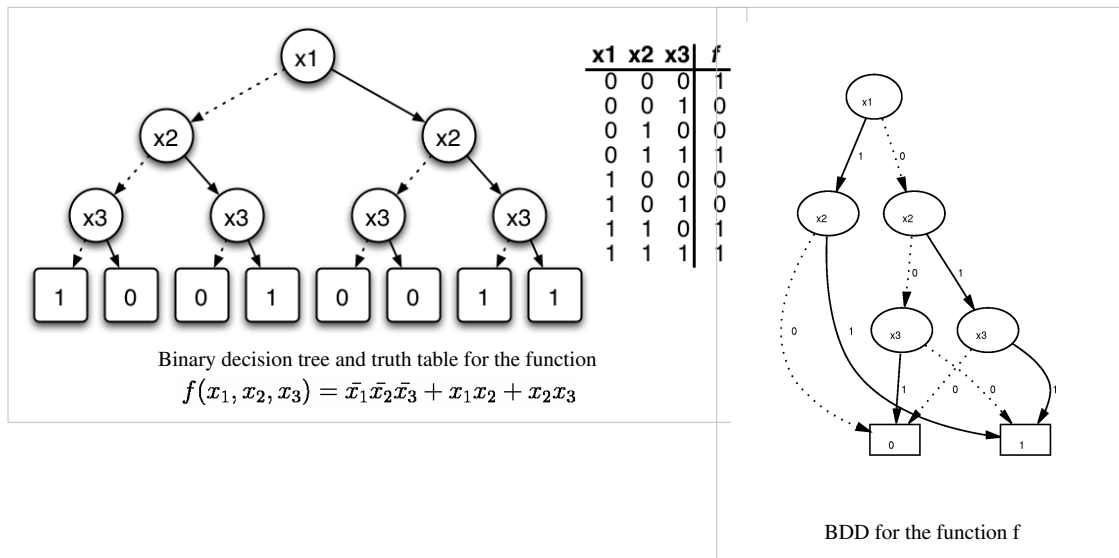
In popular usage, the term **BDD** almost always refers to **Reduced Ordered Binary Decision Diagram (ROBDD)** in the literature, used when the ordering and reduction aspects need to be emphasized). The advantage of an ROBDD is that it is canonical (unique) for a particular function and variable order.<sup>[1]</sup> This property makes it useful in functional equivalence checking and other operations like functional technology mapping.

A path from the root node to the 1-terminal represents a (possibly partial) variable assignment for which the represented Boolean function is true. As the path descends to a low (or high) child from a node, then that node's variable is assigned to 0 (resp. 1).

## Example

The left figure below shows a binary decision *tree* (the reduction rules are not applied), and a truth table, each representing the function  $f(x_1, x_2, x_3)$ . In the tree on the left, the value of the function can be determined for a given variable assignment by following a path down the graph to a terminal. In the figures below, dotted lines represent edges to a low child, while solid lines represent edges to a high child. Therefore, to find  $(x_1=0, x_2=1, x_3=1)$ , begin at  $x_1$ , traverse down the dotted line to  $x_2$  (since  $x_1$  has an assignment to 0), then down two solid lines (since  $x_2$  and  $x_3$  each have an assignment to one). This leads to the terminal 1, which is the value of  $f(x_1=0, x_2=1, x_3=1)$ .

The binary decision *tree* of the left figure can be transformed into a binary decision *diagram* by maximally reducing it according to the two reduction rules. The resulting **BDD** is shown in the right figure.



## History

The basic idea from which the data structure was created is the Shannon expansion. A switching function is split into two sub-functions (cofactors) by assigning one variable (cf. *if-then-else normal form*). If such a sub-function is considered as a sub-tree, it can be represented by a *binary decision tree*. Binary decision diagrams (BDD) were introduced by Lee,<sup>[2]</sup> and further studied and made known by Akers<sup>[3]</sup> and Boute.<sup>[4]</sup>

The full potential for efficient algorithms based on the data structure was investigated by Randal Bryant at Carnegie Mellon University: his key extensions were to use a fixed variable ordering (for canonical representation) and shared sub-graphs (for compression). Applying these two concepts results in an efficient data structure and algorithms for the representation of sets and relations.<sup>[5][6]</sup> By extending the sharing to several BDDs, i.e. one sub-graph is used by several BDDs, the data structure *Shared Reduced Ordered Binary Decision Diagram* is defined.<sup>[7]</sup> The notion of a BDD is now generally used to refer to that particular data structure.

In his video lecture *Fun With Binary Decision Diagrams (BDDs)*<sup>[8]</sup>, Donald Knuth calls BDDs "one of the only really fundamental data structures that came out in the last twenty-five years" and mentions that Bryant's 1986 paper was for some time one of the most-cited papers in computer science.

## Applications

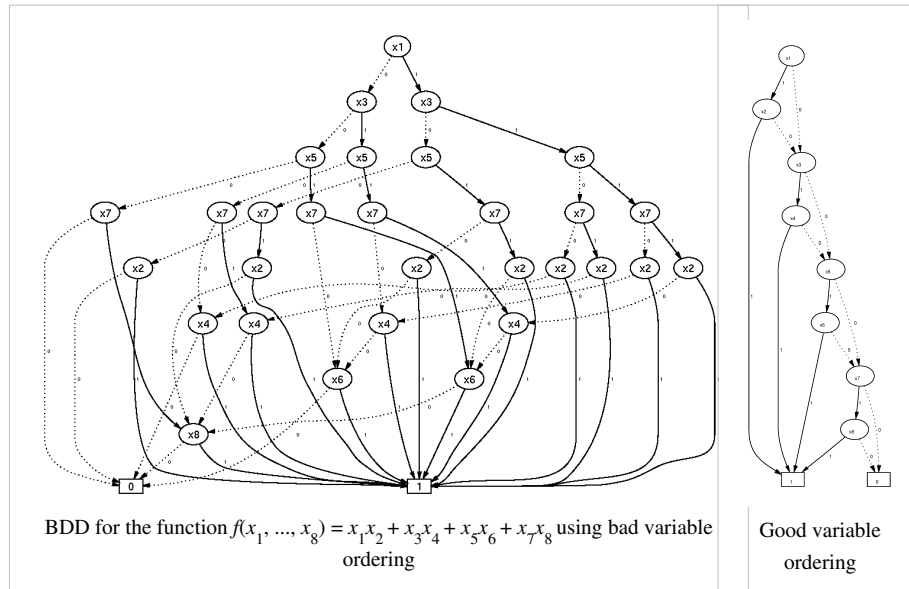
BDDs are extensively used in CAD software to synthesize circuits (logic synthesis) and in formal verification. There are several lesser known applications of BDD, including Fault tree analysis, Bayesian Reasoning, Product Configuration, and Private information retrieval<sup>[9][10]</sup>.

Every arbitrary BDD (even if it is not reduced or ordered) can be directly implemented by replacing each node with a 2 to 1 multiplexer; each multiplexer can be directly implemented by a 4-LUT in a FPGA. It is not so simple to convert from an arbitrary network of logic gates to a BDD (unlike the and-inverter graph).



## Variable ordering

The size of the BDD is determined both by the function being represented and the chosen ordering of the variables. There exist Boolean functions  $f(x_1, \dots, x_n)$  for which depending upon the ordering of the variables we would end up getting a graph whose number of nodes would be linear (in  $n$ ) at the best and exponential at the worst case (e.g., a ripple carry adder). Let us consider the Boolean function  $f(x_1, \dots, x_{2n}) = x_1x_2 + x_3x_4 + \dots + x_{2n-1}x_{2n}$ . Using the variable ordering  $x_1 < x_3 < \dots < x_{2n-1} < x_2 < x_4 < \dots < x_{2n}$ , the BDD needs  $2^{n+1}$  nodes to represent the function. Using the ordering  $x_1 < x_2 < x_3 < x_4 < \dots < x_{2n-1} < x_{2n}$ , the BDD consists of  $2n+2$  nodes.



It is of crucial importance to care about variable ordering when applying this data structure in practice. The problem of finding the best variable ordering is NP-hard.<sup>[11]</sup> For any constant  $c > 1$  it is even NP-hard to compute a variable ordering resulting in an OBDD with a size that is at most  $c$  times larger than an optimal one.<sup>[12]</sup> However there exist efficient heuristics to tackle the problem.

There are functions for which the graph size is always exponential — independent of variable ordering. This holds e. g. for the multiplication function (an indication as to the apparent complexity of factorization ).

Researchers have of late suggested refinements on the BDD data structure giving way to a number of related graphs, such as BMD (Binary Moment Diagrams), ZDD (Zero Suppressed Decision Diagram), FDD (Free Binary Decision Diagrams), PDD (Parity decision Diagrams), and MTBDDs (Multiple terminal BDDs).

## Logical operations on BDDs

Many logical operations on BDDs can be implemented by polynomial-time graph manipulation algorithms.

- conjunction
- disjunction
- negation
- existential abstraction
- universal abstraction

However, repeating these operations several times, for example forming the conjunction or disjunction of a set of BDDs, may in the worst case result in an exponentially big BDD. This is because any of the preceding operations for two BDDs may result in a BDD with a size proportional to the product of the BDDs' sizes, and consequently for

several BDDs the size may be exponential.

## References

- [1] Graph-Based Algorithms for Boolean Function Manipulation, Randal E. Bryant, 1986
- [2] C. Y. Lee. "Representation of Switching Circuits by Binary-Decision Programs". Bell Systems Technical Journal, 38:985–999, 1959.
- [3] Sheldon B. Akers. Binary Decision Diagrams (<http://ieeexplore.ieee.org/search/wrapper.jsp?arnumber=1675141>), IEEE Transactions on Computers, C-27(6):509–516, June 1978.
- [4] Raymond T. Boute, "The Binary Decision Machine as a programmable controller". EUROMICRO Newsletter, Vol. 1(2):16–22, January 1976.
- [5] Randal E. Bryant. "Graph-Based Algorithms for Boolean Function Manipulation (<http://www.cs.cmu.edu/~bryant/pubdir/ieeetc86.ps>)". IEEE Transactions on Computers, C-35(8):677–691, 1986.
- [6] R. E. Bryant, "Symbolic Boolean Manipulation with Ordered Binary Decision Diagrams" (<http://www.cs.cmu.edu/~bryant/pubdir/acmcs92.ps>), ACM Computing Surveys, Vol. 24, No. 3 (September, 1992), pp. 293–318.
- [7] Karl S. Brace, Richard L. Rudell and Randal E. Bryant. "Efficient Implementation of a BDD Package" (<http://portal.acm.org/citation.cfm?id=123222&coll=portal&dl=ACM>). In Proceedings of the 27th ACM/IEEE Design Automation Conference (DAC 1990), pages 40–45. IEEE Computer Society Press, 1990.
- [8] <http://myvideos.stanford.edu/player/slplayer.aspx?coll=ea60314a-53b3-4be2-8552-dcf190ca0c0b&co=18bcd3a8-965a-4a63-a516-a1ad74af1119&o=true>
- [9] R.M. Jensen. "CLab: A C++ library for fast backtrack-free interactive product configuration" (<http://www.cs.cmu.edu/~runej/data/papers/JSW04.pdf>). Proceedings of the Tenth International Conference on Principles and Practice of Constraint Programming, 2004.
- [10] H.L. Lipmaa. "First CIPR Protocol with Data-Dependent Computation" (<http://eprint.iacr.org/2009/395.pdf>). ICISC 2009.
- [11] Beate Bollig, Ingo Wegener. Improving the Variable Ordering of OBDDs Is NP-Complete (doi:10.1109/12.537122), IEEE Transactions on Computers, 45(9):993–1002, September 1996.
- [12] Detlef Sieling. "The nonapproximability of OBDD minimization." Information and Computation 172, 103–138. 2002.
- R. Ubar, "Test Generation for Digital Circuits Using Alternative Graphs (in Russian)", in Proc. Tallinn Technical University, 1976, No.409, Tallinn Technical University, Tallinn, Estonia, pp. 75–81.

## Further reading

- D. E. Knuth, "The Art of Computer Programming Volume 4, Fascicle 1: Bitwise tricks & techniques; Binary Decision Diagrams" (Addison–Wesley Professional, March 27, 2009) viii+260pp, ISBN 0-321-58050-8. Draft of Fascicle 1b (<http://www-cs-faculty.stanford.edu/~knuth/fasc1b.ps.gz>) available for download.
- H. R. Andersen "An Introduction to Binary Decision Diagrams (<http://www.configit.com/fileadmin/Configit/Documents/bdd-eap.pdf>)," Lecture Notes, 1999, IT University of Copenhagen.
- Ch. Meinel, T. Theobald, "Algorithms and Data Structures in VLSI-Design: OBDD – Foundations and Applications" ([http://www.hpi.uni-potsdam.de/fileadmin/hpi/FG\\_ITS/books/OBDD-Book.pdf](http://www.hpi.uni-potsdam.de/fileadmin/hpi/FG_ITS/books/OBDD-Book.pdf)), Springer-Verlag, Berlin, Heidelberg, New York, 1998. Complete textbook available for download.
- Rüdiger Ebendt; Görschwin Fey; Rolf Drechsler (2005). *Advanced BDD optimization*. Springer. ISBN 978-0-387-25453-1.
- Bernd Becker; Rolf Drechsler (1998). *Binary Decision Diagrams: Theory and Implementation*. Springer. ISBN 978-1-4419-5047-5.

## External links

### Available OBDD Packages

- ABCD (<http://fmv.jku.at/abcd/>): The ABCD package by Armin Biere, Johannes Kepler Universität, Linz.
- CMU BDD (<http://www-2.cs.cmu.edu/~modelcheck/bdd.html>), BDD package, Carnegie Mellon University, Pittsburgh
- CUDD (<http://vlsi.colorado.edu/~fabio/CUDD/>): BDD package, University of Colorado, Boulder
- Installing CUDD in Windows/Visual Studio environments. (<http://www.technical-recipes.com/category/binary-decision-diagrams/>)
- Biddy (<http://lms.uni-mb.si/biddy/>): multi-platform academic BDD package, University of Maribor, Slovenia
- JavaBDD (<http://javabdd.sourceforge.net>), a Java port of BuDDy that also interfaces to CUDD, CAL, and JDD
- The Berkeley CAL ([http://embedded.eecs.berkeley.edu/Research/cal\\_bdd/](http://embedded.eecs.berkeley.edu/Research/cal_bdd/)) package which does breadth-first manipulation
- A. Costa BFunc (<http://www.dei.isep.ipp.pt/~acc/bfunc/>), includes a BDD boolean logic simplifier supporting up to 32 inputs / 32 outputs (independently)
- DDD (<http://ddd.lip6.fr>): A C++ library with support for integer valued and hierarchical decision diagrams.
- JINC (<http://www.jossowski.de/projects/jinc/jinc.html>): A C++ library developed at University of Bonn, Germany, supporting several BDD variants and multi-threading.

## Binary moment diagram

---

A **binary moment diagram (BMD)** is a generalization of the binary decision diagram (BDD) to linear functions over domains such as booleans (like BDDs), but also to integers or to real numbers.

They can deal with boolean functions with complexity comparable to BDDs, but also some functions that are dealt with very inefficiently in a BDD are handled easily by BMD, most notably multiplication.

The most important properties of BMD is that, like with BDDs, each function has exactly one canonical representation, and many operations can be efficiently performed on these representations.

The main features that differentiate BMDs from BDDs are using linear instead of pointwise diagrams, and having weighted edges.

The rules that ensure the canonicity of the representation are:

- Decision over variables higher in the ordering may only point to decisions over variables lower in the ordering.
  - No two nodes may be identical (in normalization such nodes all references to one of these nodes should be replaced by references to another)
  - No node may have all decision parts equivalent to 0 (links to such nodes should be replaced by links to their always part)
  - No edge may have weight zero (all such edges should be replaced by direct links to 0)
  - Weights of the edges should be coprime. Without this rule or some equivalent of it, it would be possible for a function to have many representations, for example  $2x + 2$  could be represented as  $2 \cdot (1 + x)$  or  $1 \cdot (2 + 2x)$ .
-

## Pointwise and linear decomposition

In pointwise decomposition, like in BDDs, on each branch point we store result of all branches separately. An example of such decomposition for an integer function  $(2x + y)$  is:

$$\begin{cases} \text{if } x \begin{cases} \text{if } y, 3 \\ \text{if } \neg y, 2 \end{cases} \\ \text{if } \neg x \begin{cases} \text{if } y, 1 \\ \text{if } \neg y, 0 \end{cases} \end{cases}$$

In linear decomposition we provide instead a default value and a difference:

$$\begin{cases} \text{always} \begin{cases} \text{always } 0 \\ \text{if } y, +1 \end{cases} \\ \text{if } x, +2 \end{cases}$$

It can easily be seen that the latter (linear) representation is much more efficient in case of additive functions, as when we add many elements the latter representation will have only  $O(n)$  elements, while the former (pointwise), even with sharing, exponentially many.

## Edge weights

Another extension is using weights for edges. A value of function at given node is a sum of the true nodes below it (the node under always, and possibly the decided node) times the edges' weights.

For example  $(4x_2 + 2x_1 + x_0)(4y_2 + 2y_1 + y_0)$  can be represented as:

1. Result node, always  $1 \times$  value of node 2, if  $x_2$  add  $4 \times$  value of node 4
2. Always  $1 \times$  value of node 3, if  $x_1$  add  $2 \times$  value of node 4
3. Always 0, if  $x_0$  add  $1 \times$  value of node 4
4. Always  $1 \times$  value of node 5, if  $y_2$  add  $+4$
5. Always  $1 \times$  value of node 6, if  $y_1$  add  $+2$
6. Always 0, if  $y_0$  add  $+1$

Without weighted nodes a much more complex representation would be required:

1. Result node, always value of node 2, if  $x_2$  value of node 4
2. Always value of node 3, if  $x_1$  value of node 7
3. Always 0, if  $x_0$  value of node 10
4. Always value of node 5, if  $y_2$  add  $+16$
5. Always value of node 6, if  $y_1$  add  $+8$
6. Always 0, if  $y_0$  add  $+4$
7. Always value of node 8, if  $y_2$  add  $+8$
8. Always value of node 9, if  $y_1$  add  $+4$
9. Always 0, if  $y_0$  add  $+2$
10. Always value of node 11, if  $y_2$  add  $+4$
11. Always value of node 12, if  $y_1$  add  $+2$
12. Always 0, if  $y_0$  add  $+1$

# Zero-suppressed decision diagram

---

A **zero-suppressed decision diagram** (ZSDD or ZDD) is a type of binary decision diagram (BDD) where instead of nodes being introduced when the positive and the negative part are different, they are introduced when negative part is different from constant 0. A Zero-suppressed decision diagram is also commonly referred to as a **zero-suppressed binary decision diagram** (ZBDD).

They are useful when dealing with functions that are almost everywhere 0.

## Available packages

- CUDD<sup>[1]</sup>: A BDD package written in C that implements BDDs and ZBDDs, University of Colorado, Boulder
- JDD<sup>[2]</sup>, A java library that implements common BDD and ZBDD operations

## References

[1] <http://vlsi.colorado.edu/~fabio/CUDD/>

[2] <http://javaddlib.sourceforge.net/jdd/>

- Shin-ichi Minato, " Zero-suppressed BDDs for set manipulation in combinatorial problems ([http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=1600231](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1600231))", DAC '93: Proceedings of the 30th international conference on Design automation, 1993
- Ch. Meinel, T. Theobald, " Algorithms and Data Structures in VLSI-Design: OBDD - Foundations and Applications" ([http://www.hpi.uni-potsdam.de/fileadmin/hpi/FG\\_ITS/books/OBDD-Book.pdf](http://www.hpi.uni-potsdam.de/fileadmin/hpi/FG_ITS/books/OBDD-Book.pdf)), Springer-Verlag, Berlin, Heidelberg, New York, 1998.

## External links

- Alan Mishchenko, An Introduction to Zero-Suppressed Binary Decision Diagrams ([http://www.eecs.berkeley.edu/~alanmi/publications/2001/tech01\\_zdd.pdf](http://www.eecs.berkeley.edu/~alanmi/publications/2001/tech01_zdd.pdf))
- Donald Knuth, Fun With Zero-Suppressed Binary Decision Diagrams (ZDDs) (<http://myvideos.stanford.edu/player/slplayer.aspx?coll=ea60314a-53b3-4be2-8552-dcf190ca0c0b&co=af52aca1-9c60-4a7b-a10b-0e543f4f3451&o=true>) (video lecture, 2008)

# Propositional directed acyclic graph

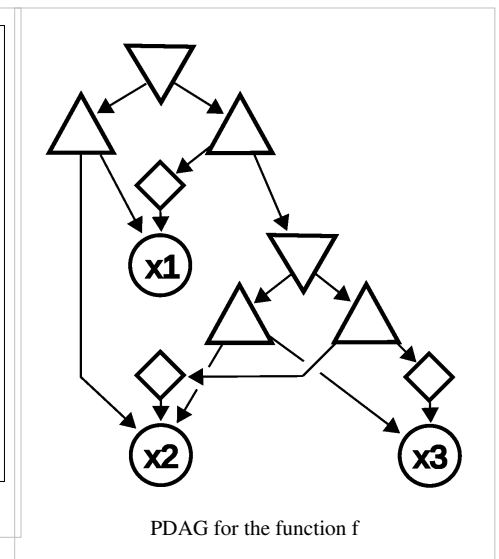
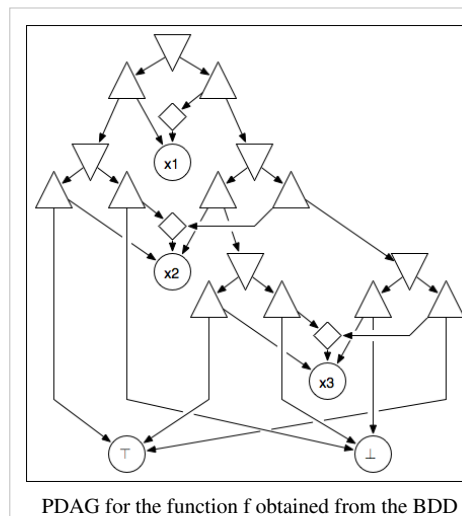
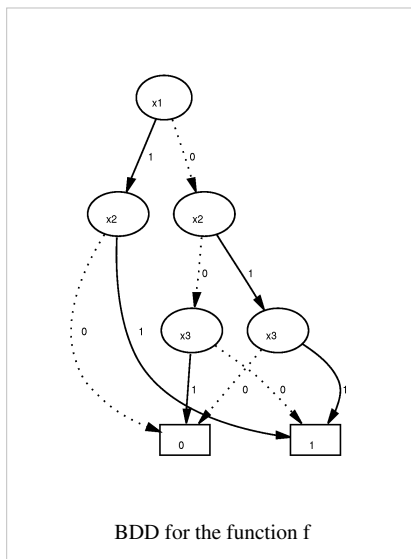
A **propositional directed acyclic graph (PDAG)** is a data structure that is used to represent a Boolean function. A Boolean function can be represented as a rooted, directed acyclic graph of the following form:

- Leaves are labeled with  $\top$  (true),  $\perp$  (false), or a Boolean variable.
- Non-leaves are  $\triangle$  (logical and),  $\nabla$  (logical or) and  $\diamond$  (logical not).
- $\triangle$  - and  $\nabla$  -nodes have at least one child.
- $\diamond$  -nodes have exactly one child.

Leaves labeled with  $\top$  ( $\perp$ ) represent the constant Boolean function which always evaluates to 1 (0). A leaf labeled with a Boolean variable  $x$  is interpreted as the assignment  $x = 1$ , i.e. it represents the Boolean function which evaluates to 1 if and only if  $x = 1$ . The Boolean function represented by a  $\triangle$  -node is the one that evaluates to 1, if and only if the Boolean function of all its children evaluate to 1. Similarly, a  $\nabla$  -node represents the Boolean function that evaluates to 1, if and only if the Boolean function of at least one child evaluates to 1. Finally, a  $\diamond$  -node represents the complementary Boolean function its child, i.e. the one that evaluates to 1, if and only if the Boolean function of its child evaluates to 0.

## PDAG, BDD, and NNF

Every **binary decision diagram (BDD)** and every **negation normal form (NNF)** are also a PDAG with some particular properties. The following pictures represent the Boolean function  $f(x_1, x_2, x_3) = -x_1 * -x_2 * -x_3 + x_1 * x_2 + x_2 * x_3$ :



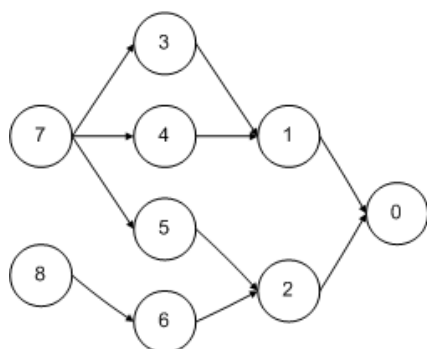
## References

- M. Wachter & R. Haenni, "Propositional DAGs: a New Graph-Based Language for Representing Boolean Functions", KR'06, 10th International Conference on Principles of Knowledge Representation and Reasoning, Lake District, UK, 2006.
- M. Wachter & R. Haenni, "Probabilistic Equivalence Checking with Propositional DAGs", Technical Report iam-2006-001, Institute of Computer Science and Applied Mathematics, University of Bern, Switzerland, 2006.
- M. Wachter, R. Haenni & J. Jonczy, "Reliability and Diagnostics of Modular Systems: a New Probabilistic Approach", DX'06, 18th International Workshop on Principles of Diagnosis, Peñaranda de Duero, Burgos, Spain, 2006.

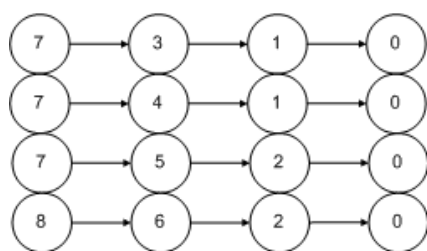
## Graph-structured stack

In computer science, a **graph-structured stack** is a directed acyclic graph where each directed path represents a stack. The graph-structured stack is an essential part of Tomita's algorithm, where it replaces the usual stack of a pushdown automaton. This allows the algorithm to encode the nondeterministic choices in parsing an ambiguous grammar, sometimes with greater efficiency.

In the following diagram, there are four stacks:  $\{7,3,1,0\}$ ,  $\{7,4,1,0\}$ ,  $\{7,5,2,0\}$ , and  $\{8,6,2,0\}$ .



Another way to simulate nondeterminism would be to duplicate the stack as needed. The duplication would be less efficient since vertices would not be shared. For this example, 16 vertices would be needed instead of 9.



# Scene graph

---

A **scene graph** is a general data structure commonly used by vector-based graphics editing applications and modern computer games. Examples of such programs include Acrobat 3D, Adobe Illustrator, AutoCAD, CorelDRAW, OpenSceneGraph, OpenSG, VRML97, and X3D.

The scene graph is a structure that arranges the logical and often (but not necessarily) spatial representation of a graphical scene. The definition of a scene graph is fuzzy because programmers who implement scene graphs in applications — and, in particular, the games industry — take the basic principles and adapt these to suit particular applications. This means there is no consensus as to what a scene graph should be.

A scene graph is a collection of nodes in a graph or tree structure. A node may have many children but often only a single parent, with the effect of a parent applied to all its child nodes; an operation performed on a group automatically propagates its effect to all of its members. In many programs, associating a geometrical transformation matrix (see also transformation and matrix) at each group level and concatenating such matrices together is an efficient and natural way to process such operations. A common feature, for instance, is the ability to group related shapes/objects into a compound object that can then be moved, transformed, selected, etc. as easily as a single object.

It also happens that in some scene graphs, a node can have a relation to any node including itself, or at least an extension that refers to another node (for instance Pixar's PhotoRealistic RenderMan because of its usage of Reyes rendering algorithm, or Adobe Systems's Acrobat 3D for advanced interactive manipulation).

## Scene graphs in graphics editing tools

In vector-based graphics editing, each leaf node in a scene graph represents some atomic unit of the document, usually a shape such as an ellipse or Bezier path. Although shapes themselves (particularly paths) can be decomposed further into nodes such as spline nodes, it is practical to think of the scene graph as composed of shapes rather than going to a lower level of representation.

Another useful and user-driven node concept is the layer. A layer acts like a transparent sheet upon which any number of shapes and shape groups can be placed. The document then becomes a set of layers, any of which can be conveniently made invisible, dimmed, or locked (made read-only). Some applications place all layers in a linear list, while others support sublayers (i.e., layers within layers to any desired depth).

Internally, there may be no real structural difference between layers and groups at all, since they are both just nodes of a scene graph. If differences are needed, a common type declaration in C++ would be to make a generic node class, and then derive layers and groups as subclasses. A visibility member, for example, would be a feature of a layer, but not necessarily of a group.

## Scene graphs in games and 3D applications

Scene graphs are useful for modern games using 3D graphics and increasingly large worlds or levels. In such applications, nodes in a scene graph (generally) represent entities or objects in the scene.

For instance, a game might define a logical relationship between a knight and a horse so that the knight is considered an extension to the horse. The scene graph would have a 'horse' node with a 'knight' node attached to it.

As well as describing the logical relationship, the scene graph may also describe the spatial relationship of the various entities: the knight moves through 3D space as the horse moves.

In these large applications, memory requirements are major considerations when designing a scene graph. For this reason, many large scene graph systems use instancing to reduce memory costs and increase speed. In our example above, each knight is a separate scene node, but the graphical representation of the knight (made up of a 3D mesh, textures, materials and shaders) is instanced. This means that only a single copy of the data is kept, which is then

---



referenced by any 'knight' nodes in the scene graph. This allows a reduced memory budget and increased speed, since when a new knight node is created, the appearance data does not need to be duplicated.

## Scene graph implementation

The simplest form of scene graph uses an array or linked list data structure, and displaying its shapes is simply a matter of linearly iterating the nodes one by one. Other common operations, such as checking to see which shape intersects the mouse pointer (e.g., in a GUI-based applications) are also done via linear searches. For small scene graphs, this tends to suffice.

Larger scene graphs cause linear operations to become noticeably slow and thus more complex underlying data structures are used, the most popular and common form being a tree. In these scene graphs, the composite design pattern is often employed to create the hierarchical representation of group nodes and leaf nodes.

**Group nodes** — Can have any number of child nodes attached to it. Group nodes include transformations and switch nodes.

**Leaf nodes** — Are nodes that are actually rendered or see the effect of an operation. These include objects, sprites, sounds, lights and anything that could be considered 'rendered' in some abstract sense.

## Scene graph operations and dispatch

Applying an operation on a scene graph requires some way of dispatching an operation based on a node's type. For example, in a render operation, a transformation group node would accumulate its transformation by matrix multiplication, vector displacement, quaternions or Euler angles. After which a leaf node sends the object off for rendering to the renderer. Some implementations might render the object directly, which invokes the underlying rendering API, such as DirectX or OpenGL. But since the underlying implementation of the rendering API usually lacks portability, one might separate the scene graph and rendering systems instead. In order to accomplish this type of dispatching, several different approaches can be taken.

In object-oriented languages such as C++, this can easily be achieved by virtual functions, where each represents an operation that can be performed on a node. Virtual functions are simple to write, but it is usually impossible to add new operations to nodes without access to the source code. Alternatively, the **visitor pattern** can be used. This has a similar disadvantage in that it is similarly difficult to add new node types.

Other techniques involve the use of RTTI (Run-Time Type Information). The operation can be realised as a class that is passed to the current node; it then queries the node's type using RTTI and looks up the correct operation in an array of callbacks or functors. This requires that the map of types to callbacks or functors be initialized at runtime, but offers more flexibility, speed and extensibility.

Variations on these techniques exist, and new methods can offer added benefits. One alternative is scene graph rebuilding, where the scene graph is rebuilt for each of the operations performed. This, however, can be very slow, but produces a highly optimised scene graph. It demonstrates that a good scene graph implementation depends heavily on the application in which it is used.

## Traversals

Traversals are the key to the power of applying operations to scene graphs. A traversal generally consists of starting at some arbitrary node (often the root of the scene graph), applying the operation(s) (often the updating and rendering operations are applied one after the other), and recursively moving down the scene graph (tree) to the child nodes, until a leaf node is reached. At this point, many scene graph engines then traverse back up the tree, applying a similar operation. For example, consider a render operation that takes transformations into account: while recursively traversing down the scene graph hierarchy, a pre-render operation is called. If the node is a transformation node, it adds its own transformation to the current transformation matrix. Once the operation finishes traversing all the

children of a node, it calls the node's post-render operation so that the transformation node can undo the transformation. This approach drastically reduces the necessary amount of matrix multiplication.

Some scene graph operations are actually more efficient when nodes are traversed in a different order — this is where some systems implement scene graph rebuilding to reorder the scene graph into an easier-to-parse format or tree.

For example, in 2D cases, scene graphs typically render themselves by starting at the tree's root node and then recursively draw the child nodes. The tree's leaves represent the most foreground objects. Since drawing proceeds from back to front with closer objects simply overwriting farther ones, the process is known as employing the Painter's algorithm. In 3D systems, which often employ depth buffers, it is more efficient to draw the closest objects first, since farther objects often need only be depth-tested instead of actually rendered, because they are occluded by nearer objects.

## Scene graphs and bounding volume hierarchies (BVHs)

Bounding Volume Hierarchies (BVHs) are useful for numerous tasks — including efficient culling and speeding up collision detection between objects. A BVH is a spatial structure, but doesn't have to partition the geometry (see spatial partitioning below).

A BVH is a tree of bounding volumes (often spheres, axis-aligned bounding boxes or oriented bounding boxes). At the bottom of the hierarchy, the size of the volume is just large enough to encompass a single object tightly (or possibly even some smaller fraction of an object in high resolution BVHs). As one ascends the hierarchy, each node has its own volume that tightly encompasses all the volumes beneath it. At the root of the tree is a volume that encompasses all the volumes in the tree (the whole scene).

BVHs are useful for speeding up collision detection between objects. If an object's bounding volume does not intersect a volume higher in the tree, it cannot intersect any object below that node (so they are all rejected very quickly).

Obviously, there are some similarities between BVHs and scene graphs. A scene graph can easily be adapted to include/become a BVH — if each node has a volume associated or there is a purpose-built 'bound node' added in at convenient location in the hierarchy. This may not be the typical view of a scene graph, but there are benefits to including a BVH in a scene graph.

## Scene graphs and spatial partitioning

An effective way of combining spatial partitioning and scene graphs is by creating a scene leaf node that contains the spatial partitioning data. This data is usually static and generally contains non-moving level data in some partitioned form. Some systems may have the systems and their rendering separately. This is fine and there are no real advantages to either method. In particular, it is bad to have the scene graph contained within the spatial partitioning system, as the scene graph is better thought of as the grander system to the spatial partitioning.

### When it is useful to combine them

In short: Spatial partitioning will/should considerably speed up the processing and rendering time of the scene graph.

Very large drawings, or scene graphs that are generated solely at runtime (as happens in ray tracing rendering programs), require defining of group nodes in a more automated fashion. A raytracer, for example, will take a scene description of a 3D model and build an internal representation that breaks up its individual parts into bounding boxes (also called bounding slabs). These boxes are grouped hierarchically so that ray intersection tests (as part of visibility determination) can be efficiently computed. A group box that does not intersect an eye ray, for example, can entirely skip testing any of its members.

A similar efficiency holds in 2D applications as well. If the user has magnified a document so that only part of it is visible on his computer screen, and then scrolls in it, it is useful to use a bounding box (or in this case, a bounding rectangle scheme) to quickly determine which scene graph elements are visible and thus actually need to be drawn.

Depending on the particulars of the application's drawing performance, a large part of the scene graph's design can be impacted by rendering efficiency considerations. In 3D video games such as Quake, for example, binary space partitioning (BSP) trees are heavily favored to minimize visibility tests. BSP trees, however, take a very long time to compute from design scene graphs, and must be recomputed if the design scene graph changes, so the levels tend to remain static, and dynamic characters aren't generally considered in the spatial partitioning scheme.

Scene graphs for dense regular objects such as heightfields and polygon meshes tend to employ quadtrees and octrees, which are specialized variants of a 3D bounding box hierarchy. Since a heightfield occupies a box volume itself, recursively subdividing this box into eight subboxes (hence the 'oct' in octree) until individual heightfield elements are reached is efficient and natural. A quadtree is simply a 2D octree.

## Standards

### PHIGS

PHIGS was the first commercial scene graph specification, and became an ANSI standard in 1988. Disparate implementations were provided by Unix hardware vendors. The HOOPS 3D Graphics System appears to have been the first commercial scene graph library provided by a single software vendor. It was designed to run on disparate lower-level 2D and 3D interfaces, with the first major production version (v3.0) completed in 1991. Shortly thereafter, Silicon Graphics released IRIS Inventor 1.0 (1992), which was a scene graph built on top of the IRIS GL 3D API. It was followed up with Open Inventor in 1994, a portable scene graph built on top of OpenGL. More 3D scene graph libraries can be found in Category:3D scenegraph APIs.

### X3D

X3D is a royalty-free open-standards file format and run-time architecture to represent and communicate 3D scenes and objects using XML. It is an ISO-ratified standard that provides a system for the storage, retrieval and playback of real-time graphics content embedded in applications, all within an open architecture to support a wide array of domains and user scenarios.

## References

### Books

- Leler, Wm and Merry, Jim (1996) *3D with HOOPS*, Addison-Wesley
  - Wernecke, Josie (1994) *The Inventor Mentor: Programming Object-Oriented 3D Graphics with Open Inventor*, Addison-Wesley, ISBN 0-201-62495-8 (Release 2)
-

## Web sites and articles

- Bar-Zeev, Avi. "Scenegraphs: Past, Present, and Future" <sup>[1]</sup>
- Carey, Rikk and Bell, Gavin (1997). "The Annotated VRML 97 Reference Manual" <sup>[2]</sup>
- Helman, Jim; Rohlf, John (1994). "IRIS Performer: A High Performance Multiprocessing Toolkit for Real-Time 3D Graphics" <sup>[3]</sup>
- Java3D <sup>[4]</sup>: Aviatrix3D <sup>[5]</sup>, LG3D <sup>[6]</sup> \* jReality <sup>[7]</sup>
- OpenSG <sup>[8]</sup>
- OpenSceneGraph <sup>[9]</sup>
- PEXTimes <sup>[10]</sup>
- Strauss, Paul (1993). "IRIS Inventor, a 3D Graphics Toolkit" <sup>[11]</sup>
- Visualization Library <sup>[12]</sup>

## References

- [1] <http://www.realityprime.com/scenegraph.php>
  - [2] <http://www.jwave.vt.edu/~engineer/vrml97book/ch1.htm>
  - [3] <http://portal.acm.org/citation.cfm?id=192262>
  - [4] <http://java3d.dev.java.net>
  - [5] <http://aviatrix3d.j3d.org>
  - [6] <https://lg3d.dev.java.net>
  - [7] <http://www.jreality.de>
  - [8] <http://www.opensg.org>
  - [9] <http://www.openscenegraph.org>
  - [10] <http://www.jch.com/jch/vrml/PEXTimes.txt>
  - [11] <http://portal.acm.org/citation.cfm?id=165889>
  - [12] <http://www.visualizationlibrary.com>
-

# Appendix

## Big O notation

In mathematics, **big O notation** is used to describe the limiting behavior of a function when the argument tends towards a particular value or infinity, usually in terms of simpler functions. It is a member of a larger family of notations that is called **Landau notation**, **Bachmann–Landau notation** (after Edmund Landau and Paul Bachmann), or **asymptotic notation**. In computer science, big O notation is used to classify algorithms by how they respond (*e.g.*, in their processing time or working space requirements) to changes in input size.

Big O notation characterizes functions according to their growth rates: different functions with the same growth rate may be represented using the same O notation. A description of a function in terms of big O notation

usually only provides an upper bound on the growth rate of the function. Associated with big O notation are several related notations, using the symbols  $o$ ,  $\Omega$ ,  $\omega$ , and  $\Theta$ , to describe other kinds of bounds on asymptotic growth rates.

Big O notation is also used in many other fields to provide similar estimates.

### Formal definition

Let  $f(x)$  and  $g(x)$  be two functions defined on some subset of the real numbers. One writes

$$f(x) = O(g(x)) \text{ as } x \rightarrow \infty$$

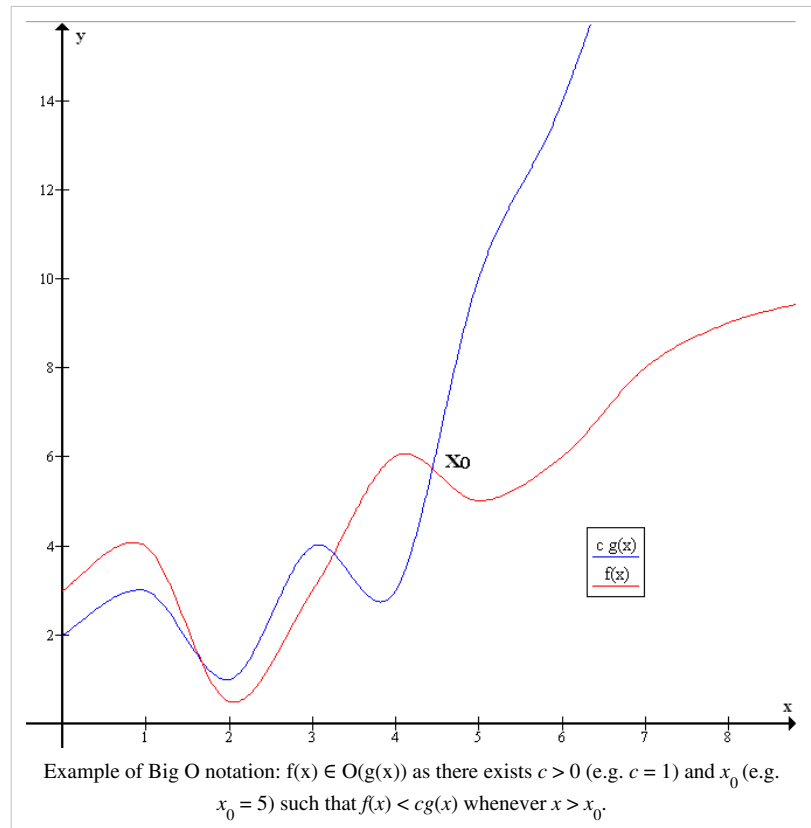
if and only if there is a positive constant  $M$  such that for all sufficiently large values of  $x$ ,  $f(x)$  is at most  $M$  multiplied by  $g(x)$  in absolute value. That is,  $f(x) = O(g(x))$  if and only if there exists a positive real number  $M$  and a real number  $x_0$  such that

$$|f(x)| \leq M|g(x)| \text{ for all } x > x_0.$$

In many contexts, the assumption that we are interested in the growth rate as the variable  $x$  goes to infinity is left unstated, and one writes more simply that  $f(x) = O(g(x))$ . The notation can also be used to describe the behavior of  $f$  near some real number  $a$  (often,  $a = 0$ ): we say

$$f(x) = O(g(x)) \text{ as } x \rightarrow a$$

if and only if there exist positive numbers  $\delta$  and  $M$  such that



$$|f(x)| \leq M|g(x)| \text{ for } |x - a| < \delta.$$

If  $g(x)$  is non-zero for values of  $x$  sufficiently close to  $a$ , both of these definitions can be unified using the limit superior:

$$f(x) = O(g(x)) \text{ as } x \rightarrow a$$

if and only if

$$\limsup_{x \rightarrow a} \left| \frac{f(x)}{g(x)} \right| < \infty.$$

## Example

In typical usage, the formal definition of  $O$  notation is not used directly; rather, the  $O$  notation for a function  $f(x)$  is derived by the following simplification rules:

- If  $f(x)$  is a sum of several terms, the one with the largest growth rate is kept, and all others omitted.
- If  $f(x)$  is a product of several factors, any constants (terms in the product that do not depend on  $x$ ) are omitted.

For example, let  $f(x) = 6x^4 - 2x^3 + 5$ , and suppose we wish to simplify this function, using  $O$  notation, to describe its growth rate as  $x$  approaches infinity. This function is the sum of three terms:  $6x^4$ ,  $-2x^3$ , and 5. Of these three terms, the one with the highest growth rate is the one with the largest exponent as a function of  $x$ , namely  $6x^4$ . Now one may apply the second rule:  $6x^4$  is a product of 6 and  $x^4$  in which the first factor does not depend on  $x$ . Omitting this factor results in the simplified form  $x^4$ . Thus, we say that  $f(x)$  is a big-oh of  $(x^4)$  or mathematically we can write  $f(x) = O(x^4)$ . One may confirm this calculation using the formal definition: let  $f(x) = 6x^4 - 2x^3 + 5$  and  $g(x) = x^4$ . Applying the formal definition from above, the statement that  $f(x) = O(x^4)$  is equivalent to its expansion,

$$|f(x)| \leq M|g(x)|$$

for some suitable choice of  $x_0$  and  $M$  and for all  $x > x_0$ . To prove this, let  $x_0 = 1$  and  $M = 13$ . Then, for all  $x > x_0$ :

$$\begin{aligned} |6x^4 - 2x^3 + 5| &\leq 6x^4 + |2x^3| + 5 \\ &\leq 6x^4 + 2x^4 + 5x^4 \\ &\leq 13x^4 \\ &\leq 13|x^4| \end{aligned}$$

so

$$|6x^4 - 2x^3 + 5| \leq 13|x^4|.$$

## Usage

Big O notation has two main areas of application. In mathematics, it is commonly used to describe how closely a finite series approximates a given function, especially in the case of a truncated Taylor series or asymptotic expansion. In computer science, it is useful in the analysis of algorithms. In both applications, the function  $g(x)$  appearing within the  $O(\dots)$  is typically chosen to be as simple as possible, omitting constant factors and lower order terms. There are two formally close, but noticeably different, usages of this notation: infinite asymptotics and infinitesimal asymptotics. This distinction is only in application and not in principle, however—the formal definition for the "big O" is the same for both cases, only with different limits for the function argument.

## Infinite asymptotics

Big O notation is useful when analyzing algorithms for efficiency. For example, the time (or the number of steps) it takes to complete a problem of size  $n$  might be found to be  $T(n) = 4n^2 - 2n + 2$ . As  $n$  grows large, the  $n^2$  term will come to dominate, so that all other terms can be neglected — for instance when  $n = 500$ , the term  $4n^2$  is 1000 times as large as the  $2n$  term. Ignoring the latter would have negligible effect on the expression's value for most purposes. Further, the coefficients become irrelevant if we compare to any other order of expression, such as an expression containing a term  $n^3$  or  $n^4$ . Even if  $T(n) = 1,000,000n^2$ , if  $U(n) = n^3$ , the latter will always exceed the former once  $n$  grows larger than 1,000,000 ( $T(1,000,000) = 1,000,000^3 = U(1,000,000)$ ). Additionally, the number of steps depends on the details of the machine model on which the algorithm runs, but different types of machines typically vary by only a constant factor in the number of steps needed to execute an algorithm. So the big O notation captures what remains: we write either

$$T(n) = O(n^2)$$

or

$$T(n) \in O(n^2)$$

and say that the algorithm has *order of  $n^2$*  time complexity. Note that "=" is not meant to express "is equal to" in its normal mathematical sense, but rather a more colloquial "is", so the second expression is technically accurate (see the "Equals sign" discussion below) while the first is a common abuse of notation.<sup>[1]</sup>

## Infinitesimal asymptotics

Big O can also be used to describe the error term in an approximation to a mathematical function. The most significant terms are written explicitly, and then the least-significant terms are summarized in a single big O term. For example,

$$e^x = 1 + x + \frac{x^2}{2} + O(x^3) \quad \text{as } x \rightarrow 0$$

expresses the fact that the error, the difference  $e^x - (1 + x + x^2/2)$ , is smaller in absolute value than some constant times  $|x^3|$  when  $x$  is close enough to 0.

## Properties

If a function  $f(n)$  can be written as a finite sum of other functions, then the fastest growing one determines the order of  $f(n)$ . For example

$$f(n) = 9 \log n + 5(\log n)^3 + 3n^2 + 2n^3 = O(n^3).$$

In particular, if a function may be bounded by a polynomial in  $n$ , then as  $n$  tends to *infinity*, one may disregard *lower-order* terms of the polynomial.  $O(n^c)$  and  $O(c^n)$  are very different. If  $c$  is greater than one, then the latter grows much faster. A function that grows faster than  $n^c$  for any  $c$  is called *superpolynomial*. One that grows more slowly than any exponential function of the form  $c^n$  is called *subexponential*. An algorithm can require time that is both superpolynomial and subexponential; examples of this include the fastest known algorithms for integer factorization.  $O(\log n)$  is exactly the same as  $O(\log(n^c))$ . The logarithms differ only by a constant factor (since  $\log(n^c) = c \log n$ ) and thus the big O notation ignores that. Similarly, logs with different constant bases are equivalent. Exponentials with different bases, on the other hand, are not of the same order. For example,  $2^n$  and  $3^n$  are not of the same order. Changing units may or may not affect the order of the resulting algorithm. Changing units is equivalent to multiplying the appropriate variable by a constant wherever it appears. For example, if an algorithm runs in the order of  $n^2$ , replacing  $n$  by  $cn$  means the algorithm runs in the order of  $c^2 n^2$ , and the big O notation ignores the constant  $c^2$ . This can be written as  $c^2 n^2 \in O(n^2)$ . If, however, an algorithm runs in the order of  $2^n$ , replacing  $n$  with  $cn$  gives  $2^{cn} = (2^c)^n$ . This is not equivalent to  $2^n$  in general. Changing of variable may affect the order of the resulting algorithm. For example, if an algorithm's running time is  $O(n)$  when measured in terms of

the number  $n$  of *digits* of an input number  $x$ , then its running time is  $O(\log x)$  when measured as a function of the input number  $x$  itself, because  $n = \Theta(\log x)$ .

### Product

$$f_1 \in O(g_1) \text{ and } f_2 \in O(g_2) \Rightarrow f_1 f_2 \in O(g_1 g_2)$$

$$f \cdot O(g) \subset O(fg)$$

### Sum

$$f_1 \in O(g_1) \text{ and } f_2 \in O(g_2) \Rightarrow f_1 + f_2 \in O(|g_1| + |g_2|)$$

This implies  $f_1 \in O(g)$  and  $f_2 \in O(g) \Rightarrow f_1 + f_2 \in O(g)$ , which means that  $O(g)$  is a convex cone.

If  $f$  and  $g$  are positive functions,  $f + O(g) \in O(f + g)$

### Multiplication by a constant

Let  $k$  be a constant. Then:

$O(kg) = O(g)$  if  $k$  is nonzero.

$$f \in O(g) \Rightarrow kf \in O(g).$$

### Multiple variables

Big  $O$  (and little  $o$ , and  $\Omega$ ...) can also be used with multiple variables. To define Big  $O$  formally for multiple variables, suppose  $f(\vec{x})$  and  $g(\vec{x})$  are two functions defined on some subset of  $\mathbb{R}^n$ . We say

$$f(\vec{x}) \text{ is } O(g(\vec{x})) \text{ as } \vec{x} \rightarrow \infty$$

if and only if

$$\exists C \exists M > 0 \text{ such that } |f(\vec{x})| \leq C|g(\vec{x})| \text{ for all } \vec{x} \text{ with } x_i > M \text{ for all } i.$$

For example, the statement

$$f(n, m) = n^2 + m^3 + O(n + m) \text{ as } n, m \rightarrow \infty$$

asserts that there exist constants  $C$  and  $M$  such that

$$\forall n, m > M: |g(n, m)| \leq C(n + m),$$

where  $g(n, m)$  is defined by

$$f(n, m) = n^2 + m^3 + g(n, m).$$

Note that this definition allows all of the coordinates of  $\vec{x}$  to increase to infinity. In particular, the statement

$$f(n, m) = O(n^m) \text{ as } n, m \rightarrow \infty$$

(i.e.,  $\exists C \exists M \forall n \forall m \dots$ ) is quite different from

$$\forall m: f(n, m) = O(n^m) \text{ as } n \rightarrow \infty$$

(i.e.,  $\forall m \exists C \exists M \forall n \dots$ ).



## Matters of notation

### Equals sign

The statement " $f(x)$  is  $O(g(x))$ " as defined above is usually written as  $f(x) = O(g(x))$ . Some consider this to be an abuse of notation, since the use of the equals sign could be misleading as it suggests a symmetry that this statement does not have. As de Bruijn says,  $O(x) = O(x^2)$  is true but  $O(x^2) = O(x)$  is not.<sup>[2]</sup> Knuth describes such statements as "one-way equalities", since if the sides could be reversed, "we could deduce ridiculous things like  $n = n^2$  from the identities  $n = O(n^2)$  and  $n^2 = O(n^2)$ ."<sup>[3]</sup> For these reasons, it would be more precise to use set notation and write  $f(x) \in O(g(x))$ , thinking of  $O(g(x))$  as the class of all functions  $h(x)$  such that  $|h(x)| \leq C|g(x)|$  for some constant  $C$ .<sup>[3]</sup> However, the use of the equals sign is customary. Knuth pointed out that "mathematicians customarily use the  $=$  sign as they use the word 'is' in English: Aristotle is a man, but a man isn't necessarily Aristotle."<sup>[4]</sup>

### Other arithmetic operators

Big O notation can also be used in conjunction with other arithmetic operators in more complicated equations. For example,  $h(x) + O(f(x))$  denotes the collection of functions having the growth of  $h(x)$  plus a part whose growth is limited to that of  $f(x)$ . Thus,

$$g(x) = h(x) + O(f(x))$$

expresses the same as

$$g(x) - h(x) \in O(f(x)).$$

### Example

Suppose an algorithm is being developed to operate on a set of  $n$  elements. Its developers are interested in finding a function  $T(n)$  that will express how long the algorithm will take to run (in some arbitrary measurement of time) in terms of the number of elements in the input set. The algorithm works by first calling a subroutine to sort the elements in the set and then perform its own operations. The sort has a known time complexity of  $O(n^2)$ , and after the subroutine runs the algorithm must take an additional  $55n^3 + 2n + 10$  time before it terminates. Thus the overall time complexity of the algorithm can be expressed as

$$T(n) = O(n^2) + 55n^3 + 2n + 10.$$

This can perhaps be most easily read by replacing  $O(n^2)$  with "some function that grows asymptotically no faster than  $n^2$ ". Again, this usage disregards some of the formal meaning of the "=" and "+" symbols, but it does allow one to use the big O notation as a kind of convenient placeholder.

### Declaration of variables

Another feature of the notation, although less exceptional, is that function arguments may need to be inferred from the context when several variables are involved. The following two right-hand side big O notations have dramatically different meanings:

$$f(m) = O(m^n),$$

$$g(n) = O(m^n).$$

The first case states that  $f(m)$  exhibits polynomial growth, while the second, assuming  $m > 1$ , states that  $g(n)$  exhibits exponential growth. To avoid confusion, some authors use the notation

$$g(x) \in O(f(x)).$$

rather than the less explicit

$$g \in O(f),$$

## Multiple usages

In more complicated usage,  $O(\dots)$  can appear in different places in an equation, even several times on each side. For example, the following are true for  $n \rightarrow \infty$

$$\begin{aligned}(n+1)^2 &= n^2 + O(n) \\ (n + O(n^{1/2}))(n + O(\log n))^2 &= n^3 + O(n^{5/2}) \\ n^{O(1)} &= O(e^n).\end{aligned}$$

The meaning of such statements is as follows: for *any* functions which satisfy each  $O(\dots)$  on the left side, there are *some* functions satisfying each  $O(\dots)$  on the right side, such that substituting all these functions into the equation makes the two sides equal. For example, the third equation above means: "For any function  $f(n) = O(1)$ , there is some function  $g(n) = O(e^n)$  such that  $n^{f(n)} = g(n)$ ." In terms of the "set notation" above, the meaning is that the class of functions represented by the left side is a subset of the class of functions represented by the right side. In this use the "=" is a formal symbol that unlike the usual use of "=" is not a symmetric relation. Thus for example  $n^{O(1)} = O(e^n)$  does not imply the false statement  $O(e^n) = n^{O(1)}$ .

## Orders of common functions

Further information: Time complexity#Table of common time complexities

Here is a list of classes of functions that are commonly encountered when analyzing the running time of an algorithm. In each case,  $c$  is a constant and  $n$  increases without bound. The slower-growing functions are generally listed first.

Notation	Name	Example
$O(1)$	constant	Determining if a number is even or odd; using a constant-size lookup table
$O(\log \log n)$	double logarithmic	Finding an item using interpolation search in a sorted array of uniformly distributed values.
$O(\log n)$	logarithmic	Finding an item in a sorted array with a binary search or a balanced search tree as well as all operations in a Binomial heap.
$O(n^c)$ , $0 < c < 1$	fractional power	Searching in a kd-tree
$O(n)$	linear	Finding an item in an unsorted list or a malformed tree (worst case) or in an unsorted array; Adding two $n$ -bit integers by ripple carry.
$O(n \log n) = O(\log n!)$	linearithmic, loglinear, or quasilinear	Performing a Fast Fourier transform; heapsort, quicksort (best and average case), or merge sort
$O(n^2)$	quadratic	Multiplying two $n$ -digit numbers by a simple algorithm; bubble sort (worst case or naive implementation), Shell sort, quicksort (worst case), selection sort or insertion sort
$O(n^c)$ , $c > 1$	polynomial or algebraic	Tree-adjoining grammar parsing; maximum matching for bipartite graphs
$L_n[\alpha, c]$ , $0 < \alpha < 1 = e^{(c+o(1))(\ln n)^\alpha (\ln \ln n)^{1-\alpha}}$	L-notation or sub-exponential	Factoring a number using the quadratic sieve or number field sieve
$O(c^n)$ , $c > 1$	exponential	Finding the (exact) solution to the travelling salesman problem using dynamic programming; determining if two logical statements are equivalent using brute-force search
$O(n!)$	factorial	Solving the traveling salesman problem via brute-force search; generating all unrestricted permutations of a poset; finding the determinant with expansion by minors.

The statement  $f(n) = O(n!)$  is sometimes weakened to  $f(n) = O(n^n)$  to derive simpler formulas for asymptotic complexity. For any  $k > 0$  and  $c > 0$ ,  $O(n^c(\log n)^k)$  is a subset of  $O(n^{c+\varepsilon})$  for any  $\varepsilon > 0$ , so may be considered as a polynomial with some bigger order.

## Related asymptotic notations

Big  $O$  is the most commonly used asymptotic notation for comparing functions, although in many cases Big  $O$  may be replaced with Big Theta  $\Theta$  for asymptotically tighter bounds. Here, we define some related notations in terms of Big  $O$ , progressing up to the family of Bachmann–Landau notations to which Big  $O$  notation belongs.

### Little-o notation

The relation  $f(x) \in o(g(x))$  is read as " $f(x)$  is little-o of  $g(x)$ ". Intuitively, it means that  $g(x)$  grows much faster than  $f(x)$ , or similarly, the growth of  $f(x)$  is nothing compared to that of  $g(x)$ . It assumes that  $f$  and  $g$  are both functions of one variable. Formally,  $f(n) = o(g(n))$  as  $n \rightarrow \infty$  means that for every positive constant  $\varepsilon$  there exists a constant  $N$  such that

$$|f(n)| \leq \varepsilon |g(n)| \quad \text{for all } n \geq N. \quad [3]$$

Note the difference between the earlier formal definition for the big- $O$  notation, and the present definition of little- $o$ : while the former has to be true for *at least one* constant  $M$  the latter must hold for *every* positive constant  $\varepsilon$ , however small.<sup>[1]</sup> In this way little- $o$  notation makes a stronger statement than the corresponding big- $O$  notation: every function that is little- $o$  of  $g$  is also big- $O$  of  $g$ , but not every function that is big- $O$  of  $g$  is also little- $o$  of  $g$  (for instance  $g$  itself is not, unless it is identically zero near  $\infty$ ).

If  $g(x)$  is nonzero, or at least becomes nonzero beyond a certain point, the relation  $f(x) = o(g(x))$  is equivalent to

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 0.$$

For example,

- $2x \in o(x^2)$
- $2x^2 \notin o(x^2)$
- $1/x \in o(1)$

Little- $o$  notation is common in mathematics but rarer in computer science. In computer science the variable (and function value) is most often a natural number. In mathematics, the variable and function values are often real numbers. The following properties can be useful:

- $o(f) + o(f) \subseteq o(f)$
- $o(f)o(g) \subseteq o(fg)$
- $o(o(f)) \subseteq o(f)$
- $o(f) \subset O(f)$  (and thus the above properties apply with most combinations of  $o$  and  $O$ ).

As with big  $O$  notation, the statement " $f(x)$  is  $o(g(x))$ " is usually written as  $f(x) = o(g(x))$ , which is a slight abuse of notation.

## Family of Bachmann–Landau notations

Notation	Name	Intuition	Informal definition: for sufficiently large $n$ ...	Formal Definition	Notes
$f(n) \in O(g(n))$	Big Omicron; Big O; Big Oh	$f$ is bounded above by $g$ (up to constant factor) asymptotically	$ f(n)  \leq g(n) \cdot k$ for some $k$	$\exists k > 0 \exists n_0 \forall n > n_0  f(n)  \leq  g(n) \cdot k $ or $\exists k > 0 \exists n_0 \forall n > n_0 f(n) \leq g(n) \cdot k$	
$f(n) \in \Omega(g(n))$	Big Omega	$f$ is bounded below by $g$ (up to constant factor) asymptotically	$f(n) \geq g(n) \cdot k$ for some positive $k$	$\exists k > 0 \exists n_0 \forall n > n_0 g(n) \cdot k \leq f(n)$	Since the beginning of the 20th century, papers in number theory have been increasingly and widely using this notation in the weaker sense that $f = o(g)$ is false.
$f(n) \in \Theta(g(n))$	Big Theta	$f$ is bounded both above and below by $g$ asymptotically	$g(n) \cdot k_1 \leq f(n) \leq g(n) \cdot k_2$ for some positive $k_1, k_2$	$\exists k_1 > 0 \exists k_2 > 0 \exists n_0 \forall n > n_0 g(n) \cdot k_1 \leq f(n) \leq g(n) \cdot k_2$	
$f(n) \in o(g(n))$	Small Omicron; Small O; Small Oh	$f$ is dominated by $g$ asymptotically	$ f(n)  \leq  g(n)  \cdot \varepsilon$ for every $\varepsilon$	$\forall \varepsilon > 0 \exists n_0 \forall n > n_0  f(n)  \leq  g(n) \cdot \varepsilon $	
$f(n) \in \omega(g(n))$	Small Omega	$f$ dominates $g$ asymptotically	$f(n) \geq g(n) \cdot k$ for every $k$	$\forall k > 0 \exists n_0 \forall n > n_0 g(n) \cdot k < f(n)$	
$f(n) \sim g(n)$	On the order of	$f$ is equal to $g$ asymptotically	$f(n)/g(n) \rightarrow 1$	$\forall \varepsilon > 0 \exists n_0 \forall n > n_0 \left  \frac{f(n)}{g(n)} - 1 \right  < \varepsilon$	

Bachmann–Landau notation was designed around several mnemonics, as shown in the As  $n \rightarrow \infty$ , eventually... column above and in the bullets below. To conceptually access these mnemonics, "omicron" can be read "o-micron" and "omega" can be read "o-mega". Also, the lower-case versus capitalization of the Greek letters in Bachmann–Landau notation is mnemonic.

- The *o-micron* mnemonic: The o-micron reading of  $f(n) \in O(g(n))$  and of  $f(n) \in o(g(n))$  can be thought of as "O-smaller than" and "o-smaller than", respectively. This *micro*/smaller mnemonic refers to: for sufficiently large input parameter(s),  $f$  grows at a rate that may henceforth be **less** than  $cg$  regarding  $g \in O(f)$  or  $g \in o(f)$ .
- The *o-mega* mnemonic: The o-mega reading of  $f(n) \in \Omega(g(n))$  and of  $f(n) \in \omega(g(n))$  can be thought of as "O-larger than". This *mega*/larger mnemonic refers to: for sufficiently large input parameter(s),  $f$  grows at a rate that may henceforth be **greater** than  $cg$  regarding  $g \in \Omega(f)$  or  $g \in \omega(f)$ .

- The **upper-case mnemonic**: This mnemonic reminds us when to use the upper-case Greek letters in  $f(n) \in O(g(n))$  and  $f(n) \in \Omega(g(n))$ : for sufficiently large input parameter(s),  $f$  grows at a rate that may henceforth be **equal** to  $\mathcal{Cg}$  regarding  $g \in O(f)$ .
- The **lower-case mnemonic**: This mnemonic reminds us when to use the lower-case Greek letters in  $f(n) \in o(g(n))$  and  $f(n) \in \omega(g(n))$ : for sufficiently large input parameter(s),  $f$  grows at a rate that is henceforth **inequal** to  $\mathcal{Cg}$  regarding  $g \in O(f)$ .

Aside from Big O notation, the Big Theta  $\Theta$  and Big Omega  $\Omega$  notations are the two most often used in computer science; the Small Omega  $\omega$  notation is rarely used in computer science.

## Use in computer science

Informally, especially in computer science, the Big O notation often is permitted to be somewhat abused to describe an asymptotic tight bound where using Big Theta  $\Theta$  notation might be more factually appropriate in a given context. For example, when considering a function  $T(n) = 73n^3 + 22n^2 + 58$ , all of the following are generally acceptable, but tightness of bound (i.e., bullets 2 and 3 below) are usually strongly preferred over laxness of bound (i.e., number 1 below).

1.  $T(n) = O(n^{100})$ , which is identical to  $T(n) \in O(n^{100})$
2.  $T(n) = O(n^3)$ , which is identical to  $T(n) \in O(n^3)$
3.  $T(n) = \Theta(n^3)$ , which is identical to  $T(n) \in \Theta(n^3)$ .

The equivalent English statements are respectively:

1.  $T(n)$  grows asymptotically no faster than  $n^{100}$
2.  $T(n)$  grows asymptotically no faster than  $n^3$
3.  $T(n)$  grows asymptotically as fast as  $n^3$ .

So while all three statements are true, progressively more information is contained in each. In some fields, however, the Big O notation (number 2 in the lists above) would be used more commonly than the Big Theta notation (bullets number 3 in the lists above) because functions that grow more slowly are more desirable. For example, if  $T(n)$  represents the running time of a newly developed algorithm for input size  $n$ , the inventors and users of the algorithm might be more inclined to put an upper asymptotic bound on how long it will take to run without making an explicit statement about the lower asymptotic bound.

## Extensions to the Bachmann–Landau notations

Another notation sometimes used in computer science is  $\tilde{O}$  (read *soft-O*):  $f(n) = \tilde{O}(g(n))$  is shorthand for  $f(n) = O(g(n) \log^k g(n))$  for some  $k$ . Essentially, it is Big O notation, ignoring logarithmic factors because the growth-rate effects of some other super-logarithmic function indicate a growth-rate explosion for large-sized input parameters that is more important to predicting bad run-time performance than the finer-point effects contributed by the logarithmic-growth factor(s). This notation is often used to obviate the "nitpicking" within growth-rates that are stated as too tightly bounded for the matters at hand (since  $\log^k n$  is always  $o(n^\varepsilon)$  for any constant  $k$  and any  $\varepsilon > 0$ ). The L notation, defined as

$$L_n[\alpha, c] = O\left(e^{(c+o(1))(\ln n)^\alpha (\ln \ln n)^{1-\alpha}}\right),$$

is convenient for functions that are between polynomial and exponential.

## Generalizations and related usages

The generalization to functions taking values in any normed vector space is straightforward (replacing absolute values by norms), where  $f$  and  $g$  need not take their values in the same space. A generalization to functions  $g$  taking values in any topological group is also possible. The "limiting process"  $x \rightarrow x_o$  can also be generalized by introducing an arbitrary filter base, i.e. to directed nets  $f$  and  $g$ . The  $o$  notation can be used to define derivatives and differentiability in quite general spaces, and also (asymptotical) equivalence of functions,

$$f \sim g \iff (f - g) \in o(g)$$

which is an equivalence relation and a more restrictive notion than the relationship " $f$  is  $\Theta(g)$ " from above. (It reduces to  $\lim f/g = 1$  if  $f$  and  $g$  are positive real valued functions.) For example,  $2x$  is  $\Theta(x)$ , but  $2x - x$  is not  $o(x)$ .

## Graph theory

It is often useful to bound the running time of graph algorithms. Unlike most other computational problems, for a graph  $G = (V, E)$  there are two relevant parameters describing the size of the input: the number  $|V|$  of vertices in the graph and the number  $|E|$  of edges in the graph. Inside asymptotic notation (and only there), it is common to use the symbols  $V$  and  $E$ , when someone really means  $|V|$  and  $|E|$ . This convention simplifies asymptotic functions and make them easily readable. The symbols  $V$  and  $E$  are never used inside asymptotic notation with their literal meaning, since the number of vertices and edges must be non-negative, so this abuse of notation does not risk ambiguity. For example  $O(E + V \log V)$  means  $O((V, E) \mapsto |E| + |V| \cdot \log |V|)$  for a suitable metric of graphs. Another common convention—referring to the values  $|V|$  and  $|E|$  by the names  $n$  and  $m$ , respectively—sidesteps this ambiguity.

## History (Bachmann-Landau, Hardy, and Vinogradov notations)

The notation was first introduced by number theorist Paul Bachmann in 1894, in the second volume of his book *Analytische Zahlentheorie* ("analytic number theory"), the first volume of which (not yet containing big O notation) was published in 1892.<sup>[5]</sup> The notation was popularized in the work of number theorist Edmund Landau; hence it is sometimes called a Landau symbol. It was popularized in computer science by Donald Knuth, who re-introduced the related Omega and Theta notations.<sup>[6]</sup> Knuth also noted that the Omega notation had been introduced by Hardy and Littlewood<sup>[7]</sup> under a different meaning " $\neq o$ " (i.e. "is not an  $o$  of"), and proposed the above definition. Hardy and Littlewood's original definition (which was also used in one paper by Landau<sup>[8]</sup>) is still used in number theory.

Hardy's symbols were (in terms of the modern  $O$  notation)

$$f \preceq g \iff f \in O(g) \quad \text{and} \quad f \prec g \iff f \in o(g);$$

(Hardy however never defined or used the notation  $\prec\prec$ , nor  $\ll$ , as it has been sometimes reported). It should also be noted that Hardy introduces the symbols  $\preceq$  and  $\prec$  (as well as some other symbols) in his 1910 tract "Orders of Infinity", and makes use of it only in three papers (1910-1913). In the remaining papers (nearly 400!) and books he constantly uses the Landau symbols  $O$  and  $o$ .

Hardy's notation is not used anymore. On the other hand, in 1947, the Russian number theorist Ivan Matveyevich Vinogradov introduced his notation  $\ll$ , which has been increasingly used in number theory instead of the  $O$  notation. We have

$$f \ll g \iff f \in O(g),$$

and frequently both notations are used in the same paper.

The big-O, standing for "order of", was originally a capital omicron; today the identical-looking Latin capital letter  $O$  is used, but never the digit zero.

## References

- [1] Thomas H. Cormen et al., 2001, *Introduction to Algorithms*, Second Edition (<http://highered.mcgraw-hill.com/sites/0070131511/>)
- [2] N. G. de Bruijn (1958). *Asymptotic Methods in Analysis* ([http://books.google.com/?id=\\_tnwmvHmVwMC&pg=PA5&vq=The+trouble+is](http://books.google.com/?id=_tnwmvHmVwMC&pg=PA5&vq=The+trouble+is)). Amsterdam: North-Holland. pp. 5–7. ISBN 978-0-486-64221-5. .
- [3] Ronald Graham, Donald Knuth, and Oren Patashnik (1994). 0-201-55802-5 *Concrete Mathematics* (<http://books.google.com/?id=pntQAAAAAAAJ&dq=editions:ISBN>) (2 ed.). Reading, Massachusetts: Addison–Wesley. p. 446. ISBN 978-0-201-55802-9. 0-201-55802-5.
- [4] Donald Knuth (June/July 1998). "Teach Calculus with Big O" (<http://www.ams.org/notices/199806/commentary.pdf>). *Notices of the American Mathematical Society* **45** (6): 687. . ( Unabridged version (<http://www-cs-staff.stanford.edu/~knuth/ocalc.tex>))
- [5] Nicholas J. Higham, *Handbook of writing for the mathematical sciences*, SIAM. ISBN 0-89871-420-6, p. 25
- [6] Donald Knuth. *Big Omicron and big Omega and big Theta* (<http://doi.acm.org/10.1145/1008328.1008329>), ACM SIGACT News, Volume 8, Issue 2, 1976.
- [7] G. H. Hardy and J. E. Littlewood, *Some problems of Diophantine approximation*, Acta Mathematica 37 (1914), p. 225
- [8] E. Landau, Nachr. Gesell. Wiss. Gött. Math-phys. Kl. 1924, 137–150.

## Further reading

- Paul Bachmann. *Die Analytische Zahlentheorie. Zahlentheorie*. pt. 2 Leipzig: B. G. Teubner, 1894.
- Edmund Landau. *Handbuch der Lehre von der Verteilung der Primzahlen*. 2 vols. Leipzig: B. G. Teubner, 1909.
- G. H. Hardy. *Orders of Infinity: The 'Infinitärcalcul' of Paul du Bois-Reymond*, 1910.
- Donald Knuth. *The Art of Computer Programming*, Volume 1: *Fundamental Algorithms*, Third Edition. Addison–Wesley, 1997. ISBN 0-201-89683-4. Section 1.2.11: Asymptotic Representations, pp. 107–123.
- Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw–Hill, 2001. ISBN 0-262-03293-7. Section 3.1: Asymptotic notation, pp. 41–50.
- Michael Sipser (1997). *Introduction to the Theory of Computation*. PWS Publishing. ISBN 0-534-94728-X. Pages 226–228 of section 7.1: Measuring complexity.
- Jeremy Avigad, Kevin Donnelly. *Formalizing O notation in Isabelle/HOL* (<http://www.andrew.cmu.edu/~avigad/Papers/bigo.pdf>)
- Paul E. Black, "big-O notation" (<http://www.nist.gov/dads/HTML/bigOnotation.html>), in *Dictionary of Algorithms and Data Structures* [online], Paul E. Black, ed., U.S. National Institute of Standards and Technology. 11 March 2005. Retrieved December 16, 2006.
- Paul E. Black, "little-o notation" (<http://www.nist.gov/dads/HTML/littleOnotation.html>), in *Dictionary of Algorithms and Data Structures* [online], Paul E. Black, ed., U.S. National Institute of Standards and Technology. 17 December 2004. Retrieved December 16, 2006.
- Paul E. Black, " $\Omega$ " (<http://www.nist.gov/dads/HTML/omegaCapital.html>), in *Dictionary of Algorithms and Data Structures* [online], Paul E. Black, ed., U.S. National Institute of Standards and Technology. 17 December 2004. Retrieved December 16, 2006.
- Paul E. Black, " $\omega$ " (<http://www.nist.gov/dads/HTML/omega.html>), in *Dictionary of Algorithms and Data Structures* [online], Paul E. Black, ed., U.S. National Institute of Standards and Technology. 29 November 2004. Retrieved December 16, 2006.
- Paul E. Black, " $\Theta$ " (<http://www.nist.gov/dads/HTML/theta.html>), in *Dictionary of Algorithms and Data Structures* [online], Paul E. Black, ed., U.S. National Institute of Standards and Technology. 17 December 2004. Retrieved December 16, 2006.

## External links

- Introduction to Asymptotic Notations (<http://www.soe.ucsc.edu/classes/cmps102/Spring04/TantaloAsymp.pdf>)
- Landau Symbols (<http://mathworld.wolfram.com/LandauSymbols.html>)
- O-Notation Visualizer: Interactive Graphs of Common O-Notations (<https://students.ics.uci.edu/~zmohiudd/ONotationVisualizer.html>)

# Amortized analysis

---

In computer science, **amortized analysis** is a method of analyzing algorithms that considers the entire sequence of operations of the program. It allows for the establishment of a worst-case bound for the performance of an algorithm irrespective of the inputs by looking at all of the operations. At the heart of the method is the idea that while certain operations may be extremely costly in resources, they cannot occur at a high-enough frequency to weigh down the entire program because the number of less costly operations will far outnumber the costly ones in the long run, "paying back" the program over a number of iterations.<sup>[1]</sup> It is particularly useful because it guarantees worst-case performance rather than making assumptions about the state of the program.

## History

Amortized analysis initially emerged from a method called aggregate analysis, which is now subsumed by amortized analysis. However, the technique was first formally introduced by Robert Tarjan in his paper *Amortized Computational Complexity*, which addressed the need for a more useful form of analysis than the common probabilistic methods used. Amortization was initially used for very specific types of algorithms, particularly those involving binary trees and union operations. However, it is now ubiquitous and comes into play when analyzing many other algorithms as well.<sup>[1]</sup>

## Method

The method requires knowledge of which series of operations are possible. This is most commonly the case with data structures, which have state that persists between operations. The basic idea is that a worst case operation can alter the state in such a way that the worst case cannot occur again for a long time, thus "amortizing" its cost.

There are generally three methods for performing amortized analysis: the aggregate method, the accounting method, and the potential method. All of these give the same answers, and their usage difference is primarily circumstantial and due to individual preference.<sup>[2]</sup>

- Aggregate analysis determines the upper bound  $T(n)$  on the total cost of a sequence of  $n$  operations, then calculates the average cost to be  $T(n) / n$ .<sup>[2]</sup>
  - The accounting method determines the individual cost of each operation, combining its immediate execution time and its influence on the running time of future operations. Usually, many short-running operations accumulate a "debt" of unfavorable state in small increments, while rare long-running operations decrease it drastically.<sup>[2]</sup>
  - The potential method is like the accounting method, but overcharges operations early to compensate for undercharges later.<sup>[2]</sup>
-



## Common use

- In common usage, an "amortized algorithm" is one that an amortized analysis has shown to perform well.
- Online algorithms commonly use amortized analysis.

## References

- Allan Borodin and Ran El-Yaniv (1998). *Online Computation and Competitive Analysis* <sup>[3]</sup>. Cambridge University Press. pp. 20,141.
- [1] Rebecca Fiebrink (2007), *Amortized Analysis Explained* ([http://www.cs.princeton.edu/~fiebrink/423/AmortizedAnalysisExplained\\_Fiebrink.pdf](http://www.cs.princeton.edu/~fiebrink/423/AmortizedAnalysisExplained_Fiebrink.pdf)), , retrieved 2011-05-03
- [2] Vijaya Ramachandran (2006), *CS357 Lecture 16: Amortized Analysis* (<http://www.cs.utexas.edu/~vlr/s06.357/notes/lec16.pdf>), , retrieved 2011-05-03
- [3] <http://www.cs.technion.ac.il/~rani/book.html>

# Locality of reference

---

In computer science, **locality of reference**, also known as the **principle of locality**, is the phenomenon of the same value or related storage locations being frequently accessed. There are two basic types of reference locality. Temporal locality refers to the reuse of specific data and/or resources within relatively small time durations. Spatial locality refers to the use of data elements within relatively close storage locations. Sequential locality, a special case of spatial locality, occurs when data elements are arranged and accessed linearly, e.g., traversing the elements in a one-dimensional array.

Locality is merely one type of predictable behavior that occurs in computer systems. Systems which exhibit strong *locality of reference* are good candidates for performance optimization through the use of techniques like the cache and instruction prefetch technology for memory or the advanced branch predictor at the pipelining of processors.

## Locality of reference

The locality of reference, also known as the locality principle, is the phenomenon that the collection of the data locations referenced in a short period of time in a running computer often consists of relatively well predictable clusters. Important special cases of locality are *temporal*, *spatial*, *equidistant* and *branch* locality.

- **Temporal locality:** if at one point in time a particular memory location is referenced, then it is likely that the same location will be referenced again in the near future. There is a temporal proximity between the adjacent references to the same memory location. In this case it is common to make efforts to store a copy of the referenced data in special memory storage, which can be accessed faster. Temporal locality is a very special case of the spatial locality, namely when the prospective location is identical to the present location.
  - **Spatial locality:** if a particular memory location is referenced at a particular time, then it is likely that nearby memory locations will be referenced in the near future. In this case it is common to attempt to guess the size and shape of the area around the current reference for which it is worthwhile to prepare faster access.
  - **Branch locality:** if there are only few amount of possible alternatives for the prospective part of the path in the **spatial-temporal coordinate space**. This is the case when an instruction loop has a simple structure, or the possible outcome of a small system of conditional branching instructions is restricted to a small set of possibilities. Branch locality is typically not a spatial locality since the few possibilities can be located far away from each other.
  - **Equidistant locality:** it is halfway between the spatial locality and the branch locality. Consider a loop accessing locations in an equidistant pattern, i.e. the path in the **spatial-temporal coordinate space** is a dotted line. In this case, a simple linear function can predict which location will be accessed in the near future.
-

In order to make benefit from the very frequently occurring *temporal* and *spatial* kind of locality, most of the information storage systems are hierarchical; see below. The *equidistant* locality is usually supported by the diverse nontrivial increment instructions of the processors. For the case of *branch* locality, the contemporary processors have sophisticated branch predictors, and on the base of this prediction the memory manager of the processor tries to collect and preprocess the data of the plausible alternatives.

## Reasons for locality

There are several reasons for locality. These reasons are either goals to achieve or circumstances to accept, depending on the aspect. The reasons below are not disjoint; in fact, the list below goes from the most general case to special cases.

- **Predictability:** In fact, locality is merely one type of predictable behavior in computer systems. Luckily, many of the practical problems are decidable and hence the corresponding program can behave predictably, if it is well written.
- **Structure of the program:** Locality occurs often because of the way in which computer programs are created, for handling decidable problems. Generally, related data is stored in nearby locations in storage. One common pattern in computing involves the processing of several items, one at a time. This means that if a lot of processing is done, the single item will be accessed more than once, thus leading to temporal locality of reference. Furthermore, moving to the next item implies that the next item will be read, hence spatial locality of reference, since memory locations are typically read in batches.
- **Linear data structures:** Locality often occurs because code contains loops that tend to reference arrays or other data structures by indices. Sequential locality, a special case of spatial locality, occurs when relevant data elements are arranged and accessed linearly. For example, the simple traversal of elements in a one-dimensional array, from the base address to the highest element would exploit the sequential locality of the array in memory.<sup>[1]</sup> The more general *equidistant locality* occurs when the linear traversal is over a longer area of adjacent data structures having identical structure and size, and in addition to this, not the whole structures are in access, but only the mutually corresponding same elements of the structures. This is the case when a matrix is represented as a sequential matrix of rows and the requirement is to access a single column of the matrix.

## Use of locality in general

If most of the time the substantial portion of the references aggregate into clusters, and if the shape of this system of clusters can be well predicted, then it can be used for speed optimization. There are several ways to make benefit from locality. The common techniques for optimization are:

- to increase the locality of references. This is achieved usually on the software side.
- to exploit the locality of references. This is achieved usually on the hardware side. The *temporal* and *spatial* locality can be capitalized by hierarchical storage hardware. The *equidistant* locality can be used by the appropriately specialized instructions of the processors, this possibility is not only the responsibility of hardware, but the software as well, whether its structure is suitable for compiling a binary program which calls the specialized instructions in question. The *branch* locality is a more elaborate possibility, hence more developing effort is needed, but there is much larger reserve for future exploration in this kind of locality than in all the remaining ones.

## Use of spatial and temporal locality: hierarchical memory

Hierarchical memory is a hardware optimization that takes the benefits of spatial and temporal locality and can be used on several levels of the memory hierarchy. Paging obviously benefits from *temporal and spatial locality*. A cache is a simple example of exploiting temporal locality, because it is a specially designed faster but smaller memory area, generally used to keep recently referenced data and data near recently referenced data, which can lead to potential performance increases. Data in cache does not necessarily correspond to data that is spatially close in main memory; however, data elements are brought into cache one cache line at a time. This means that spatial locality is again important: if one element is referenced, a few neighboring elements will also be brought into cache. Finally, temporal locality plays a role on the lowest level, since results that are referenced very closely together can be kept in the machine registers. Programming languages such as C allow the programmer to suggest that certain variables are kept in registers.

Data locality is a typical memory reference feature of regular programs (though many irregular memory access patterns exist). It makes the hierarchical memory layout profitable. In computers, memory is divided up into a hierarchy in order to speed up data accesses. The lower levels of the memory hierarchy tend to be slower, but larger. Thus, a program will achieve greater performance if it uses memory while it is cached in the upper levels of the memory hierarchy and avoids bringing other data into the upper levels of the hierarchy that will displace data that will be used shortly in the future. This is an ideal, and sometimes cannot be achieved.

Typical memory hierarchy (access times and cache sizes are approximations of typical values used as of 2006 for the purpose of discussion; actual values and actual numbers of levels in the hierarchy vary):

- CPU registers (8-128 registers) – immediate access
- L1 CPU caches (32 KiB to 512 KiB) – fast access
- L2 CPU caches (128 KiB to 24 MiB) – slightly slower access
- Main physical memory (RAM) (256 MiB to 64 GiB) – slow access
- Disk (file system) (100 GiB to 10 TiB) – very slow
- Remote Memory (such as other computers or the Internet) (Practically unlimited) – speed varies

Modern machines tend to read blocks of lower memory into the next level of the memory hierarchy. If this displaces used memory, the operating system tries to predict which data will be accessed least (or latest) and move it down the memory hierarchy. Prediction algorithms tend to be simple to reduce hardware complexity, though they are becoming somewhat more complicated.

## Spatial and temporal locality example: matrix multiplication

A common example is matrix multiplication:

```
for i in 0..n
  for j in 0..m
    for k in 0..p
      C[i][j] = C[i][j] + A[i][k] * B[k][j];
```

When dealing with large matrices, this algorithm tends to shuffle data around too much. Since memory is pulled up the hierarchy in consecutive address blocks, in the C programming language it would be advantageous to refer to several memory addresses that share the same row (spatial locality). By keeping the row number fixed, the second element changes more rapidly. In C and C++, this means the memory addresses are used more consecutively. One can see that since *j* affects the column reference of both matrices *C* and *B*, it should be iterated in the innermost loop (this will fix the row iterators, *i* and *k*, while *j* moves across each column in the row). This will not change the mathematical result, but it improves efficiency. By switching the looping order for *j* and *k*, the speedup in large matrix multiplications becomes dramatic. (In this case, 'large' means, approximately, more than 100,000 elements in

each matrix, or enough addressable memory such that the matrices will not fit in L1 and L2 caches.)

Temporal locality can also be improved in the above example by using a technique called *blocking*. The larger matrix can be divided into evenly-sized sub-matrices, so that the smaller blocks can be referenced (multiplied) several times while in memory.

```
for (ii = 0; ii < SIZE; ii += BLOCK_SIZE)
    for (kk = 0; kk < SIZE; kk += BLOCK_SIZE)
        for (jj = 0; jj < SIZE; jj += BLOCK_SIZE)
            for (i = ii; i < ii + BLOCK_SIZE && i < SIZE; i++)
                for (k = kk; k < kk + BLOCK_SIZE && k < SIZE; k++)
                    for (j = jj; j < jj + BLOCK_SIZE && j < SIZE; j++)
                        C[i][j] = C[i][j] + A[i][k] * B[k][j];
```

The temporal locality of the above solution is provided because a block can be used several times before moving on, so that it is moved in and out of memory less often. Spatial locality is improved because elements with consecutive memory addresses tend to be pulled up the memory hierarchy together.

## Bibliography

- P.J. Denning, The Locality Principle, Communications of the ACM, Volume 48, Issue 7, (2005), Pages 19–24
- P.J. Denning, S.C. Schwartz, Communications of the ACM, Volume 15, Issue 3 (March 1972), Pages 191-198

## References

- [1] Aho, Lam, Sethi, and Ullman. "Compilers: Principles, Techniques & Tools" 2nd ed. Pearson Education, Inc. 2007

# Standard Template Library

---

The **Standard Template Library (STL)** is a C++ software library which heavily influenced many parts of the C++ Standard Library. It provides four components called *algorithms*, *containers*, *functional*, and *iterators*.<sup>[1]</sup>

The STL provides a ready-made set of common classes for C++, such as containers and associative arrays, that can be used with any built-in type and with any user-defined type that supports some elementary operations (such as copying and assignment). STL algorithms are independent of containers, which significantly reduces the complexity of the library.

The STL achieves its results through the use of templates. This approach provides compile-time polymorphism that is often more efficient than traditional run-time polymorphism. Modern C++ compilers are tuned to minimize any abstraction penalty arising from heavy use of the STL.

The STL was created as the first library of generic algorithms and data structures for C++, with four ideas in mind: generic programming, abstractness without loss of efficiency, the Von Neumann computation model,<sup>[2]</sup> and value semantics.

## Composition

### Containers

The STL contains sequence containers and associative containers. The standard sequence containers include

```
vector
```

```
,
```

```
deque
```

```
, and
```

```
list
```

. The standard associative containers are

```
set
```

```
,
```

```
multiset
```

```
,
```

```
map
```

```
, and
```

```
multimap
```

. There are also *container adaptors*

```
queue
```

```
,
```

```
priority_queue
```

```
, and
```

```
stack
```

, that are containers with specific interface, using other containers as implementation.

Container	Description
<b>Simple Containers</b>	
<b>pair</b>	The pair container is a simple associative container consisting of a 2-tuple of data elements or objects, called 'first' and 'second', in that fixed order. The STL 'pair' can be assigned, copied and compared. The array of objects allocated in a map or hash_map (described below) are of type 'pair' by default, where all the 'first' elements act as the unique keys, each associated with their 'second' value objects.
<b>Sequences (Arrays/Linked Lists): ordered collections</b>	
<b>vector</b>	a dynamic array, like C array (i.e., capable of random access) with the ability to resize itself automatically when inserting or erasing an object. Inserting and removing an element to/from back of the vector at the end takes amortized constant time. Inserting and erasing at the beginning or in the middle is linear in time. A specialization for type bool exists, which optimizes for space by storing bool values as bits.
<b>list</b>	a doubly linked list; elements are not stored in contiguous memory. Opposite performance from a vector. Slow lookup and access (linear time), but once a position has been found, quick insertion and deletion (constant time).
<b>deque</b> ( <i>double-ended queue</i> )	a vector with insertion/erase at the beginning or end in amortized constant time, however lacking some guarantees on iterator validity after altering the deque.
<b>Container adaptors</b>	

queue	Provides FIFO queue interface in terms of
	push
	/
	pop
	/
	front
	/
	back
	operations. Any sequence supporting operations
	front ()
	,
	back ()
	,
	push_back ()
	, and
	pop_front ()
	can be used to instantiate queue (e.g.
	list
	and
	deque
	).

priority_queue	Provides priority queue interface in terms of
	push/pop/top
	operations (the element with the highest priority is on top). Any random-access sequence supporting operations
	front ()
	,
	push_back ()
	, and
	pop_back ()
	can be used to instantiate priority_queue (e.g.
	vector
	and
	deque
	).
	Elements should additionally support comparison (to determine which element has a higher priority and should be popped first).



<b>stack</b>	<p>Provides LIFO stack interface in terms of</p> <p><code>push/pop/top</code></p> <p>operations (the last-inserted element is on top). Any sequence supporting operations</p> <p><code>back()</code></p> <p>,</p> <p><code>push_back()</code></p> <p>, and</p> <p><code>pop_back()</code></p> <p>can be used to instantiate stack (e.g.</p> <p><code>vector</code></p> <p>,</p> <p><code>list</code></p> <p>, and</p> <p><code>deque</code></p> <p>).</p>
<b>Associative containers: unordered collections</b>	
<b>set</b>	<p>a mathematical set; inserting/erasing elements in a set does not invalidate iterators pointing in the set. Provides set operations union, intersection, difference, symmetric difference and test of inclusion. Type of data must implement comparison operator</p> <p><code>&lt;</code></p> <p>or custom comparator function must be specified; such comparison operator or comparator function must guarantee strict weak ordering, otherwise behavior is undefined. Typically implemented using a self-balancing binary search tree.</p>
<b>multiset</b>	same as a set, but allows duplicate elements.
<b>map</b>	<p>an associative array; allows mapping from one data item (a key) to another (a value). Type of key must implement comparison operator</p> <p><code>&lt;</code></p> <p>or custom comparator function must be specified; such comparison operator or comparator function must guarantee strict weak ordering, otherwise behavior is undefined. Typically implemented using a self-balancing binary search tree.</p>
<b>multimap</b>	same as a map, but allows duplicate keys.

<b>hash_set</b> <b>hash_multiset</b> <b>hash_map</b> <b>hash_multimap</b>	<p>similar to a set, multiset, map, or multimap, respectively, but implemented using a hash table; keys are not ordered, but a hash function must exist for the key type. These containers are not part of the C++ Standard Library, but are included in SGI's STL extensions, and are included in common libraries such as the GNU C++ Library in the</p> <p><code>__gnu_cxx</code></p> <p>namespace. These are scheduled to be added to the C++ standard as part of TR1, with the slightly different names of</p> <p><code>unordered_set</code></p> <p>,</p> <p><code>unordered_multiset</code></p> <p>, <code>unordered_map</code> and</p> <p><code>unordered_multimap</code></p> <p>.</p>
<b>Other types of containers</b>	
<b>bitset</b>	stores series of bits similar to a fixed-sized vector of bools. Implements bitwise operations and lacks iterators. Not a Sequence.
<b>valarray</b>	another C-like array like vector, but is designed for high speed numerics at the expense of some programming ease and general purpose use. It has many features that make it ideally suited for use with vector processors in traditional vector supercomputers and SIMD units in consumer-level scalar processors, and also ease vector mathematics programming even in scalar computers.

## Iterators

The STL implements five different types of iterators. These are *input iterators* (that can only be used to read a sequence of values), *output iterators* (that can only be used to write a sequence of values), *forward iterators* (that can be read, written to, and move forward), *bidirectional iterators* (that are like forward iterators, but can also move backwards) and *random access iterators* (that can move freely any number of steps in one operation).

It is possible to have bidirectional iterators act like random access iterators, as moving forward ten steps could be done by simply moving forward a step at a time a total of ten times. However, having distinct random access iterators offers efficiency advantages. For example, a vector would have a random access iterator, but a list only a bidirectional iterator.

Iterators are the major feature that allow the generality of the STL. For example, an algorithm to reverse a sequence can be implemented using bidirectional iterators, and then the same implementation can be used on lists, vectors and dequeues. User-created containers only have to provide an iterator that implements one of the five standard iterator interfaces, and all the algorithms provided in the STL can be used on the container.

This generality also comes at a price at times. For example, performing a search on an associative container such as a map or set can be much slower using iterators than by calling member functions offered by the container itself. This is because an associative container's methods can take advantage of knowledge of the internal structure, which is opaque to algorithms using iterators.

## Algorithms

A large number of algorithms to perform operations such as searching and sorting are provided in the STL, each implemented to require a certain level of iterator (and therefore will work on any container that provides an interface by iterators).

## Functors

The STL includes classes that overload the function operator (

```
operator ()
```

). Classes that do this are called functors or function objects. They are useful for keeping and retrieving state information in functions passed into other functions. Regular function pointers can also be used as functors.

A particularly common type of functor is the predicate. For example, algorithms like

```
find_if
```

take a unary predicate that operates on the elements of a sequence. Algorithms like `sort`, `partial_sort`, `nth_element` and all sorted containers use a binary predicate that must provide a strict weak ordering, that is, it must behave like a membership test on a transitive, irreflexive and antisymmetric binary relation. If none is supplied, these algorithms and containers use `less` <sup>[3]</sup> by default, which in turn calls the `less-than-operator` `<`.

## History

The architecture of STL is largely the creation of Alexander Stepanov. In 1979 he began working out his initial ideas of generic programming and exploring their potential for revolutionizing software development. Although David Musser had developed and advocated some aspects of generic programming already by year 1971, it was limited to a rather specialized area of software development (computer algebra).

Stepanov recognized the full potential for generic programming and persuaded his then-colleagues at General Electric Research and Development (including, primarily, David Musser and Deepak Kapur) that generic programming should be pursued as a comprehensive basis for software development. At the time there was no real support in any programming language for generic programming.

The first major language to provide such support was Ada (ANSI standard 1983), with its generic units feature. In 1985, the Eiffel programming language became the first object-oriented language to include intrinsic support for generic classes, combined with the object-oriented notion of inheritance.<sup>[4]</sup> By 1987 Stepanov and Musser had developed and published an Ada library for list processing that embodied the results of much of their research on generic programming. However, Ada had not achieved much acceptance outside the defense industry and C++ seemed more likely to become widely used and provide good support for generic programming even though the language was relatively immature. Another reason for turning to C++, which Stepanov recognized early on, was the C/C++ model of computation that allows very flexible access to storage via pointers, which is crucial to achieving generality without losing efficiency.

Much research and experimentation were needed, not just to develop individual components, but to develop an overall architecture for a component library based on generic programming. First at AT&T Bell Laboratories and later at Hewlett-Packard Research Labs (HP), Stepanov experimented with many architectural and algorithm formulations, first in C and later in C++. Musser collaborated in this research and in 1992 Meng Lee joined Stepanov's project at HP and became a major contributor.

This work undoubtedly would have continued for some time being just a research project or at best would have resulted in an HP proprietary library, if Andrew Koenig of Bell Labs had not become aware of the work and asked

Stepanov to present the main ideas at a November 1993 meeting of the ANSI/ISO committee for C++ standardization. The committee's response was overwhelmingly favorable and led to a request from Koenig for a formal proposal in time for the March 1994 meeting. Despite the tremendous time pressure, Alex and Meng were able to produce a draft proposal that received preliminary approval at that meeting.

The committee had several requests for changes and extensions (some of them major), and a small group of committee members met with Stepanov and Lee to help work out the details. The requirements for the most significant extension (associative containers) had to be shown to be consistent by fully implementing them, a task Stepanov delegated to Musser. It would have been quite easy for the whole enterprise to spin out of control at this point, but again Stepanov and Lee met the challenge and produced a proposal that received final approval at the July 1994 ANSI/ISO committee meeting. (Additional details of this history can be found in Stevens.) Subsequently, the Stepanov and Lee document 17 was incorporated into the ANSI/ISO C++ draft standard (1, parts of clauses 17 through 27). It also influenced other parts of the C++ Standard Library, such as the string facilities, and some of the previously adopted standards in those areas were revised accordingly.

In spite of STL's success with the committee, there remained the question of how STL would make its way into actual availability and use. With the STL requirements part of the publicly available draft standard, compiler vendors and independent software library vendors could of course develop their own implementations and market them as separate products or as selling points for their other wares. One of the first edition's authors, Atul Saini, was among the first to recognize the commercial potential and began exploring it as a line of business for his company, Modena Software Incorporated, even before STL had been fully accepted by the committee.

The prospects for early widespread dissemination of STL were considerably improved with Hewlett-Packard's decision to make its implementation freely available on the Internet in August 1994. This implementation, developed by Stepanov, Lee, and Musser during the standardization process, became the basis of many implementations offered by compiler and library vendors today.

## Criticisms

### Quality of implementation

The Quality of Implementation (QoI) of the C++ compiler has a large impact on usability of STL (and templated code in general):

- Error messages involving templates tend to be very long and difficult to decipher. This problem has been considered so severe that a number of tools have been written that simplify and prettyprint STL-related error messages to make them more comprehensible.
  - Careless use of STL templates can lead to code bloat. This has been countered with special techniques within STL implementation (using `void*` containers internally) and by improving optimization techniques used by compilers.
  - Template instantiation tends to increase compilation time and memory usage (even by an order of magnitude). Until the compiler technology improves enough, this problem can be only partially eliminated by very careful coding and avoiding certain idioms.
-

## Other issues

- Initialization of STL containers with constants within the source code is not as easy as data structures inherited from C (addressed in C++11 with initializer lists).
- STL containers are not intended to be used as base classes (their destructors are deliberately non-virtual); deriving from a container is a common mistake.<sup>[2][5]</sup>
- The concept of iterators as implemented by STL can be difficult to understand at first: for example, if a value pointed to by the iterator is deleted, the iterator itself is then no longer valid. This is a common source of errors. Most implementations of the STL provide a debug mode that is slower, but can locate such errors if used. A similar problem exists in other languages, for example Java. Ranges have been proposed as a safer, more flexible alternative to iterators.<sup>[6]</sup>
- Certain iteration patterns do not map to the STL iterator model. For example, callback enumeration APIs cannot be made to fit the STL model without the use of coroutines,<sup>[7]</sup> which are platform-dependent or unavailable, and are outside the C++ standard.
- Compiler compliance does not guarantee that Allocator objects, used for memory management for containers, will work with state-dependent behavior. For example, a portable library can't define an allocator type that will pull memory from different pools using different allocator objects of that type. (Meyers, p. 50) (addressed in C++11).
- The set of algorithms is not complete: for example, the

```
copy_if
```

algorithm was left out,<sup>[8]</sup> though it has been added in C++11.<sup>[9]</sup>

- The interface of some containers (in particular string) is argued to be bloated (Sutter and Alexandrescu, p. 79); others are argued to be insufficient.
- Hashing containers were left out of the original standard, but have been added in C++11 and in Technical Report 1, a recent extension to C++.

## Implementations

- Original STL implementation by Stepanov and Lee. 1994, Hewlett-Packard. No longer maintained.
- SGI STL, based on original implementation by Stepanov & Lee. 1997, Silicon Graphics. No longer maintained.
- libstdc++ from gnu (was part of libg++)
- libc++ from clang
- STLPort, based on SGI STL
- Rogue Wave standard library (HP, SGI, SunSoft, Siemens-Nixdorf)
- Dinkum STL library by P.J. Plauger
  - The Microsoft STL<sup>[10]</sup> which ships with Visual C++ is a licensed derivative of Dinkum's STL.
- Apache C++ Standard Library<sup>[11]</sup> (The copyright notes in the source files indicate that this is based on - or the same as - the Rogue Wave standard library.)

## Notes

- [1] Holzner, Steven (2001). *C++ : Black Book*. Scottsdale, Ariz.: Coriolis Group. p. 648. ISBN 1-57610-777-9. "The STL is made up of *containers, iterators, function objects, and algorithms*"
- [2] Musser, David (2001). *STL tutorial and reference guide: C++ programming with the standard template library*. Addison Wesley. ISBN 0-201-37923-6.
- [3] <http://www.sgi.com/tech/stl/less.html>
- [4] Meyer, Bertrand. *Genericity versus inheritance*, in ACM Conference on Object-Oriented Programming Languages Systems and Applications (OOPSLA), Portland (Oregon), September 29 - October 2, 1986, pages 391-405.
- [5] Sutter, Herb; Alexandrescu, Andrei (2004). *C++ Coding Standards: 101 Rules, Guidelines, and Best Practices*. Addison-Wesley. ISBN 0-321-11358-6.
- [6] Andrei Alexandrescu (6 May 2009). "Iterators Must Go" ([https://github.com/boostcon/2009\\_presentations/raw/master/wed/iterators-must-go.pdf](https://github.com/boostcon/2009_presentations/raw/master/wed/iterators-must-go.pdf)). BoostCon 2009. . Retrieved 19 March 2011.
- [7] Matthew Wilson (February 2004). "Callback Enumeration APIs & the Input Iterator Concept" (<http://www.ddj.com/cpp/184401766>). *Dr. Dobb's Journal*. .
- [8] Bjarne Stroustrup (2000). *The C++ Programming Language* (3rd ed.). Addison-Wesley. ISBN 0-201-70073-5.<sup>p.530</sup>
- [9] More STL algorithms (revision 2) (<http://www.open-std.org/jtc1/sc22/wg21/docs/papers/2008/n2666.pdf>)
- [10] [http://msdn.microsoft.com/en-us/library/c191tb28\(v=vs.80\).aspx](http://msdn.microsoft.com/en-us/library/c191tb28(v=vs.80).aspx)
- [11] <http://stdcxx.apache.org>

## References

- Alexander Stepanov and Meng Lee, The Standard Template Library. HP Laboratories Technical Report 95-11(R.1), 14 November 1995. (Revised version of A. A. Stepanov and M. Lee: The Standard Template Library, Technical Report X3J16/94-0095, WG21/N0482, ISO Programming Language C++ Project, May 1994.) (<http://www.stepanovpapers.com>)
- Alexander Stepanov (2007) (PDF). *Notes on Programming* (<http://www.stepanovpapers.com/notes.pdf>). Stepanov reflects about the design of the STL.
- Nicolai M. Josuttis (2000). *The C++ Standard Library: A Tutorial and Reference*. Addison-Wesley. ISBN 0-201-37926-0.
- Scott Meyers (2001). *Effective STL: 50 Specific Ways to Improve Your Use of the Standard Template Library*. Addison-Wesley. ISBN 0-201-74962-9.
- Al Stevens (March 1995). "Al Stevens Interviews Alex Stepanov" (<http://www.sgi.com/tech/stl/drddobbs-interview.html>). *Dr. Dobb's Journal*. Retrieved 18 July 2007.
- David Vandevoorde and Nicolai M. Josuttis (2002). *C++ Templates: The Complete Guide*. Addison-Wesley Professional. ISBN 0-201-73484-2.
- Atul Saini and David R. Musser, *STL Tutorial and Reference Guide: C++ Programming with the Standard Template Library*. Foreword by Alexander Stepanov; [Copyright Modena Software Inc.] Addison-Wesley ISBN 0-201-63398-1

## External links

- C/C++ STL reference (<http://en.cppreference.com/w/cpp/container>), includes C++0x features
- STL programmer's guide (<http://www.sgi.com/tech/stl/>) guide from SGI
- Apache (formerly Rogue Wave) C++ Standard Library Class Reference (<http://stdcxx.apache.org/doc/stdlibref/index.html>)
- Apache (formerly Rogue Wave) C++ Standard Library User Guide (<http://stdcxx.apache.org/doc/stdlibug/index.html>)
- Bjarne Stroustrup on The emergence of the STL (<http://www.research.att.com/~bs/DnE2005.pdf>) (Page 5, Section 3.1)
- A series of blog postings of C++ STL examples (<http://www.technical-recipes.com/stl-boost/>)

# Article Sources and Contributors

**Data structure** *Source:* <http://en.wikipedia.org/w/index.php?oldid=505440284> *Contributors:* -- April, 195.149.37.xxx, 24.108.14.xxx, Abd, Abhishek.kumar.ak, Adrianwn, Ahoerstemeier, Ahy1, Alansohn, Alexius08, Alhoori, Allan McInnes, Altenmann, Anderson, Andre Engels, Andreas Kaufmann, Antonielli, Ap, Apoclyptic, Arjayay, Arvindn, Babbage, Bharatshettybarkur, BioPupil, Bluemoose, BurntSky, Bushytails, CRGreathouse, Caiaffa, Caltas, Carlette, Chandraguptamaurya, Chris Lundberg, Closedmouth, Cncplayer, Coldfire82, Conversion script, Corti, Cpl Syx, Craig Stuntz, DAndC, DCDuring, DRAGON BOOSTER, DancingPhilosopher, Danim, DavidCary, Dcoetzee, Demicx, Derbeth, Digisus, Dmoss, Dougher, Easyas12c, EconoPhysicist, EdEColbert, Edaelon, Er Komandante, Esap, Eurooppa, Eve Hall, Falcon8765, FinalMinuet, Fraggel, Frap, Fresheneesz, GPhilip, Garyzx, GeorgeBills, Ghyll, Giftlite, Gilliam, Glenn, Gmharhar, Googl, GreatWhiteNortherner, Hairy Dude, Haiviet, Ham Pastrami, Helix84, Hypersonic12, IGeMiNix, Iridescent, JLaTondre, Jacob grace, Jerryobject, Jiang, Jim1138, Jimmytharpe, Jirka6, Jncraton, Jorge Stolfi, Jorgenev, Justin W Smith, Karl E. V. Palmen, Kh31311, Khukri, Kingpin13, Kingturtle, Kjetil r, Koavf, LC, Lancekt, Lanov, Laurențiu Dascălu, Liao, Ligulem, Liridon, Lithui, Loadmaster, Lotje, MTA, Mahanga, Mandarax, Marcin Suwalczan, Mark Renier, MasterRadius, Materialscientist, Mdd, MertyWiki, Methcub, Michael Hardy, Mindmatrix, Minesweeper, Mipadi, MisterSheik, MithrandirAgain, Miym, Morel, Mr Stephen, MrOllie, MrJeff, Mushroom, Nanshu, Nick Levine, Nikola Smolenski, Nnp, Noah Salzman, Noldoaran, Nskillen, Obradovic Goran, Ohnoitsjamie, Oicumayberight, Orderud, PaePae, Pale blue dot, Panchobook, Pascal.Tesson, Pausali, Peterdjones, Pgallert, Pgan002, Piet Delpoit, Populus, Prari, Publichealthguru, Pur3r4ngelw, Qwyrxian, Ramkumaran7, Raveendra Lakpriya, Reedy, Requestion, Rettetast, RexNL, ReyBrujo, Rhwawn, Richfaber, Ripper234, Rodhullandemu, Rrwright, Ruud Koot, Ryan Roos, Sallupandit, Sanjay742, Seth Ilys, Sethwoodworth, Sgord512, Shadowjams, Shanes, Sharcho, Siroxo, SoniyaR, Soumyasch, Spellinger180, SteelPangolin, Strife911, Sundar sando, Tablizer, TakuyaMurata, Tanvir Ahmed, Tas50, Tbotch, Teles, Thadius856, The Thing That Should Not Be, Thecheesekid, Thinktdub, Thompsonb24, Thunderboltz, Tide rolls, Tobias Bergemann, Tom 99, Tony1, Traroth, TreveX, TuukkaH, Uriah123, User A1, UserGoogol, Varma rockzz, Vicarious, Vineetzone, Vipinhari, Viriditas, Vishnu0919, Vortexrealm, Walk&check, Wbm1058, Widefox, Wikilolo, Wmbolle, Wrp103, Wwmbes, XJamRastafire, Yama, Yashykt, Yorice, ماتي, ماتي, 454 anonymous edits

**Linked data structure** *Source:* <http://en.wikipedia.org/w/index.php?oldid=502941011> *Contributors:* Andreas Kaufmann, Anuradha tupsundare, BD2412, Bunnyhop11, ChrisHodgesUK, Cybercobra, Gmiller4th, Headbomb, IshitaMundada, Jeffrey.Rodriguez, Jorge Stolfi, Katieh5584, Miym, SchreiberBike, Spartlow, Strait, 20 anonymous edits

**Succinct data structure** *Source:* <http://en.wikipedia.org/w/index.php?oldid=492030944> *Contributors:* Algotime, Andreas Kaufmann, Camilo Sanchez, David Eppstein, Glacialfox, JyBy, Kmanmike15, Magioladitis, Ryk, Sciyoshi, Shalom Yechiel, Smhanov, Whyrn, 9 anonymous edits

**Implicit data structure** *Source:* <http://en.wikipedia.org/w/index.php?oldid=500006640> *Contributors:* Andreas Kaufmann, Azylber, Borseashwini, Dcoetzee, Emilkeyder, Headbomb, Ilyathemuromets, Jesse V., RainbowCrane, Rangilo Gujarati, Runtime, Ruud Koot, Schizobullet, Shrutika girme, 6 anonymous edits

**Compressed data structure** *Source:* <http://en.wikipedia.org/w/index.php?oldid=350086433> *Contributors:* Algotime, Andreas Kaufmann, Eekster

**Search data structure** *Source:* <http://en.wikipedia.org/w/index.php?oldid=496528278> *Contributors:* Andreas Kaufmann, Clayhalliwell, Devendermishra, GTBacchus, Headbomb, Hosamaly, Jorge Stolfi, Quibik, Scandum, Trigger hurt, Vicarious, Woodshed, X7q, 13 anonymous edits

**Persistent data structure** *Source:* <http://en.wikipedia.org/w/index.php?oldid=508572804> *Contributors:* Allen3, Belovedeagle, BiT, Billgordon1099, Cdiggin, Charles Matthews, Chris the speller, Chris0804, David Eppstein, David1217, Dcoetzee, Dfeuer, DivideByZero14, Edward, Haeleth, Harej, Headbomb, Hooperbloob, JiBB, John Nowak, Jwang210, Ken Hirsch, MarSch, MelbourneStar, Peruvianllama, Pomte, Qwertyus, Ratan203, Rory O'Kane, Ruud Koot, Scarverwiki, Seunosewa, Sophie means wisdom, Spoon!, Svick, Tehom2000, Tobias Bergemann, Tomaxer, 39 anonymous edits

**Concurrent data structure** *Source:* <http://en.wikipedia.org/w/index.php?oldid=473162734> *Contributors:* Andreas Kaufmann, Grafen, Headbomb, Hervegirod, Hyang04, John of Reading, Khizmax, LilHelpa, Rjwilmsi, Ronjhones, Senthryl, Uriah123, 14 anonymous edits

**Abstract data type** *Source:* <http://en.wikipedia.org/w/index.php?oldid=506009601> *Contributors:* 209.157.137.xxx, A5b, Adrianwn, Aitias, Alansohn, Andreas Kaufmann, Anna Lincoln, Antonielli, Appicharlask, Aqualung, Arjun024, Ark, Armin Rigo, B4hand, Babayaggyypsies, Baudway, BenRG, Blaisorblade, Bluebusy, Boing! said Zebedee, Brick Thrower, Cadadril, Chevymontecarlo, Chris the speller, Cobi, Conversion script, Corti, Cybercobra, DGaw, Daniel Brockman, Daniel.Burckhardt, Debresser, Demonkoryu, Diego Moya, Dismantle101, Don4of4, Double Dribble, Dreadstar, Dunks58, Efphf, Everton137, Felipe1982, Fishnet37222, Garyzx, GhettoBlaster, Giftlite, Gnfrf, Haakon, Hoorayforturtles, Ideogram, JC Chu, Japanese Searobin, Jonathan.mark.lingard, Jorge Stolfi, Jpvnall, Kbdank71, Khrose, Kendrick Hang, Knutux, Leif, Lights, Liztamp, Mark Renier, Marudubshinki, Merphant, Miaow Miaow, Michael Hardy, Mild Bill Hiccup, Mr Adequate, Nhey24, Noldoaran, Only2sea, Pcap, PeterV1510, Petri Krohn, Phuzion, Pink18, Populus, R. S. Shaw, RJHall, Reconsider the static, Rich Farmbrough, Ruud Koot, SAE1962, Sagaciousuk, SchiftyThree, Sean Whitton, Sector7agent, Silvonen, Skysmith, SolKarma, SpallettaAC1041, Spoon!, Svick, The Arbiter, The Thing That Should Not Be, Thecheeskykid, Thehelpfulone, Tobias Bergemann, TuukkaH, W7cook, Wapcaplet, Wavelength, Wernher, Widefox, Wrp103, Yerro, Zacchiro, 237 anonymous edits

**List** *Source:* <http://en.wikipedia.org/w/index.php?oldid=482704255> *Contributors:* Adrianwn, Alihaq717, Altenmann, Andreas Kaufmann, Andrew Eisenberg, Angela, Bomazi, BradBeattie, Brick Thrower, Calexico, Chevan, Chowbok, Chris the speller, Christian List, Classicaecon, Cmdrjameson, Crater Creator, Cybercobra, Daniel Brockman, Delirium, Denispir, Dgreen34, Dijkstra, Dismantle101, Docu, Drag, Eao, Ed Cormany, Elaz85, Elf, Elwikipedista, EugeneZelenko, Falk Lieder, Fredrik, Gaius Cornelius, Glenn, HQCentral, Ham Pastrami, Hyacinth, Jan Hidders, Jareha, Jeff3000, Jeffrey Mall, Jimmisbl, Jorge Stolfi, Joseghr, Josh Parris, Joswig, Ketiltrout, Liao, ManN, Mav, Mic, Michael Hardy, Mickeymousechen, Mike.nicholaides, Mipadi, Neilc, Noldoaran, P0nc, Paddy3118, Palmard, Patrick, Paul G, Paul foord, Pcap, Peak, Poor Yorick, Prumpf, Puckly, R. S. Shaw, Ruud Koot, Salix alba, Samuelsen, Spoon!, Stormie, TShilo12, TakuyaMurata, The Thing That Should Not Be, Tokek, VictorAnyakin, WODUP, Wavelength, Wbm1058, WillNess, Wmahan, Wnissen, Wwwwolf, XJamRastafire, ZeroOne, 83 anonymous edits

**Stack** *Source:* <http://en.wikipedia.org/w/index.php?oldid=452924945> *Contributors:* 144.132.75.xxx, 1exec1, 202.144.44.xxx, 2607:F140:400:1008:61A0:5A79:C3CA:1D2E, 2mcm, Aaron Rotenberg, Aavvof, Abhidesai, Adam majewski, Adam78, Agateller, Ahluka, Aillema, Aitias, Al Lemos, Alxeedo, Andre Engels, Andreas Kaufmann, Andrej, Angusmcellan, Arch dude, Arkahot, Arvindn, BenFrantzDale, BiT, BliizzmasterPilch, Bobo192, Boivie, Bookmaker, Borgx, Bsmntmbombdood, Caerwine, Calliopejen1, CanisRufus, Ceriak, Cheusov, Chris the speller, ChrisGualtieri, Christian List, Cjhoyle, Clx321, Cmcormick8, Colin meier, Conversion script, CoolKoon, Corti, CosineKitty, Cybercobra, Dan Granahan, David Eppstein, David.Federman, Davidhormen, Dcoetzee, Dhardik007, Dilleca, Dinoen, ENeville, Edgar181, ElNuevoEinstein, F15x28, Faysol037, Fernandopabon, Finlay McWalter, Flaqueteo, Fragment, Fredrik, FrontLine, Funandtrvl, Funky Monkey, Gegc, Gggh, GhettoBlaster, Giftlite, Gonzen, Graham87, Grue, Guy Peters, Gwermol, Hackwrench, Ham Pastrami, Hangfromthefloor, Hariva, Headbomb, Hgfernan, Hook43113, Hqb, IITManojit, IanOsmond, Ianbogges, Individual X, IntrigueBlue, Ionutzomuse, Iridescent, Ixfd64, Jake Nelson, James Foster, Jeff G., JensMueller, Jesse Viviano, Jfmanitis, Jheiv, Johnunig, Jprg1966, Jzalaee, Karl-Henner, Kbdank71, Klower, KnightRider, Kushalbiswas777, L Kensington, Liao, Loadmaster, LobStoR, Luciform, Maashatra11, Macrakis, Mahlon, Manassehkatz, Mandarax, Marc Mongenet, Mark Renier, MartinHarper, MattGiuca, Maxim Razin, Maximaximax, Mbessey, Mdd, MegaHasher, Melizg, Mentifisto, Michael Hardy, Michael Slone, Mindmatrix, Mipadi, Mlpkr, Modster, Mohinib27, Mr. Stradivarius, Murray Langton, Musiphil, Myasuda, Nakarumaka, Nbarth, Netkinetic, Nipunbayas, NoirNoir, Noldoaran, Notheruser, Nova2358, Nutster, Obradovic Goran, OLEnglish, Oli Filth, Patrick, PeterJeremy, Physicistjedi, Pion, Poccil, Pomte, Postrach, PranavAmbhore, Quantran202, R'nB, R. S. Shaw, RDBrown, RTC, Raviemani, Reikon, RevRagnarok, ReyBrujo, Robbe, Robert impey, Robin400, Rpv.imcc, Rustamabd, Ruud Koot, Rwxrwxrwx, Salocin-yel, Sanjay742, Seaphoto, Seth Manapio, Shuipzv3, SimenH, SiobhanHansa, Slgrandson, Spieren, Spoon!, SpyMagician, Stan Shebs, StanBally, Stephenb, Stevenbird, Streac, Stw, TakuyaMurata, Tapkeerrambo007, Tasc, The Anome, Thumperward, Tom.Reding, Traroth, Tsf, TuukkaH, Ultranaut, Unara, VTBassMatt, VampWillow, Vasiyli Faronov, Vishal G.Dhavale., Vystrix Nexoth, Whosasking, Wikidan829, Wikilolo, WiseWoman, Wj32, Wlievens, Xdenizen, Yammessicka, Zchenyu, Ztothefifth, 345 anonymous edits

**Queue** *Source:* <http://en.wikipedia.org/w/index.php?oldid=452925452> *Contributors:* 16@r, Ahoerstemeier, Akerans, Almkglor, Andre Engels, Andreas Kaufmann, Arsenic99, Atiflz, Banditlord, BenFrantzDale, Bobo2000, Brain, Caerwine, Caesura, Calliopejen1, Carlosguitar, Cdills, Chairboy, Chelseafan528, Chris the speller, Christian75, Kcatz, Clehner, Conan, Contactbanish, Conversion script, Corti, Cybercobra, Dabear, DavidLevinson, Dcoetzee, Deflective, Detonadorado, Discospinster, Dmitrysobolev, Edward, Egerhart, Emperorbna, Ewlyahoocom, Fredrik, Fswangke, Furrykef, Garfieldnate, Gbduende, Ggia, Giftlite, Glenn, GrahamDavies, Gralfca, Gunslinger47, Ham Pastrami, Hariva, Helix84, Hires an editor, Honza Záruba, Howcheng, Indil, Iprathik, Ixfd64, J. M., JC Chu, Jesin, Jguk, JohJak2, John lindgren, Joseph.w.s, JosephBarillari, Jrtayloriv, Jusjih, Keilana, Kenyon, Kflorenc, Kletos, Kusli10, Kushalbiswas777, Kwamikagami, LapoLuchini, Liao, Loupeter, Lopez2029, M2MM4M, MahlerFive, Mark Renier, Marry314113, Massysett, MattGiuca, Maw, Maxwellttery, Mecanismo, Mehrenberg, Metasquares, Michael Hardy, Mike1024, MikeDunlavey, Miklccet, Mindmatrix, Mlpkr, MrOllie, Nanshu, Nemo Kartikeyan, Noldoaran, Nutster, Nwbeeson, Oli Filth, OliverTwisted, Olivier Teuliere, Patrick, Peng, PhilipR, PhuksyWiki, Pissant, PrometheusFeu, PseudoSudo, Qwertyus, Rachel1, Rahulghose, Rasmus Faber, Rdsmith4, Redhanker, Ruby.red.roses, Ruud Koot, Sanjay742, SensuiShinobu1234, Sharcho, SimenH, SiobhanHansa, SoSaysChappy, Some jerk on the Internet, Sorancio, Spoon!, SpuriousQ, Statsats, Stephenb, Thadius856, Thesuperslacke, Tobias Bergemann, TobiasPenson, Traroth, Tsemii, Uriuamme, VTBassMatt, Vanmaple, Vegpuff, W3bbo, Wikibarista, Wikilolo, Woohookitty, Wouter.oet, Wrp103, X96lee15, Zachlpton, Zanaferx, Zoney, Zotel, Ztothefifth, Zvar, ماتي, 229 anonymous edits

**Deque** *Source:* <http://en.wikipedia.org/w/index.php?oldid=499941465> *Contributors:* Amairlang, Andreas Kaufmann, Anonymous Dissident, Bcbell, BenFrantzDale, CesarB, Chowbok, Chris the speller, Conversion script, Cparkr, Cybercobra, David Eppstein, Dcoetzee, Drmrngy, E. Necess, Esrogs, Fabartus, Fberg, Felix C. Stegerman, Ffaarr, Foxj, Frecklefoot, Fredrik, Funnyfarmofdoom, Furrykef, Ham Pastrami, Hawk777, Headbomb, JengelH, KILNA, Kbrose, Kuryk, LittleWink, Luder, Mandarax, Merovingian, Mindmatrix, Mwhitlock, Narahat, Offby1, Oli Filth, Omicronpersei8, Psphiorg, Pt, Puetzk, Ramiromagalhaes, Rasmus Faber, Ripper234, Rosen, Rrmsjp, Ruud Koot, Sae1962, Schellhammer, Shire Reeve, Silly rabbit, Sj, Sneftel, Spoon!, SpuriousQ, The Anome, TimBentley, Tobias Bergemann, VictorAnyakin, Wikibofh, Wolkykim, Zoicon5, Ztothefifth, 112 anonymous edits

**Priority queue** *Source:* <http://en.wikipedia.org/w/index.php?oldid=506979802> *Contributors:* 1exec1, Andreas Kaufmann, Arkenflame, BACbKA, Bdonlan, Bestiasonica, BlackFingolfin, Bobo192, Byrial, Chrisahn, Clicketyclack, Coder Dan, Conversion script, CorpX, CosineKitty, Cybercobra, David Eppstein, Dbeardsl, Dcoetzee, Decrypt3, El C, Emeraldemon, EnOreg, FUZxxl, Frecklefoot, Fredrik, Ghetoblaster, Giftlite, Gilliam, Happyuk, Hdante, HenryAyoola, HighwayJitchhiker, Hobsonlane, Ilingod, Itusg15q4user, J.C. Labbrev, Jeff02, Jeltz, Jncraton, John Reed Riley, Jutiphann, Kenyon, Kostmo, Kowey, Krungie factor, Kushalbiswas777, LastKingpin, LeeG, Mbloore, MeekMelange, Mentifisto, Michael Hardy, Ninjagecko, Nixdorf, Nyenyec, Oli Filth, Oliphant, Omicronpersei8, Orangeroof, Orderud, PaulWright, Pdelong, Pete142, Qwertyus, RHaden, Red Act, Redroof, Rhanekom, Riedl, Robbins, Ruud Koot, RyanGerbil10, Sabik, Sanxiyn, SchreiberBike, ShelfSkewed, Silly rabbit, Spl, Spoon!, StuartBrady, Stux, Thejoshwolfe, ThomasTenCate, Thore Husfeldt, Vadmiun, Vield, Volkan YAZICI, Wayiran, Woohookitty, Zigger, Ziothefifth, Zvar, 126 anonymous edits

**Map** *Source:* <http://en.wikipedia.org/w/index.php?oldid=500496278> *Contributors:* Agorf, Ajo Mama, Alansohn, Altenmann, Alvin-cs, AmiDaniel, Andreas Kaufmann, Anna Lincoln, Antonielly, AvramYU, B4hand, Bart Massey, Bartledan, Bevo, Bluemoose, Bobo192, Boothy443, Bosmon, Brianiac, Brianski, Catbar, Cfallin, Chaos5023, CheesyPuffs144, Comet--berkeley, Countchoc, CultureDrone, Cybercobra, Damian Yerrick, David Epstein, DavidCary, DavidDouthitt, Davidwhite544, Dcoetzee, Dcsaba70, Debresser, Declyte, Deineka, Dggoldst, Doc aberdeen, Doug Bell, Dreftymac, Dysprosia, EdC, Edward, Efadae, Ericamick, EvanED, Floatingdecimal, Fredrik, Frostus, Fubar Obfusco, George100, Graue, Hashar, Hirzel, Hugo-cs, Int19h, Inter, Irishjugg, JForget, JLaTondre, James b crocker, JannuBl22t, Jdh30, Jeff02, Jerryobject, Jesdisciple, Jleedev, Jokes Free4Me, JonathanCross, Jorge Stolfi, Jpo, Karol Langner, Kdau, Kglavin, KnowledgeOfSelf, Koavf, Krischik, Kusunose, Kwamikagami, LeeHunter, Macrakis, Maerk, Malbrain, Marcos canbeiro, Margin1522, Maslin, Maury Markowitz, Michael Hardy, Mindmatrix, Minesweeper, Minghong, Mintleaf, Mirzabab, Mithrasgregoriae, MrSteve, Mt, Neil Schipper, Neile, Nemo20000, Nick Levine, Noldoaran, ObsidianOrder, Oddity-, Orbnauticus, Orderud, PP Jewel, Paddy3118, PanagosTheOther, Patrick, Paul Ebermann, Pcap, Pfast, Pfunk42, Pgr94, PhilLho, Pimlottc, Pne, Radagast83, RainbowOfLight, RevRagnarok, RexNL, Robert Merkel, Ruakh, RzR, Sae1962, Sam Pointon, Samuelsen, Scandum, Shellreef, Signalhead, Silvonen, Sligocki, Spoon!, Swmcd, TShilo12, TheDoctor10, Tobe2199, Tobias Bergemann, TommyG, Tony Sidaway, Tushar858, TuukkaH, Vegard, Wlievens, Wmbolle, Wolfkeeper, Yurik, Zven, 249 anonymous edits

**Bidirectional map** *Source:* <http://en.wikipedia.org/w/index.php?oldid=504901327> *Contributors:* Andreas Kaufmann, Cgdecker, Cobi, Cyc, Efadae, GregorB, Mahdavi110, Mattbierner, 3 anonymous edits

**Multimap** *Source:* <http://en.wikipedia.org/w/index.php?oldid=502608982> *Contributors:* Aminorex, Andreas Kaufmann, Bantman, BluePyth, Bluebusy, Ceklock, Cybercobra, Enochlau, Excirial, Fuhghettaboutit, GCarty, JakobVoss, JonathanWakely, Macha, Not enough names left, Ruud Koot, Spoon!, Svick, TheDJ, TuukkaH, Wmbolle, 13 anonymous edits

**Set** *Source:* <http://en.wikipedia.org/w/index.php?oldid=452925673> *Contributors:* Amniarix, Andreas Kaufmann, CBM, Casablanca2000in, Classicalecon, Cybercobra, Damian Yerrick, Davecrosby uk, Dcoetzee, Denispir, Dreadstar, EdH, Elaz85, Fredrik, Gracenotes, Hetar, Hosamaly, Hut 8.5, Incnis Mrsi, Irene31, Jorge Stolfi, Linforce, Loupeter, Lt basketball, Lvr, MartinPoulter, MegaHasher, Mintguy, Modster, Mxn, Otus, Oxymoron83, Patrick, Pcap, Pederjones, Pfunk42, Polaris408, Poonam7393, QmunkE, Quintaylor, RJFJR, Raise exception, Rhanekom, Ruud Koot, SoSaysChappy, Spoon!, TakuyaMurata, Twri, Tyamath, Umasoni30, Urhixidur, Vimalwatwani, Wikilolo, William Avery, 37 anonymous edits

**Tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=507711306> *Contributors:* 209.157.137.xxx, 2help, Abce2, Acdx, Adrianwn, Adubi Stephen, Afrozenator, Alfredo J. Herrera Lago, Alvestrand, Andrevan, Anonymous Dissident, Anoshak, Arthena, Aseld, BAXelord, Bernard François, BlckKnight, Bobo192, Boleslav Bobcik, Bryan Derksen, Btyner, Can't sleep, clown will eat me, Chraga, Cfallin, Chris857, Coastrolver1994, Conversion script, Corti, Creidieki, Cruccione, Cybercobra, DRAGON BOOSTER, David Eppstein, Dcandeto, Dcoetzee, Defza, DevastatorIIC, Digitalme, Dillard421, Dixtosa, Dwandelt, Dysprosia, Eh kia, Evercat, Extra999, Fabiogramas, Fabricationary, FleetCommand, Fredrik, Fresheneesz, FvdP, Garyzx, Giftlite, Gioto, Giraffedata, Glenn, Graham87, Grilxlaxl, Ham Pastrami, Happynomad, Hashbrowncipher, HiteshLinus, Hotlorp, INkubusse, Iamfscked, Iamtheudeus, Jagged 85, Jarble, Jfroeilich, Jim1138, Jitse Niesen, Jiy, Jokes Free4Me, Jorge Stolfi, JulesH, Julesd, Justin W Smith, Kaswini15, Kausikghatak, KellyCoinGuy, Kinetician, Kjetil r, Knutux, KoRnholio8, Kpeeters, Kragen, Kuru, Kurykh, Ladsgroup, Liao, Lisamh, Loisel, Lukevanin, MER-C, MagiMaster, Majorly, Marechal Ney, Mark Renier, Martnym, Mattisgoo, Mckaysalisbury, Mdd, Mdd4696, Mdebets, Mentifisto, Mercury1992, Michael A. White, Mike314113, Mindmatrix, Mmemex, Mnogo, Mrwojo, Myanw, N5iln, NHRHS2010, Nbarth, Nbhatala, Nempti, Night Gyr, Nikai, Nivix, Nsaa, Nuno Tavares, Ollydbg, Ovakim, Paul August, Perique des Palottes, Pgan002, Pi is 3.14159, Pingveno, Pradeepvaishy, Qwertyus, R. S. Shaw, Ramzysamman, RazorICE, Reetep, Rich Farmbrough, Ripper234, Robertbowerman, RoyBoy, Rp, Ruakh, Ruud Koot, RzR, Sara wiki, Sbowers3, Seaphoto, Seb, Shamus1331, Silly rabbit, SiobhanHansa, Skittleys, Skraz, SlowJog, Solbris, SteveWitham, Svick, T Long, Taemyr, Tarquin, The Anome, The Thing That Should Not Be, Thomas Björkan, Thunderboltz, Tide rolls, Tobias Bergemann, Tomhubbard, Tommy2010, Trusilver, Two Bananas, Twri, TyA, Utcursch, VKokielov, Vdm, Wavelength, Wenttomowameadow, WhyBeNormal, Wsloand, Wtmitchell, Wxidea, Xevior, Ykh Wong, 297 anonymous edits

**Array data structure** *Source:* <http://en.wikipedia.org/w/index.php?oldid=504504813> *Contributors:* 111008066it, 16@r, 209.157.137.xxx, A'bad group, AbstractClass, Ahyl1, Alfio, Alksentrs, Alksub, Andre Engels, Andreas Kaufmann, Anonymous Dissident, Anwar saadat, Apers0n, Army1987, Atanamir, Awbell, B4hand, Badanedwa, Bargomm, Beej71, Beetstra, Beland, Beliaivsky, BenFrantzDale, Berland, Betacommand, Bill37212, Blue520, Borgx, Brick Thrower, Btx40, Caltas, Cameltrader, Cgs, Chetan chopade, Chris glenne, Christian75, Conversion script, Corti, Courcelles, Cybercobra, DAGwyn, Danakil, Darkspots, DavidCary, Dcoetzee, Derek farn, Dmason, Don4of4, Dreftymac, Dysprosia, ESKog, EconoPhysicist, Ed Poor, Engelec, Fabartus, Footballfan190, Fredrik, Funandtrvl, Func, Fvw, G worroll, Garde, Gaydudes, George100, Gerbrant, Giftlite, Graham87, Graue, Grika, GwydionM, Heavyrain2408, Henry513414, Hide&Reason, Highegg, Icairns, Ieee andy, Immortal Wowbagger, Intgr, Ipsign, Jdelanoy, JLaTondre, JaK81600, Jackollie, Jandalhandler, Jeff3000, Jfmantis, Jh51681, Jimbryho, Jkl, Jleedev, Jlmerrill, Jogloran, John, Johnuniq, Jonathan Grynspan, Jorge Stolfi, Josh Cherry, JulesH, Kaldosh, Karol Langner, Kbdank71, Kbrose, Ketiltrout, Kimchi.sg, Krischik, Kukini, LAX, Lardarse, Laurențiu Dascălu, Liempt, Ligulem, Ling.Nut, Lockeownzj00, Lowellian, Macrakis, Massysett, Masterdriverz, Mattb90, Mearuso, Mdd, Merlinsorca, Mfaheem007, Mfb52, Michael Hardy, Mike Van Emmerik, Mikeblas, Mindmatrix, MisterSheik, Mr Adequate, Mrstonky, Muzadded, Mwtoews, Narayane, Neelix, Nicvaroce, Nixdorf, Norm, Orderud, Oxymoron83, Patrick, PhilLho, Piet Delpot, Poor Yorick, Princeatapi, Pseudomonas, Qutezuze, Quuxplusone, R000t, RTC, Rbj, Redacteur, ReyBrujo, Rgrig, Rich Farmbrough, Rilak, Roger Wellington-Oguri, Rossami, Ruud Koot, SPTWriter, Sagaciousuk, Sewing, Sharkface217, Simeon, Simoneau, SiobhanHansa, Skittleys, Slakr, SlogswEEP, Smremde, Spoon!, Squidionius, Ssd, Stephenb, Strangelv, Supertouch, Suruena, Svick, TakuyaMurata, Tamfang, Tauwasser, Thadius856, The Anome, The Thing That Should Not Be, The Utahraptor, Themanian, Thingg, Timneu22, Travelbird, Trevyn, Trojo, Tsja, TylerWilliamRoss, User A1, Visor, Waywardhorizons, Wbm1058, Wernher, Wws, Yamamoto Ichiro, ZeroOne, Zzedar, 317 anonymous edits

**Row-major order** *Source:* <http://en.wikipedia.org/w/index.php?oldid=501231237> *Contributors:* AaronWL, Arthena, Awaterl, BenFrantzDale, Czarkoff, Drf5n, JordiGH, Jungrung, Kiefer, Wolfowitz, Liberio, Mcl, Mikeblas, Mstahl, Nzroller, Pearle, Rhymeswthorange, Rubik-wuerfel, Sintaku, Splintercellguy, Stevenj, Tedernst, Tiroche, Trebb, Welsh, Woohookitty, ZMughal, Zawersh, 58 anonymous edits

**Dope vector** *Source:* <http://en.wikipedia.org/w/index.php?oldid=491138659> *Contributors:* Amalas, Andreas Kaufmann, Dougher, Finn-Zoltan, Gennaro Prota, Gongshow, LivinInTheUSA, Phresnel, Sjordford, Stamos20, Wik, Xexeo, 6 anonymous edits

**Iliffe vector** *Source:* <http://en.wikipedia.org/w/index.php?oldid=497102572> *Contributors:* Alynna Kasmira, Andreas Kaufmann, Cmdrjameson, Cybercobra, Doradus, Feydey, Llinkade, Mgreenbe, Michael Hardy, NapoliRoma, Ospalh, Pgan002, RainbowCrane, Tobias Bergemann, Wotnarg, 12 anonymous edits

**Dynamic array** *Source:* <http://en.wikipedia.org/w/index.php?oldid=508170333> *Contributors:* Aekton, Alex.vatchenko, Andreas Kaufmann, Beetstra, Cobi, Ctxppc, Cybercobra, Damian Yerrick, David Eppstein, Dcoetzee, Declyte, Did293, Dpm64, Edward, Forbsey, Fresheneesz, Furrykef, Garyzx, Green caterpillar, Icep, Ixfid64, Jorge Stolfi, Karol Langner, MegaHasher, MisterSheik, Moxfyre, Mutinus, Octahedron80, Patmorin, Phoe6, Ryk, Rönin, SPTWriter, Simonykill, Spoon!, Wavelength, Wdscxsj, Wikilolo, WillNess, Wrp103, ZeroOne, 虞海, 44 anonymous edits

**Hashed array tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=467096408> *Contributors:* Cobi, Dcoetzee, Garyzx, MegaHasher, Queenmomcat, RainbowCrane, Surachit, WillNess

**Gap buffer** *Source:* <http://en.wikipedia.org/w/index.php?oldid=486860006> *Contributors:* Alf, Andreas Kaufmann, Aniketpate, Ansumang, Charles Matthews, Chronodm, Cybercobra, Damian Yerrick, Dennis714, Dpm64, Fyrael, Hitendrashukla, Hosamaly, Hydrostatic, J04n, Jaberwocky6669, Jacobslusser, Jogloran, LittleDan, MER-C, MisterSheik, Niceguyedc, Pgan002, Populus, Tobias Bergemann, 9 anonymous edits

**Circular buffer** *Source:* <http://en.wikipedia.org/w/index.php?oldid=498360883> *Contributors:* Amikake3, Andreas Kaufmann, Anonymi, Asimsalam, Astronouth7303, BastianVenthur, Bloodust, Calliopejen1, Churnett, Chocolateboy, Chris the speller, ChrisCPearson, Cybercobra, DrZoomEN, Eight40, Exfuent, Foobaz, Headbomb, Hoo man, Hosamaly, Jennavecia, Joeyadams, Julesd, KiloByte, Lucius Annaeus Seneca, Malcohol, Marokwitz, Mayukh iitbomby 2008, Mhi, Mike65535, MrOllie, Msoltyspl, Ohnoitsjamie, OlivierEM, OrlinKolev, Para15000, Parthashome, Paulitex, Pok148, Rhanekom, Rrelf, Serkan Kenar, Shabble, Shengliangsong, Shervinemami, Shirik, Siegelord, Silly rabbit, Strategist333, Sysabod, Tennenrshin, Tobias Bergemann, WolfWings, Ybungalobill, Zoxc, 醜い女, 99 anonymous edits

**Sparse array** *Source:* <http://en.wikipedia.org/w/index.php?oldid=509105360> *Contributors:* 16@r, BPositive, Cadillac, Charles Matthews, Engelec, FeatherPluma, Hawk777, Hgrangqvist, Incnis Mrsi, Kalyani Kulkarni, Karol Langner, Kolrok, KurtRaschke, Megha92, MisterSheik, Monedula, Pako, RainbowCrane, Rjwilmsi, Timrollpickering, 18 anonymous edits

**Bit array** *Source:* <http://en.wikipedia.org/w/index.php?oldid=508952084> *Contributors:* AJim, Andreas Kaufmann, Archimerge, Bluebusy, Bluemoose, Boud, Bovbl, Bvds, CRGreathouse, Chip123456, Chris the speller, Cobi, Cocciasik, DanielLemire, Davnor, Dcoetzee, Doug Bell, Furrykef, Glen Pepicelli, Hack-Man, Intgr, JLaTondre, Jacobolus, JesseW, Karol Langner, Kubanczyk, Minakshinajardhane, Neelix, Notinasnoid, Onomou, Orderud, Paul August, Pcoedes, Peter Flass, Pnm, Psychless, R'nB, R. S. Shaw, Rgrig, RomainThibaux, Sam Hovecar, Skwa, Spoon!, StuartBrady, Thajdog10, TheChrisD, Themanian, Tide rolls, TomJF, Vadmiun, 48 anonymous edits



**Bitboard** *Source:* <http://en.wikipedia.org/w/index.php?oldid=493068923> *Contributors:* Andreas Kaufmann, Avalon, Bigtimepeace, Brainix, Bubba73, CharlesGillingham, DGJM, Damian Yerrick, Daniel.Cardenas, Danielmachin, Ddx, Dissident, Doug Bell, Dr.Szlachedzki, Dhweller, Dylan Lake, Epolk, Furrykef, Glen Pepicelli, Hiram 99, IanOsgood, Jleedev, Kaimiddleton, Kenneth Cochran, LokaClock, M412k, MER-C, MRPProgrammer, Marudubshinki, Miko3k, Mynamesnotpj, Nibuod, Nippashish, Notheruser, Nv8200p, Pawnkingshree, Pearle, Pjoe, PrometheusX303, Psu256, QTCaptain, Quackor, RJFJR, Rkalyankumar, RoyBoy, Slamb, Srieffer, Sun Creator, Tromp, Trovatore, Updatehelper, WiiWillieWiki, Will Beback Auto, Zargulon, ZeroOne, 103 anonymous edits

**Parallel array** *Source:* <http://en.wikipedia.org/w/index.php?oldid=489732575> *Contributors:* AlanBarrett, Alik Kirillovich, Andreas Kaufmann, Charles Matthews, Dcoetzee, Dgpop, Garyzx, GregorB, Ironholds, Karol Langner, MisterSheik, Peter bertok, Ragzouken, TakuyaMurata, TheProgrammer, Thorwald, Veledan, 12 anonymous edits

**Lookup table** *Source:* <http://en.wikipedia.org/w/index.php?oldid=501057580> *Contributors:* Andreas Kaufmann, Angela, Armandas j, Arteitle, Berland, CRGreathouse, Cal-linux, Charles Matthews, Cheese Sandwich, Chris the speller, Cybercobra, David@sickmiller.com, Davidmaxwaterman, Dcoetzee, Decora, Dylan Lake, Dysprosia, Epok, Falcor84, Fingew, Fredrik, Gazoot, Gennaro Prota, Giftlite, Girolamo Savonarola, Graemec2, Headbomb, JLaTondre, Jelsova, Jengelh, KSmrq, Kdakin, Kencf0618, Khalid hassani, Kostmo, Kyng, Melesse, Michael Hardy, Mortense, MrDomino, MrOllie, Nbarth, NicM, Ohnoitsjamie, Omeatron, Patrick, Peter S., Petter.kallstrom, Pnm, Poco a poco, Qef, Ralian, Sadi, Sterrys, SteveBaker, Supernoo, Welsh, Woohookitty, Yellowdesk, Yflicker, ZAB, 114 anonymous edits

**Linked list** *Source:* <http://en.wikipedia.org/w/index.php?oldid=508946751> *Contributors:* 12.234.213.xxx, 16@r, 65.91.64.xxx, 75th Trombone, ATren, Achurch, Adamking, Adrian, Aeon, Afromayun, Agrammenos, Ahyl, Albertus Aditya, Amog, Anandvachhani, Andreas Kaufmann, Angela, Antaeus Feldspar, Apeculiaz, Apollo2991, Arakunem, Arivne, Arjun024, Arneth, Arthena, Astronautics, Avoided, BD2412, BL, BZRatfink, Beejaye, Beetstra, BenFrantzDale, BigDunc, Binaryedit, BkckNght, Bluezy, Bobo192, Borgx, Bughunter2, Carmichael, Celticeric, Cesarsorn, Chris the speller, Christian \*Yendi\* Severin, Clayhalliwell, Cload81, Cobi, ColdFusion650, Colonies Chris, Computersagar, Constructive editor, Conversion script, Corti, Creidicki, CryptoDerk, Cybercobra, Danarmstrong, Daniel Mietchen, Daniel5Ko, Dark knight, Darklilac, David Eppstein, Dcoetzee, Deeptrivia, Dekai Wu, Deshrag, Dillesca, Dixie91, Docboat, Dreslough, Elf, Eniagrom, Enochlaur, Etienne Lehnart, Fawcett5, Ferdinand Pienaar, FlashSheridan, Fredrik, Freqsh0, Fubar Obfusco, G7mcluvn, Garyzx, Giftlite, Giraffedata, Grafen, Graham87, Haeleth, Ham Pastrami, HappyInGeneral, Hariva, Headbomb, Hook43113, Hu12, IE, IceKarma, IgushevEdward, Ijsf, Integr, Iridescent, John7r0n, JFreeman, Jan.Smolik, Japo, Jarsyl, Javawizard, Jengelh, Jfmantis, Jim1138, Jin, JohnnyDog, JonHarder, Jorge Stolfi, Justin W Smith, Karingo, Karl-Henner, Katieh5584, Kbrose, Keilana, Kelmar, Kenny Moens, Kenny sh, Kenyon, Kku, Klilidiplomus, Koweja, KralSS, Kurniasan, Larryisgood, Lasindi, Leafboat, Levin, Lindsayfilmour, Lord Pistachio, MER-C, MaGa, Magister Mathematicae, Mantipula, Marc van Leeuwen, Marek69, Mark Renier, Materialscentist, MatrixFrog, Mblumber, Mboverload, Mc6809e, Mdd, MegaHasher, Meiskam, Meshari alnaim, Mild Bill Hiccup, Millermk, Mindmatrix, Minesweeper, Minimax, Minna Sora no Shita, MisterSheik, Miyum, Moggie2002, MrOllie, Mtsaic, Mwe 001, Mygerardromance, NHSKR, NJZombie, Nasty psycho, Neil916, Neilc, NewName, Nick Number, Nickj, Nixdorf, Noah Saltzman, Patmorin, Paul Ebermann, PavelY, Pengo, Perique des Palottes, Peter Karlsen, PhilKnight, Philg88, Pi is 3.14159, Pluke, Poccl, Prari, Quantran202, Quantumobserver, Quentin malmz, R. S. Shaw, RJFJR, Raghaven, Ravak, Raven4x4x, Redmarkvialnist, RichF, RickK, Rmosler2100, Rnt20, Ruud Koot, Rönin, SPTWriter, Sam Hoecevar, Sam Pointon, Samber, SaurabhKB, Shadowninja, Shanes, Silvonen, Simply.ari1201, Singlaive, SiobhanHansa, Smack, Smiler jerg, Sonett72, Sophus Bie, Spoon!, Sprocter, Ste4k, Stevage, Stevertigo, StuartBrady, Stw, Supadawg, Ta bu shi da you, Tabletop, TakuyaMurata, TalhalrfanKhan, Tango4567, TechTony, TedPostol, Tetracube, Tgeaim, Thadius856, The Thing That Should Not Be, TheCoffee, Theodore Kloba, Thorncrag, Thue, Tiddly Tom, TimBentley, Tobias Bergemann, TobiasPersson, Tomaxer, Traroth, Trusilver, Tyomitch, Unixxx, Upnishad, Uriyan, VTBassMatt, Versageek, Vhcomptech, Vipinhari, Vrenator, Waltercruz, Wereon, Wikilolo, WillNess, Willemo, Wizmo, Wj12, Wolfrock, Wolkykim, Wrp103, Wtmitchell, Ww, Xenochria, Yacht, Zvn, 699 anonymous edits

**XOR linked list** *Source:* <http://en.wikipedia.org/w/index.php?oldid=502557257> *Contributors:* Ais523, Amniarix, Andreas Kaufmann, Cacycle, Chris the speller, Cjoev, Cybercobra, Cyp, Damian Yerrick, Dav4is, Dcoetzee, Ebehn, Equor, Evil saltine, Haeleth, Itai, J04n, Jfmantis, Jketola, Jon Awbrey, Lusum, Nick, Nixdorf, Pakaran, Paul August, Quotemstr, Quuxplusone, Sander123, Siroxo, Sleske, Tomaxer, Vadmiun, Zoicon5, 26 anonymous edits

**Unrolled linked list** *Source:* <http://en.wikipedia.org/w/index.php?oldid=501842130> *Contributors:* Afrikaner, Andreas Kaufmann, Chris the speller, Costaluz, Damian Yerrick, DavidCary, Dcoetzee, Decora, Fx2, Gampuzampe, Garyzx, Headbomb, Jfmantis, Johnbod, Kcsomisetty, Kdau, MattGiuca, Neilc, Potatoswatter, Ryszard Szopa, Shigeru23, Warren, WillNess, Yuriybrisk, 9 anonymous edits

**VList** *Source:* <http://en.wikipedia.org/w/index.php?oldid=492131151> *Contributors:* Andreas Kaufmann, Cdiggins, Chris the speller, Cybercobra, Dawdler, Dcoetzee, Fredrik, GnuCivodul, Irrbloss, Jbolden1517, Ken Hirsch, MegaHasher, Mrwojo, Qwertys, Ripounet, Sandstein, WriterHound, 16 anonymous edits

**Skip list** *Source:* <http://en.wikipedia.org/w/index.php?oldid=507738124> *Contributors:* Alan Dawrst, Almkglor, Altenmann, Andreas Kaufmann, Antaeus Feldspar, Bovineone, Braddunbar, CRGreathouse, Carslort, Cerebllo, Chadmdaniel, Charles Matthews, Cybercobra, Dcoetzee, Devynici, Doradus, Dysprosia, Fredrik, Gene91, Integr, JaGa, Jim10701, Jorge Stolfi, Jrockway, Kukolar, Laurieenne Bell, Menahem.fuchs, MinorContributor, Mrwojo, Musically ut, Nkour, Noldoaran, Nsfmc, Overred, Patmorin, Pet3ris, Piet Delpont, Populus, R. S. Shaw, Rdhettinger, Rhanekom, Rpk512, Sanchom, Silly rabbit, Stevenj, Svick, Viebel, Vishalvishnoi, Waffleguy4, Wavelength, Wojciech mula, Xcez-be, Zr2d2, 152 anonymous edits

**Self-organizing list** *Source:* <http://en.wikipedia.org/w/index.php?oldid=505682644> *Contributors:* Aclwon, Andreas Kaufmann, Ansumang, AshLin, Bearcat, Headbomb, JonHarder, Khazar, Lilwik, PigFlu Oink, Roshmorris, SaurabhKB, Utcursch, VTBassMatt, 8 anonymous edits

**Binary tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=508621340> *Contributors:* 15turnsm, 7, ABCD, Aarsalankhalid, Abu adam, Adam majewski, Ahoerstemeier, Ahyl, Airplaneman, Alex.vatchenko, Alienus, Allefant, Altenmann, Andreas Kaufmann, AndrewKepert, AnotherPerson 2, Antaeus Feldspar, Aroundthewayboy, Ateeq.s, B4hand, Behrangsa, Beland, Belovedeagle, Bender2k14, Bhadani, BigDunc, Bkell, BkckNght, Bluebusy, Bobo192, Bojan1989, Bonada, Brentsmith101, Brianga, BruceLee, Caltas, Calvin zxc, Card Zero, Chraga, Cdiggins, Charles Matthews, Chicodroid, Chris the speller, Chris857, Ck lostsword, Classicaecon, Cncplyr, Coelacan, Conversion script, Cybercobra, Cyhawk, Czar.pino, DMacks, Darangho, David Eppstein, David Shay, David-Sarah Hopwood, Dawynn, Dcoetzee, Dguido, Djcollom, Dkasak, Dominus, Dontoito, Doriftu, Doug s, Dr. Sunglasses, Drano, DuaneBailey, Duoduoduo, Dysprosia, Ekeb, Encognito, Ferkelparade, Frankel444, Frozendice, FvdP, Garyzx, Gdevanla, Giftlite, Gilliam, Gimmetrow, Gsmodi, Happyuk, Hazmat2, Heirpixel, ISTB351, IanS1967, Ilana, Itchy Archibald, Itman, JabberWok, Jafet, Jerome Charles Potts, Jerryobject, JimsMaher, John Quincy Adding Machine, Jonfore, Josephskeller, Jruderman, Julesd, Justin W Smith, Kamath.nakul, Kamirao, KaragouniS, Kholino, Kgashek, Kgaucorn28, Klilidiplomus, Kuru, LC, LandruBek, Liao, Liftarn, LightingDragon, Linas, LithiumBreather, Loisel, LokaClock, Loolo, Loopwhile1, Lotje, MONGO, Mahahahaneapneap, Malleus Fatuorum, Marc van Leeuwen, Mark Renier, Martinp23, Materialscentist, MathijsM, Matir, Maurice Carbonaro, Mbdeir, Mclid, Mdnahas, Metricopolus, Mhayes46, Michael Angelkovich, Michael Hardy, Michael Slone, Microbizz, Mike Christie, Minesweeper, Mjm1964, Mrwojo, Neomagic100, Nippoo, Noldoaran, Nonexistent, Oblivious, Oli Filth, Ontariolot, Opelio, Orphic, Otterdam, ParticleMan, Petr, Pkg, Philip Trueman, Pih, Pohl, Pp007, Ppelteli, RG2, RadioFan, Rege, Reinderien, Rhanekom, Rich Farmbrough, Rohitgoyal100, Roreuqnoc, Roystrow, Rspcer, Rzelnik, Rönin, SGBailey, Sapeur, Shentino, ShinjiXman, Shmomuffin, Shoujun, Silver hr, Simeon, Smallpond, SmartGuy Old, Someone else, SpaceFlight89, Spottedowl, Sss41, Stickee, Sun Creator, Taemyr, TakuyaMurata, Tarquin, Tarrahaticas, Tdsmith, The Thing That Should Not Be, Thrapper, Vegpuff, Waggars, WillNess, Wtarreau, Waelgæst wæfre, XJamRastafire, Xevior, Ynhockey, Yuubimbako, Zero sharp, Zetawoof, Zipdisc, Zvn, 445 anonymous edits

**Binary search tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=506477311> *Contributors:* 2620:0:1000:1B01:F08A:D18F:B5D4:3D36, 4get, Abednigo, Abu adam, Adamuu, AgentSnoop, Ahyl1, Alansohn, Alexsh, Allan McInnes, Andreas Kaufmann, Anoopjohnson, Avermapub, Awu, BAxelrod, BPositive, Banaticus, Beetstra, Bernard François, Bkell, Booyabazooka, Bryan Derksen, Butros, Calbaer, Capricorn42, Casted, Chery, Cochito, Conversion script, Cybercobra, D6, Damian Yerrick, Danadocus, Dcoetzee, Devastator1IC, Dicklyon, Dimchord, Djcollom, Doctordiehard, Doradus, Dysprosia, Dzikasosna, Ecb29, Enochlaur, Evil Monkey, Ezhiki, Frankro444, Fredrik, Func, GRHhooked, Gaius Cornelius, Garoth, Giftlite, Glenn, Googl, Gorffy, GregorB, Grunt, Hadal, Ham Pastrami, Hathawayc, Havardk, Hu12, Hyad, IgushevEdward, Ilana, Ivan Kuckir, Ixfd64, JForget, James pic, Jdm64, Jdurham6, Jerryobject, Jin, Jms49, Jorgers, Karl-Henner, Kate, Kewlito, Konnetikut, Konstantin Pest, Krage, Kulp, LOL, Lanov, Liao, LilHelpa, LittleDan, Loren.wilton, Madhan virgo, Matekm, MatrixFrog, Maximamax, Maximus Rex, Mb1000, MclareN212, MegaHasher, Metalmax, Mgius, Michael Hardy, Michael Slone, Mikeputnam, Mindmatrix, Minesweeper, MladenWiki, MrOllie, MrSomeone, Mrwojo, Mweber, Nakarumaka, Nerdgerl, Neurodivergent, Nils schmidt hamburg, Nixdorf, Nneonneo, Nomen4Omen, Nux, Ohnoitsjamie, Oleg Alexandrov, Oli Filth, Oliphant, One half 3544, Oni Lukos, Onomou, Ontariolot, Oskar Sigvardsson, Pönc, Phil Boswell, PhilipMW, Phishman3579, Pion, Postdlf, Qiq, Qleem, Quuxplusone, Qwertys, RJK1984, Rdemar, Regnar, Rhanekom, Richardj311, RoySmith, Rudo.Thomas, Ruud Koot, S3000, SPTWriter, Salrizvy, Shen, Shmomuffin, Sketch-The-Fox, Skier Dude, Smallman12q, Solsan88, Spadgos, Spiff, Sss41, SteveAyre, Swapsy, Taejo, Taw, Tbvdm, The Parting Glass, TheMandarin, Theone256, Thesevenses, Theta4, Timwi, Tomt22, TrainUnderwater, Trevor Andersen, VKokielov, Vdm, Vectorpaladin13, Vocaro, Vromascan, Wavelength, WikHead, WikiWizard, WillNess, Wmayer, Wtarreau, Wtmitchell, X1024, Xevior, Yaderbh, Your Lord and Master, ZeroOne, ماني, 336 anonymous edits

**Self-balancing binary search tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=500482193> *Contributors:* ABCD, ActuallyRationalThinker, Andreas Kaufmann, Andreas4965, Angela, Baluba, Bluebusy, CRGreathouse, Chochopk, Cybercobra, DJ Clayworth, David Eppstein, Dcoetzee, Diberri, Drilnoth, Dysprosia, Enochlaur, Fredrik, Integr, Jacob grace, Jacobolus, Jafet, Japanese Searobin, Jeltz, Jon Awbrey, Jorge Stolfi, Jruderman, Kdau, Lamro, Larkinzhang1993, Larrycz, Light current, Michael Hardy, Momet, Moskvax, MrDrBob, Neilc, Noldoaran, Pgan002, Plastikspork, Plyd, Qwertys, RJFJR, Ripe, Shlomif, Srganeshs, Steaphan Greene, SteveJochen, Wideofox, Wightkeeper, 42 anonymous edits

**Tree rotation** *Source:* <http://en.wikipedia.org/w/index.php?oldid=497402149> *Contributors:* Abarry, Adamuu, Altenmann, Andreas Kaufmann, B4hand, BkckNght, Boykbock, Castorvx, Chub, Conversion script, David Eppstein, Dcoetzee, Dysprosia, Förbidik, Headbomb, Hyperionred, Joriki, Kjkjava, Knowledgeofthekrell, Leonard G., LilHelpa, Mav, Michael Devore, Michael Hardy, Mr Bound, Mtanti, Neilc, Oleg Alexandrov, Pako, Qwertys, Ramasamy, SCriBu, Salxvar, Skaraoke, Svick, Trainra, Vegasprof, Waylonflinn, Whitepaw, Wizzar, Woblosch, Xevior, ماني, 38 anonymous edits

**Weight-balanced tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=435813754> *Contributors:* Andreas Kaufmann, Bhadani, Bobmath, Ciphergoth, DVD R W, Etrigan, JaGa, Jointds, KenE, Norm mit, Pascal.Tesson, Snarius, The former 134.250.72.176, Zureks, 11 anonymous edits

**Threaded binary tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=496659875> *Contributors:* A3 nm, Adam murgitroyd, Amicon, Andreas Kaufmann, Ansumang, Atharvaborkar, Brad101, Edlee, Grafen, Headbomb, Hut 8.5, Jaccos, Jagat sastry, John of Reading, Konstantin Veretennicov, LokiClock, MER-C, Michael Hardy, Moberg, MoraSique, Mr2001, Pearle, Ppelteti, R. S. Shaw, Razgouken, Sanju2193p, Sdenn, Sunnygunner, Svick, Themanía, Vaibhavvc1092, Yeskw, 24 anonymous edits

**AVL tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=509654332> *Contributors:* Adamd1008, Adamuu, Aent, Agrawalyogesh, Akerbos, Alex Kapranoff, AlexGreat, Altenmann, Anant sogani, Andreas Kaufmann, Andrew Weintraub, Apanag, Astral, Auntof6, Avicennasis, Axe-Lander, BenBac, Benoit fraikin, Binnacle, Bkil, Blackllotus, Bluebusy, Byrial, Castorvx, Caviare, ChrisMP1, Codingreicipes, Compusense, Conversion script, CostinulAT, Cybercobra, Cyhawk, Daewollama, Damian Yerrick, Darangho, David.Mestel, Dawynn, Dcamp314, Dcoetzee, Denisarona, Dicklyon, DmitriyVilkov, Docboat, Doradus, Dtrebbin, Dysprosia, Eluether, Enviroboy, Epachamo, Euchiasmus, Fredrik, FvdP, G0gogcsc300, Gaius Cornelius, Geek84, Geoff55, Gnowor, Greenrd, Greg Tyler, Gruu, Gulliveig, Gurch, Gökhan, II MusLiM HyBRiD II, Iamnitin, Infrogmation, Intger, J.delanoy, Jac16888, Jeepday, Jeff02, Jennavecia, Jim1138, Jirka6, Jll, Joeyadams, KGasso, Kain2396, Kdau, Kenyon, Kingpin13, Kjkolb, Ksulli10, Kukolar, Kuru, Kushalbiswas777, LOL, Lankiveil, Larry V, Leszek Jaficzuk, Leuko, M, MarkHeily, Materialscientist, MattyIX, Mauritsmaartendejong, Mckaysalisbury, Mellerho, Merovingian, Michael Hardy, Michael M Clarke, Michael miceli, Mike Rosoft, Mikm, Minesweeper, Mjkoo, MladenWiki, Mnogo, Moberg, Mohammad ahad, Momet, Mr.Berna, Msanchez1978, Mtanti, Mzruya, NawlinWiki, Neile, Nguyễn Hữu Dung, Nixdorf, Nnemo, Noldoaran, Northamerica1000, Nysin, Obradovic Goran, Oleg Alexandrov, Oliversisson, Ommyi-Pangaeus, Orimosenzon, Paul D. Anderson, Pavel Vozenilek, Pedrito, Pgan002, Phishman3579, Pnorcks, Poor Yorick, RJFJR, Resper, Rockslave, Ruud Koot, ST47, Safety Cap, Sebculture, Seyen, Shlomif, Shmnomuffin, Smalljim, Srivesh, Ste4k, Tamfang, Tide rolls, Tobias Bergemann, Toby Douglass, Tphyahoo, Tsemii, Tsoft, UnwashedMeme, Uw.Antony, Vektor330, Vlad.c.manea, West.andrew.g, Xevior, Yksyksyks, Zian, 386 anonymous edits

**Red-black tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=441577312> *Contributors:* 203.37.81.xxx, 206.176.2.xxx, Abu adam, Adamuu, Ahoerstemeier, Ahy1, Akerbos, AlanUS, Altenmann, Andreas Kaufmann, Aplusbi, Awakenrz, Banej, Belfry, Belovedeagle, BenFrantzDale, Bezenek, Binnacle, Bioskope, Blackllotus, Blow, Bonadea, Bovineone, Brona, C. A. Russell, Cababunga, Card Zero, Caviare, Cburnett, Cjcollier, Connelly, Consed, Conversion script, Cybercobra, Cyp, David Epstein, Dcoetzee, Deepakabhyankar, Dfeuer, Dicklyon, Drak, Dreamshade, Drebs, Dysprosia, Dlugosz, Enochlau, Epr123, ErikHaugen, Fawcett5, Fcp2007, Fragglet, Fredrik, FvdP, Ghakko, Ghewgill, Giftlite, Gimbo13, Giraffedata, Glrx, Gnathan87, Graham87, Grandphuba, H2g2bob, Hao2lian, Hariva, Hawke666, Headbomb, Hermel, Hgkamath, Hnn79, Hugh Aguilar, Humpback, IDogbert, Iav, JMBucknall, Jameshfisher, Jaxl, JengelH, Jerome.Abela, Jleedev, Jodawi, Johnuiq, Joriki, Joshibschman, Jtsiomb, Jzcool, Karakak, Karl-Henner, Karlhendrikse, Kenyon, Khalil Sawant, Kmels, Kragen, Kukolar, Kyle Hardgrave, Laurier12, Leonard G., LiDaobing, Linuxrocks123, Loading, Lukax, Lunakeet, Madhurtanwani, MatrixFrog, Maxis fw, Mgrand, Michael Hardy, Mindmatrix, Minesweeper, MladenWiki, Mnogo, N3bulous, Nanobear, Narlami, Naxik, Nishantjr, Notheruser, OMouse, Ontariolot, Patmorin, Pgan002, Pgk, Phil Boswell, Phishman3579, Pnuessler, Potatoswatter, Pqrstuv, RJFJR, Regnaron, Ripper234, Rjwilmsi, Roleplayer, Rsrikanth05, Ruud Koot, Ryan Stone, SGreen, SLi, Schellhammer, Sc72, Sdenn, Sebbe, Sepreece, Sesquianual, Shizhao, Shyammurarka, SickTwist, Silly rabbit, Silverdirk, Simoncropp, Smangano, Smlindog2000, Sreeakashay, Stanislaw Nowak, Stdazi, Stone628, Storabled, Strebe, Supertigerman, Tbvdm, TheWarlock, Themanía, Tide rolls, Timwi, Tobias Bergemann, Toby Douglass, Tristero, Uncle Milty, Uniwalk, Urod, Userask, Versus22, Warut, Wikipedian to the max, Will2k, Wittjeff, Xevior, Xiteer, Xjianz, Xmarios, Yakov1122, Zehntor, Zetawoof, 351 anonymous edits

**AA tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=503960290> *Contributors:* AliasXYZ, Ash211, Balabiot, Bporopat, Charles Matthews, Confuzzled, Cybercobra, Damian Yerrick, Darguz Parsilvan, Dcoetzee, Firsfron, Floodyberry, FvdP, Gazpacho, Goochelaar, HyperQuantum, Jjungree, Koavf, Kukolar, MONGO, Mecki78, Mnogo, Mu Mind, Nroets, Optikos, Ppelteti, Qj0n, Qutezuze, RJFJR, Rkleckner, RobertG, Stanislav Nowak, Triona, Vladislav.kuzkokov, Vromascau, Why Not A Duck, 51 anonymous edits

**Scapegoat tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=507737977> *Contributors:* AbsolutBildung, AlecTaylor, Andreas Kaufmann, Aplusbi, Cybercobra, Danadocus, Dcoetzee, Firsfron, FvdP, Hankjo, Jarsyl, Joey Parrish, Kukolar, MegaHasher, Mnogo, Patmorin, Rich Farmbrough, Robert Ullmann, Ruakh, Sam Hovevar, Slike2, Stevelothen, Themanía, WillUther, Wknight94, 38 anonymous edits

**Splay tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=509809642> *Contributors:* 1exec1, Abu adam, Ahmad87, Ahy1, AlanUS, Andreas Kaufmann, Anna Lincoln, Apankrat, Aplusbi, Arunshankarb, Atavi, Axlape, B.pradeep143, Christopher Mahan, Coldzero1120, Const86, Conversion script, Crimer, Cybercobra, CyborgTosser, Dcoetzee, Dekart, Dicklyon, Dodno, Drz, Dysprosia, Dzikasosna, Dlugosz, Foober, Fredrik, Freeside3, Fresheneesz, FvdP, Giftlite, Gscshoyru, Gwern, HPRappaport, Hannan1212, HereToHelp, Honza Záruba, HueSatLum, Jameshfisher, Jamie King, Josh Guffin, Jwiliia3, KConWiki, Karl Dickman, Kukolar, Ligulem, Lqs, Martlau, Mav, MegaHasher, Michael Hardy, MladenWiki, Mnogo, Momet, Nanobear, Nixdorf, Octahedron80, Ohconfcucus, Ontariolot, P2004a, Pako, Phil Boswell, Phishman3579, Pnm, Qutezuze, Rhanekom, Rich Farmbrough, Roman Munich, Romanc19s, Russelljr, Safek, Shmnomuffin, Sihag.deepak, Silly rabbit, Silvonon, Snietfeld, Sss41, Stephan Schulz, SteveAyre, Tabletop, Terber, That Guy, From That Show!, Tjdw, Unyoyega, Urhixidur, VTBassMatt, Vector, Versus, Wavelength, Wiki.ajaygautam, Wiml, Winniehell, Wj32, Wolfkeeper, Ybungalobill, Yonkeltron, Zuphilih, 121 anonymous edits

**T-tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=500759577> *Contributors:* Acn, Andreas Kaufmann, Bezenek, Bryan Derksen, Charles Matthews, Cybercobra, Damian Yerrick, Dlugosz, Fuhghettaboutit, FvdP, Gaius Cornelius, GreatWhiteNortherner, Hholzgra, Hmendocna, Jamie Lokier, Jonah.harris, Kroepeke, Mnogo, Omnipaedia, Paxcoder, Psychonaut, Rich Farmbrough, Ted nw, The Thing That Should Not Be, WriterHound, 21 anonymous edits

**Rope** *Source:* <http://en.wikipedia.org/w/index.php?oldid=509808094> *Contributors:* Almkglor, Alvin Seville, Andreas Kaufmann, ArtemGr, AzraelUK, Chkno, Cybercobra, David-Sarah Hopwood, DavidCary, Dcoetzee, Dhruvbird, Doradus, Emersoni, Fridolin, Fwielst, Headbomb, Jacroe, Jokes Free4Me, Kimbly, Lfstevens, Meng Yao, Mfwitten, Mrwojo, Mstroec, Pengo, Preetum, Qwertyus, Snaxe920, Spoon!, Thumperward, Tobias Bergemann, Trivialesefault, Vdm, Wtanaka, Ybungalobill, 33 anonymous edits

**Top Trees** *Source:* <http://en.wikipedia.org/w/index.php?oldid=423759868> *Contributors:* Pgan002, Rjwilmsi, 3 anonymous edits

**Tango Trees** *Source:* <http://en.wikipedia.org/w/index.php?oldid=412748471> *Contributors:* Acroterion, AnonMoos, C.Fred, Card Zero, Chris the speller, Do not want, Giraffedata, Grafen, Headbomb, Inomyabcs, Iridescent, Jasper Deng, JengelH, Malcolmá, Mephistophelian, Nathan Johnson, Nick Number, Nyttend, Ontariolot, Philg88, Qwertyus, RHaworth, Rayhe, Rjwilmsi, Sfan00 IMG, Tango tree, Vector, S, 24 anonymous edits

**Van Emde Boas tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=502768425> *Contributors:* A5b, Adrianwn, Argav, B4hand, BACbKA, Brutaldeluxe, Charles Matthews, Cybercobra, Cyhawk, David Cooke, Dbenbenn, Dcoetzee, Doradus, ElhamKhodae, Fresheneesz, Fx4m, Gailcarmichael, Gulliveig, Jackrae, Jeff02, Kragen, Mangarah, Michael Hardy, Neelix, Nickj, Panarchy, Patmorin, Phil Boswell, Piet Delpont, Quuxplusone, Qwertyus, Rjwilmsi, Snoopy67, Svick, 20 anonymous edits

**Cartesian tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=470335355> *Contributors:* Andreas Kaufmann, Bbi5291, Cobi, Cybercobra, David Epstein, Giftlite, GrEp, Pichpich, Tabletop, 9 anonymous edits

**Treap** *Source:* <http://en.wikipedia.org/w/index.php?oldid=488355692> *Contributors:* AHMartin, Andreas Kaufmann, Arbor, Bencmq, Blaisorblade, Brutaldeluxe, C02134, Cdb273, Chris the speller, Cshinyee, Cybercobra, David Epstein, Edward, Eequor, Gustavb, Hans Adler, Hdante, ICEAGE, Itai, James.nvc, Jleedev, Jogloran, Jsaxton86, Justin W Smith, Jörg Knappen, KnowledgeOfSelf, Kukolar, MaxDel, MegaHasher, Milkmandan, Miyom, MoreNet, Naxik, Oleg Alexandrov, Pfh, Poor Yorick, Qef, RainbowCrane, Regnaron, Ruud Koot, Saccade, Wsload, 32 anonymous edits

**B-tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=508669371> *Contributors:* 128.139.197.xxx, ABCD, AaronSw, Abrech, Ahy1, Alansohn, Alaric, Alfalfahotshots, AlistairMcMillan, Altenmann, Altes, AlyM, Anakin101, Anders Kaseorg, Andreas Kaufmann, Andytigg, AnnaFrance, Antimatter15, Appoose, Aubrey Jaffer, Avono, BAxelrod, Battamer, Beeson, Betzaar, Bezenek, Bkell, Bladefix2, Bor4kip, Bovineone, Bryan Derksen, Btwied, CanadianLinuxUser, Carbuncle, Cbraga, Chadloder, Charles Matthews, Chmod007, Chris the speller, Ciphergoth, Ck lostsword, ContivityGoddess, Conversion script, Cp3149, Ctxppc, Curps, Cutter, Cybercobra, Daev, DanielKlein24, Dcoetzee, Decrease789, Dlæ, Dmn, Don4of4, Dpotter, Dravecky, Dysprosia, EEMIV, Ed g2s, Eddvella, Edward, Ethan, Fabriciodosanjosvilla, FatalError, Fgdafsdgfdsgfd, Flying Bishop, Fragglet, Fredrik, FreplySpang, Fresheneesz, FvdP, Gdr, Giftlite, Glrx, GoodPeriodGal, Ham Pastrami, Hao2lian, Hariva, Hbent, Headbomb, I do not exist, Inquisitus, Iohannes Animosus, JCLately, JWSchmidt, Jacosi, Jeff Wheeler, Jirka6, Jjdawson7, Joahnnes, Joe07734, John lindgren, John of Reading, Jorge Stolfi, JoshuSasori, Jy00912345, Kate, Ketil, Kinema, Kinu, Knutux, Kovianyó, Kpjäs, Kukolar, Lamdk, Lee J Haywood, Leibniz, Levin, Lfstevens, Loadmaster, Luna Santin, MIT Trekkie, MachineRebel, Makkuro, Malbrain, Mattoothman, Merit 07, Mhss, Michael Angelkovich, Michael Hardy, Mikeblas, Mindmatrix, Minesweeper, Mnogo, MoAgnome, MorgothX, Mrnaz, Mrwojo, NGPriest, Nayuki, Neile, Nishantjr, Noodlez84, Norm mit, Oldsharp, Oli Filth, P199, PKT, Patmorin, Pausali, Peter bertok, Pgan002, Phishman3579, Postrach, Priyank bolia, PrologFan, Psyphen, Ptheoch, Psychobbens, Quadrescence, Qutezuze, R. S. Shaw, RMcPhillip, Redrose64, Rich Farmbrough, Rp, Rpajares, Ruud Koot, Sandeep.a.v, Sandman@llgp.org, SickTwist, Simon04, SirSeal, Slady, Slike, Spiff, Ssbohio, Stephan Leclercq, Stevemidgley, Strake, Ta bu shi da yu, Taldean, Teles, The Fifth Horseman, Tjdw, Tobias Bergemann, TomYHChan, Trusilver, Tuolumne0, Twimoki, Uday, Uw.Antony, Verbal, Wantnot, WinampLlama, Wipe, Wkaiiley, Wolfkeeper, Wout.mertens, Wsload, Wtanaka, Wtmitchell, Yakushima, Zearin, 402 anonymous edits

**B+ tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=509035342> *Contributors:* Afaviram, Ahy1, Ajarov, Arny, Arrow, Arthana, Ash211, Bad Romance, Bluebusy, Boing! said Zebedee, Bornhj, Bovineone, Bryan Derksen, Capi, CesarB, Cherezov, Cmdrjameson, ColinTempler, CommonsDelinker, Cybercobra, Cychoi, Daniel1, Decrease789, Dmn, DomQ, Eddie595, Encyclops, Eurleif, Favianon, Foeckler, Fresheneesz, Garciaad5658, Giftlite, Giovanni Kock Bonetti, GregorB, Grundprinzip, Gurch, Happyvalley, Holy-foek, Hongooi, I do not exist, Igor Yalovecky, Imachuck, Immunize, Inking, Itmozart, James.walmsley, Jbalint, Josh Cherry, Jsnx, Julthep, Jwango1, Kastauyra, K14m, Knutties, LOL, Leibniz, Leksey, Leujohn, Lightst, LilHelpa, Lowercase Sigma, LrdChaos, Luc4, Lupo, MRLacey, Marc omorain, Matthew D Dillon, Max613, Mdmkolbe, Mfedyk, Mhss, MikeDierken, Mikeblas, Mogentianae, Mqchen, Mrcowden, Muro de Aguas, Mydimle, Nat2, Neile, Nishantjr, Nqzero, Nuworld, Nyenyec, Obradovic Goran, Oleg Alexandrov, OliviaGuest, Penumbra2000, PhilippWeissenbacher, Pinethicket, Pnm, Priyank bolia, RHaworth, Raatikka, Reaper Eternal, S.Örvarr.S, SchreiberBike, Scrool, SheffieldSteel, Snarius, Ste4k, TZOTZIOY, TheProject, Thunderpenguin, Tim Starling, Tlesher, Tommy2010, Tresiden, TutterMouse, Tuxthepenguin933, Txomin, Ubuntu2, UncleDouggy, Unixguy, Vevek, Vikreykja, WhiteDragon, WorldsApart, Wout.mertens, Yamla, Yellowstone6, Ysoroka, Yungoe,



**Suffix array** *Source:* <http://en.wikipedia.org/w/index.php?oldid=502250705> *Contributors:* Andreas Kaufmann, Arnabdotorg, Beland, BenRG, Bkil, Chris83, Cobi, EchoBlaze94, Edward, Gailcarmichael, Gaius Cornelius, Garyzx, Giftlite, Headbomb, JoaquinFerrero, Jwarhol, Karol Langner, Kiwibird, Malbrain, Mboverload, MeltBanana, Mjordan, Nils Grimsmo, Nroets, Ruud Koot, Singleheart, Tobias Bergemann, TripleF, Viksit, Wolfgang-gerlach, 60 anonymous edits

**Compressed suffix array** *Source:* <http://en.wikipedia.org/w/index.php?oldid=496253788> *Contributors:* Andreas Kaufmann, G.ottaviano, Headbomb, RHaworth, Ruud Koot, Stringologist, 1 anonymous edits

**FM-index** *Source:* <http://en.wikipedia.org/w/index.php?oldid=506427734> *Contributors:* Andreas Kaufmann, Cventus, Fingerz, G.ottaviano, Mild Bill Hiccup, Noodlez84, Omnipaedista, Orenburg1, Raven1977, Ruud Koot, SDC, Sun Creator, Vyahhi, 4 anonymous edits

**Generalised suffix tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=502141455> *Contributors:* Aditya, Andreas Kaufmann, Bbi5291, Bernhard Bauer, Dcoetzee, Giftlite, JanCeuleers, Joey-das-WBF, Malik Shabazz, Michael Hardy, Nils Grimsmo, Ninjagecko, Ruud Koot, Wsloand, 18 anonymous edits

**B-trie** *Source:* <http://en.wikipedia.org/w/index.php?oldid=492663969> *Contributors:* 1ForTheMoney, Andreas Kaufmann, Debacker1, Fram, Headbomb, Heteri, Merosonox, MindstormsKid, R'n'B, 1 anonymous edits

**Judy array** *Source:* <http://en.wikipedia.org/w/index.php?oldid=491132454> *Contributors:* Alynna Kasmira, Andreas Kaufmann, B.suhasini, C. A. Russell, ChrisGualtieri, Cowgod14, Cybercobra, Danmoberly, Dcoetzee, Djmckee1, Dlohrer2003, Doug Bell, EIFY, Ejrh, Fresheneesz, Furrykef, Garyzx, Gmaxwell, JudayArry, Malbrain, Mellery, Minghong, Nave.notnilc, Nigosh, Pnm, RainbowCrane, Rolingyid, The Anome, Uberdude85, Voomoo, XHire, 16 anonymous edits

**Directed acyclic word graph** *Source:* <http://en.wikipedia.org/w/index.php?oldid=507502997> *Contributors:* Andreas Kaufmann, Archie172, Balrog-kun, BiT, Bkonrad, Bo Lindbergh, Bokeh.Sensei, Chimz, Chkno, Damian Yerrick, David Eppstein, Gwernol, Headbomb, JonHarder, MC10, Mandarax, Marasmusine, Nightmaare, Norman Ramsey, Ort43v, Phil Boswell, Radagast83, RI, Rofl, Sadads, Smanov, Smyth, Thelb4, Watcher, 29 anonymous edits

**Ternary search tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=500059669> *Contributors:* 2607:F8F0:C10:FFF:200:5EFE:8E68:442B, Abdsahin, Antaeus Feldspar, Arkanosis, Awnedion, Booyabazooka, Cwolfsheep, Cybercobra, Fxsjy, Jds, Klapautius, Knowtheory, Magnus, Magioladitis, Mayank.kulkarni, Patrick O'Jackson, Potatoswatter, Pyroflame 91, Raanoo, Rangilo Gujarati, Richardj311, Selket, Shadypalm88, SkyWalker, Tedickey, Twimoki, Wavelength, Xuul, Zureks, 73 anonymous edits

**And-or tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=495982894> *Contributors:* Addshore, Allanrbo, AxelBoldt, CBM, Classicalecon, Fratrep, Garion96, H taammoli, Hooperbloob, Integr, Jamjam337, Jpbowen, Justmeherenow, Logperson, Mnp, Nfwu, Pete.Hurd, Routa-olio, Rwww, Ryan Roos, Saumaun, 3 anonymous edits

**(a,b)-tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=499238592> *Contributors:* AlkyAlky, Bovineone, Cacadril, Centrx, Cybercobra, Gardar Rurak, JonHarder, Linas, Lockley, MER-C, Nick, Onodevo, Sapeur, Skapur, Skysmith, Yxejamir, 4 anonymous edits

**Link/cut tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=489186560> *Contributors:* Cybercobra, FactorialG, Giftlite, GregorB, Martlau, Pgan002, Rkleckner, 6 anonymous edits

**SPQR tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=502004709> *Contributors:* Andreas Kaufmann, Auntof6, Cybercobra, David Eppstein, Eng 104\*, Harrigan, Headbomb, Kakila, Khazar2, Kintaro, LilHelpa, MaxEnt, Riedel, Rjwilmsi, Twri, 5 anonymous edits

**Spaghetti stack** *Source:* <http://en.wikipedia.org/w/index.php?oldid=490938485> *Contributors:* Alf, Antandrus, Congruence, Cybercobra, Dcoetzee, Dreftymac, Edward, EmptyString, Eric119, Eyal0, Furrykef, Grenavitar, It's-is-not-a-genitive, Jerryobject, Karada, Lindsey Kuper, Masharabinovich, Pichpich, Ripounet, Ruakh, Ruud Koot, Sean r lynch, Skaller, Somercet, Tagishsimon, The Anome, TimBentley, Timwi, Tyc20, Woohookitty, 18 anonymous edits

**Disjoint-set data structure** *Source:* <http://en.wikipedia.org/w/index.php?oldid=508043670> *Contributors:* Adrianwn, Akim Demaille, Alksentrs, Andreas Kaufmann, Archimerged, Backpackadam, Beige Tangerine, Bigaln2, Bkkbrad, Boydski, Charles Matthews, Chipchap, Chkno, Cldoyle, David Eppstein, Dcoetzee, Deewiant, Dominus, Dr Greg, Edward Vielmetti, Fresheneesz, Gfonsecabr, Giftlite, Grendelkhan, Headbomb, IanLiu, Integr, Isaac Dupree, Jamelan, Jonel, Kasei-jin, Kenahoo, Kevtrice, LOL, Lambiam, LittleDan, Mankarse, MasterAchilles, MathijsM, Michael Hardy, Msh210, NawlinWiki, Nikaustr, Nyenyec, Oaf2, Oli Filth, Pakaran, Pgan002, Qunwanges157, Qutezuze, Qwertyus, Rb Brewer42, Rednas1234, ReyBrujo, Ripper234, Rjwilmsi, Ruud Koot, Salix alba, SamRushing, Sceptre, Shmilymann, Spike Wilbury, Spirko, Stellmach, Superjoe30, Tevildo, The Anome, Tonycao, Vanisheduser12a67, Wavelength, Williacb, Wmayer, Zhouji2010, 81 anonymous edits

**Space partitioning** *Source:* <http://en.wikipedia.org/w/index.php?oldid=451613601> *Contributors:* Altenmann, Arminahmady, Bernhard Bauer, CBM, Flyinglemon, Frostyandy2k, Giftlite, Glen, Headbomb, Kjmathew, M-le-mot-dit, Michael Hardy, Mod mng, Niky cz, Nlu, Oleg Alexandrov, PaulJones, Reedbeta, Ruud Koot, Timrb, Yumf, 8 anonymous edits

**Binary space partitioning** *Source:* <http://en.wikipedia.org/w/index.php?oldid=509588926> *Contributors:* Abdull, Altenmann, Amanaplanacanalpanama, Amritchowdhury, Angela, AquaGeneral, Ariade, B4hand, Bomazi, Brucenaylor, Brutal deluxe, Bryan Derksen, Cbraga, Cgbuff, Chan siuman, Charles Matthews, Chrisjohnson, CyberSkull, Cybercobra, DanielPharos, David Eppstein, Dcoetzee, Dionyziz, Dysprosia, Fredrik, Frencheigh, Gbruin, GregorB, Headbomb, Immibis, Jafet, Jamesontai, Jkwchui, JohnnyMrNinja, Kdau, Kelvie, KnightRider, Kri, LOL, Leithian, LogiNevermore, M-le-mot-dit, Mdob, Michael Hardy, Mild Bill Hiccup, Miquonranger03, Noxin911, NtpNtp, NuclearFriend, Obiwhonn, Oleg Alexandrov, Operator link, Palmin, Percivall, Prikipedia, QuasarTE, RPHv, Reedbeta, Spodi, Stephan Leeds, Svick, Tabletop, Tarquin, The Anome, Twri, Wiki alf, WikiLaurent, WiseWoman, Wmahan, Wonghang, Yar Kramer, Zetawoof, 68 anonymous edits

**Segment tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=497454676> *Contributors:* AdultSwim, Aida147, Alfredo J. Herrera Lago, Csernica, Cybercobra, Dean p foster, Dhruvbird, Dicklyon, Diomidis Spinellis, Dstrash, Jj137, Mikhail Dvorkin, MoreNet, Optigan13, Phishman3579, Portalian, Rahmtin rotabi, Tgeairn, Tortoise 74, X7q, XLerate, Yintianjiao, 22 anonymous edits

**Interval tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=505089990> *Contributors:* 28421u2232nfencenc, Breuwi, Brutal deluxe, CharlesGillingham, Cobi, CommonsDelinker, Cosmoskramer, CostinulAT, Daiyuda, David Eppstein, Dcoetzee, Dean p foster, Dicklyon, Digulla, Favonian, Gareth Griffith-Jones, Isomeme, Jamelan, Kungi01, Mkcmkc, Mloskot, Phishman3579, Porao, RJFJR, Raboof, Raduberinde, Rodgling, Spakin, Svick, Tortoise 74, X7q, Zhybear, 44 anonymous edits

**Range tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=497061121> *Contributors:* Andreas Fabri, Breuwi, Caesura, Cybercobra, David Eppstein, Dcoetzee, Infvwl, RJFJR, Rfmge, Sanchom, Smalljim, 16 anonymous edits

**Bin** *Source:* <http://en.wikipedia.org/w/index.php?oldid=499301912> *Contributors:* Agent007ravi, AndrewHowse, Andycjp, Chris the speller, David Eppstein, Lifebaka, Lockley, Michael Hardy, Ospalh, Yumf, 4 anonymous edits

**k-d tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=434119400> *Contributors:* Achurch, Algomaster, Altenmann, Amit man, Andreas Kaufmann, Angus Lepper, Arkitus, AwamerT, BD2412, Bajsejohannes, Bearcat, BenFrantzDale, Bisqwit, Borax, Bosik GN, Braddodson, Brandonwood, BrotherE, Bryan Derksen, Btyner, Bunyk, C. Iorenz, Casbah, Ceran, CesarB, Chad.burrs, Colingodsey, Connelly, Cybercobra, Cymru.lass, David Eppstein, Dcoetzee, Dicklyon, Don Reba, EQ5afN2M, Equendil, Favonian, Ficeup, Formerly the IP-Address 24.22.227.53, Genieser, Giftlite, Grendelkhan, Huan086, Iamchenzetian, Ilyathemuromets, Jeff Wheeler, Jfmiller28, Jokes Free4Me, Jtweng45, Justin W Smith, Karlhendrikse, Karloman2, Kirigiri, KiwiSunset, Kri, Leyo, Lfstevens, MYguel, Madduck, Matt think so, Mcld, Mephistophelian, Miym, MladenWiki, Mohamedadaly, MusicScience, Neile, NikolasCo, Nintendude, Paul A, Pixor, Reedbeta, Revjp, Rich Farmbrough, Riitoken, Rjwilmsi, Rummelsworth, Ruud Koot, SJP, Sfandino, Skinkie, SteinbDJ, Steheng3, Svick, Szabolcs Nagy, The Thing That Should Not Be, TheParanoidOne, Trolololololol, Ts3r0, Uncle Milty, Uranor, User A1, Vegaswikian, Waldir, Woohookitty, Wsloand, Wwheeler, Yumf, ماني, 178 anonymous edits

**Implicit k-d tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=434178778> *Contributors:* Andreas Kaufmann, BD2412, Bduvenhage, Cybercobra, Dicklyon, Genieser, Iridescent, Karlhendrikse, KoenDelaere, Kri, Meawoppl, Nick Number, Paul A, R'n'B, Rjwilmsi, Rockfang, Svick, Tesi1700, Tim Q. Wells, 2 anonymous edits

**min/max kd-tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=471784555> *Contributors:* Andreas Kaufmann, Bduvenhage, Genieser, Jokes Free4Me, KoenDelaere, Nick Number, Rockfang, Signalhead, 3 anonymous edits

**Adaptive k-d tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=471784172> *Contributors:* David Eppstein, Favonian, Onodevo, Qutezuze, Wsloand

**Quadtree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=508226327> *Contributors:* 4pq1injbok, ALoopingIcon, Aburad, Aeternus, AllanBz, Arthena, BZRatfink, Backpackadam, Bryan Derksen, Cdc, Cuaxdon, Cybercobra, Danny, David Eppstein, Dcoetzee, Dgreenheck, DrWolfen, Dysprosia, Enochlau, Frap, Fredrik, Gaius Cornelius, GeorgeMoney, Giftlite, Gimmetrow, Happyvalley, Headbomb, Hermann.tropf, Hyperplane, Interior, Jason Davies, Johnflux, Jprg1966, Lancect, Lfstevens, Maglev2, Miko3k, MilerWhite, MorganRoderick, Nickblack, Oleg Alexandrov, Peterjanbroomans, Piet Delpert, RHaden, RJFJR, Ronz, Saffles, Shencoop, Spottedowl, TechnologyTrial, Tomo, Tony1212, Wojciech mula, Wsloand, Yrithind, 55 anonymous edits

**Octree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=499292252> *Contributors:* 23u982j32r29j92, Alienskull, Alksub, Andreas Kaufmann, Arthena, AySz88, Balabiot, Bkkbrad, Bryan Derksen, CesarB, Claystu, Cybercobra, Defiantredpill, DragonflySixtyseven, Dysprosia, Eike Welk, Exe, Ffaarr, Fredrik, Furrykef, Giftlite, Indiana State, JaGa, Jacob grace, JosephCatrambone, June8th, Kierano, KoenDelaere, Kri, Lfstevens, Litherum, MIT Trekkie, MagiMaster, Melaen, Noerrorsfound, Nomis80, Nothings, Olsonse, Rahul220, Ruud Koot, Sadangel, Scott Paeth, Scott5114, SimonP, Sscomp2004, TechnologyTrial, Tomjenkins52, Viriditas, WhiteTimberwolf, Wolfkeeper, Wsloand, 88 anonymous edits

**Linear octrees** *Source:* <http://en.wikipedia.org/w/index.php?oldid=417019860> *Contributors:* Eeekster, Jncraton, Pgan002, 1 anonymous edits

**Z-order** *Source:* <http://en.wikipedia.org/w/index.php?oldid=424157495> *Contributors:* Andreas Kaufmann, Beland, BenFrantzDale, Black Falcon, Blup, CBM, Cariaso, CesarB, Daniel Minor, David Eppstein, DnetSvlg, Edratzer, Einstein9073, Ephraim33, Fisherjs, Giftlite, Hermann.tropf, Hesperian, Joriki, Kku, Kri, Lambiam, Lendorien, Lfstevens, Magioladitis, Michael Hardy, Paul Foxworthy, Pnm, Robertd, Shadowjams, Sligocki, Sverdrup, Tonsyskjellum, VivaEmilyDavies, Wpegden, Ynhockey, Zotel, 18 anonymous edits

**UB-tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=503189259> *Contributors:* Andreas Kaufmann, Cybercobra, Hermann.tropf, Honnza, Jason.surratt, Lfstevens, Rich Farmbrough, Svick, Travelbird, Welsh, 5 anonymous edits

**R-tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=497271799> *Contributors:* AAA!, Alejos, AllanBz, Altenmann, Alynna Kasmira, Avono, Bernhard Bauer, BrotherE, CALR, Chire, Curps, Cutelyaware, Cybercobra, David Eppstein, Digital Organism, Download, Elwood j blues, EncMstr, FatalError, Foobaz, Freekh, G.hartig, Gwern, Hadal, Happyvalley, Huisho, Hydrogen Iodide, Imbcndth, Jarble, Jason.surratt, Jodawi, Lfstevens, MIT Trekkie, MacTed, Minghong, Mwtoews, Nick Pizarro, Jr., NoldorinElf, Oik, Okoky, Peterjanbroomans, Pqrstuv, Qutezuze, Radim Baca, Skinkie, Soumyasch, Spexios, Stolee, Svick, TRKeen, Tastyjew123, The Anome, Tony1212, Twri, Ziyuang, 85 anonymous edits

**R+ tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=451612737> *Contributors:* Alksentrs, AllanBz, Blup, Brian0918, CesarB, Creidieki, Fuzzie, Happyvalley, Headbomb, J.pellicoli, Jodawi, Kerowyn, Lfstevens, Minghong, NikolasCo, Nintendude, OsamaBinLogin, Pearle, Pmdusso, RJFJR, Svick, Tim Starling, Turnstep, Westquote, WikHead, 14 anonymous edits

**R\* tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=500905627> *Contributors:* Chire, Frouaix, Happyvalley, Jodawi, Kounoupis, Lfstevens, Minghong, Neilc, Svick, The Anome, Virtuald, 19 anonymous edits

**Hilbert R-tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=488374663> *Contributors:* Andreas Kaufmann, BD2412, Belizefan, Chire, D6, Damsgård, Erechtheus, FelixR23, Hammertime, JCarlos, Jodawi, Johnfn, Michael Hardy, Mwtoews, Niczar, Niteris, Okoky, Robin K. Albrecht, SimonD, Svick, Twri, 29 anonymous edits

**X-tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=386389426> *Contributors:* Chire, Cybercobra, Danceswithzerglings, Drunguy8800, Hyperfusion, Lfstevens, Mild Bill Hiccup, 6 anonymous edits

**Metric tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=471784552> *Contributors:* Brighterorange, Bryan Derksen, Chire, Cybercobra, Itman, Monkeyget, Mrzuniga333, Nsk92, Staffelde, 4 anonymous edits

**vP-tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=469317168> *Contributors:* Brutal deluxe, Bryan Derksen, Cybercobra, Dicklyon, Hadal, Itman, Mcl d, Piet Delpoit, Sjesse, User A1, 3 anonymous edits

**BK-tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=502185339> *Contributors:* Bryan Derksen, Cybercobra, DoriSmith, Ehamberg, Endpoint, Fikus, Itman, Mkarlesky, Piet Delpoit, Quuxplusone, Staffelde, Templatetypedef, Volkan YAZICI, Wossi, 11 anonymous edits

**Hash table** *Source:* <http://en.wikipedia.org/w/index.php?oldid=508149750> *Contributors:* 2001:468:701:4004:754B:6F0A:8D59:2FB5, ASchmoo, Acdx, AdamRetchless, Adrianwn, Ahoerstemeier, Ahyl1, Ajo Mama, Akuchling, Alksub, Allstarecho, Aloksukhwani, Altenmann, Andreas Kaufmann, Antaeus Feldspar, Antaeus Feldspar, Anthony Appleyard, Anurnmi, Apanag, Arbalest Mike, Arpitm, Ashwin, AxelBoldt, Axlrosen, AznBurger, Baka toroi, Baliame, BenFrantzDale, Bevo, BckKnght, Bobo192, Bug, C4chandu, CRGreathouse, CWenger, CanadianLinuxUser, CanisRufus, Carmichael, CesarB's unprivileged account, Cfai lde, Cgma, Ched, Cic, Cntras, Cobi, Cometstyles, Conversion script, Criebe, CryptoDerk, Cryptoid, Cybercobra, DNewhall, DSatz, Damian Yerrick, Daniel Waterworth, David Eppstein, DavidCary, Davidfstr, Davidgothberg, Dcoetzee, Decltype, Demokoryu, Denispir, Derek Parnell, Derek farn, Deshraj, Digwuren, Dmbstudio, Dmcomer, Doug Bell, Drbreznjev, Drilnoth, Dysprosia, Ecb29, Eddof13, Eminull, EncMstr, Erel Segal, Esh, Everyking, FeralOink, Filu, Floodyberry, Fragglet, Frap, Frecklefoot, Fredrik, Frehley, Fresheneesz, Furrykef, Giftlite, Glrx, GregorB, Gremel123, Groxx, Gulliveig, Gustavb, HFuruseth, Happyuk, Helios2k6, Helix84, Hetori, Hosamaly, Hydrox, Ibbn, Iekpo, IghushevEdward, Imran, Incompetence, Intgr, Iron Wallaby, IronGargoyle, Jjuran, JLaTondre, Jim1138, Jk2q3jrklse, Johnunig, Jorge Stolfi, Josephsieh, Josephskeller, Justin W Smith, JustinWick, Karuthedam, Kastchei, Kaustuv, Kbdank71, Khrose, Khalid, Kinu, Knutux, Krakhan, Kungfuadam, LOL, LapoLuchini, Larry V, Lavenderbunny, Linguica, Luqui, Luxem, MER-C, Magnus Bakken, MaxEnt, Mayrel, Mcom320, Mdd, Meand, MegaHasher, Meneth, Michael Hardy, Mike.aizatsky, Mild Bill Hiccup, Miles, Mipadi, Mousehousemd, MrOllie, Mrwojo, Nanshu, Narendrak, Neilc, Nethgibr, Neurodivergent, Nightkhaos, Nixdorf, Nneonneo, Not-just-yeti, Nuno Tavares, ObfuscatePenguin, Om Sao, Omegatron, Oravec, Ouishoebean, Pagh, Pakaran, Patmorin, Paul Mackay, Paulsheer, Pbruneau, Peter Horn, PeterCanthropus, Pgan002, Piccolomomo, Pichpich, Pixor, PizzaMargherita, Pnm, Purple, QrczakMK, R3m0t, Radagast83, Raph Levien, Rawafmail, Rememberway, Rgamble, RichLewis, Rjwilmsi, Sae1962, Sam Hovecar, Sandos, Scandum, Schwarzbichler, Sebleblanc, Secretlondon, Seizethedave, Shadowjams, Shafigoldwasser, Shanes, Shmageggy, Shuchung, Simonsarris, Simulationelson, SiobhanHansa, Sleske, Sligocki, Sonjaaa, Stannered, Svick, Sycomonster, T Long, TShilo12, Tackline, TakuyaMurata, Teacup, Teapeat, Th1rt3en, The Anome, TheTraveler3, Thermon, Ticklempink42, Tikiwont, Tjdw, Tomchiuck, Tostie14, Triwbe, UtherSRG, Varuna, Velociostich, W Nowicki, Waltpohl, Watcher, Wavelength, Wernher, Wikilolo, Winecellar, Wmahan, Wolfkeeper, Wolkykim, Woohookitty, Wrp103, X7q, Zundark, 474 anonymous edits

**Hash function** *Source:* <http://en.wikipedia.org/w/index.php?oldid=509702609> *Contributors:* 10metreh, 2620:0:1000:1500:BAAC:6FFF:FE80:D37E, Adolphus79, Agent X2, Akerans, Alexcollins, Altenmann, Andipi, Antaeus Feldspar, Apoc2400, Applesed, Applrpn, Atlant, Bender235, BiT, BckKnght, Bluechimera0, Bomac, Borezd, BrianRice, Brona, Brookie, Butterwell, Capricorn42, CesarB's unprivileged account, Chalst, Christofpaar, Chuck Simmons, Claygate, Cntras, Coffee2theorems, Connelly, Conversion script, Courcelles, Crashtatch, Cuddlyable3, CyberSkull, Damian Yerrick, Dappawit, David spector, Davidgothberg, Dcoetzee, Demosta, Derek Ross, Dfinkel, Dkf11, Doradus, Doug4, Dranorter, Drostie, Dwheeler, EddEdmondson, Eisel, Emc2, EngineerScotty, Epr123, Erebus Morgaine, Eurosang, Evand, Everyking, Fender123, Folic Acid, Francis Tyers, FusionNow, Fæ, GVOLT, GarbledLecture933, Garyzx, Gfoley4, Ggia, Giftlite, Globbet, Graham87, GroveGuy, Gulliveig, Gurch, Hamsterlopitecus, Harmil, Harpreet Osahan, Helix84, Hfastedge, Iamhigh, Imran, InShaneeen, Incompetence, Intgr, InverseHypercube, Ixfd64, J-Wiki, J.delanoy, Jachto, JakeD409, Jbalint, Jc, Jeaise, Jediknil, Jeff Erickson, Jeltz, JesseStone, Jfroeich, Jopxton, Jorge Stolfi, KHAAAAAAAAAAN, Karl-Henner, Karuthedam, Kastchei, Kazvorpal, Keith Cascio, KelySYC, Kgfleischmann, Kirrages, Knutux, Kotha arun2005, KrizzyB, Kurivaim, Kuru, Kusunose, L337 kybldmstr, LOL, Lambiam, Leedeth, Linas, LittleBenW, LjL, M2millenium, MIT Trekkie, Malycetenar, MarkWegman, Materialschemist, MathStuf, Matt Crypto, Mattoothman, Maurice Carbonaro, Mblumber, Mdebt, MegaHasher, Meiskam, Mghgtg, Michael Hardy, MichealH, Midnightcomm, Miguel, Mikeblas, Mikepelley, Mild Bill Hiccup, Mmmready, Mnb9rca, MrOllie, Ms2ger, Mking, Music Sorter, My76Strat, Nageh, Nanshu, Nbarth, Neelix, Neuronatron, Nguyen Thanh Quang, Ninly, Noformation, Noldoaran, Nyttend, Obli, Octahedron80, Oli Filth, OverlordQ, Pabix, Patelm, Paul Richter, Paulschou, Peter bertok, Peyre, Pfunk42, Phil Spectre, PhilHibbs, PierreAbbat, Pinethicket, Pndfam05, Powerdesi, Pseudomonas, QmunkE, Quota, Qwertys, R'n'B, Raghaw, RattusMaximus, Rememberway, Rich Farmbrough, Rishi.bedi, Rjwilmsi, SamHartman, SavantEdge, Schmittey, ShakingSpirit, ShawnVW, Shi Hou, Shingra, Sjones23, Sklender, Snoops, Stangaa, StealthFox, SteveJothan, Sven Manguard, Svnsn, Tachyon01, Talrias, Taw, Teapeat, Tech77, Tgeairn, That Guy, From That Show!, TheNightFly, Thenowhereman, TomViza, Tonsyan, Toolan, TruthIIPower, TutterMouse, Twri, Ultimtis, Updatehelper, Utcursch, Vanis, Velella, Vrenator, Vstarre, Watcher, WereSpielChequers, White Trillium, Wikilolo, Winston Chuen-Shih Yang, Wjaguar, Wolfkeeper, Ww, Xelgen, Zeno of Elea, ঞ, 504 anonymous edits

**Open addressing** *Source:* <http://en.wikipedia.org/w/index.php?oldid=500091958> *Contributors:* Cobi, DataWraith, Fresheneesz, Gildos, Guzzrocha, Headbomb, LOL, O keyes, Skwa, ماني, 9 anonymous edits

**Lazy deletion** *Source:* <http://en.wikipedia.org/w/index.php?oldid=508673639> *Contributors:* A Generic User, AManWithNoPlan, Alksub, Bearcat, Fresheneesz, The Doink, Xezbeth, 11 anonymous edits

**Linear probing** *Source:* <http://en.wikipedia.org/w/index.php?oldid=508949599> *Contributors:* A3 nm, Andreas Kaufmann, Bearcat, C. A. Russell, CesarB's unprivileged account, Chris the speller, Danmoberly, David Eppstein, Dixtosa, Enochlau, Gazpacho, Infinity ive, Jeberle, Jngnyc, JonHarder, Linas, Negrulio, OliviaGuest, Patmorin, RJFJR, Sbluen, SpuriousQ, Tas50, Tedzdog, Themanian, 27 anonymous edits

**Quadratic probing** *Source:* <http://en.wikipedia.org/w/index.php?oldid=506498153> *Contributors:* Andreas Kaufmann, Bavla, C. A. Russell, CesarB's unprivileged account, Cybercobra, David Eppstein, Dcoetzee, Enochlau, Idanz, Kmgrpattyush, Mikeblas, Oleg Alexandrov, Philip Trueman, Rich Farmbrough, Robertvan1, Ryk, Simeon, Vaibhav1992, Yashykt, ZeroOne, 36 anonymous edits

**Double hashing** *Source:* <http://en.wikipedia.org/w/index.php?oldid=503371801> *Contributors:* Amiceli, Angela, AxelBoldt, Bdawson1982, CesarB, CesarB's unprivileged account, Cobblet, DasBrose, Dcoetzee, Gurch, Hashbrowncipher, Imposing, JForget, Only2sea, Pfunk42, RJFJR, Smallman12q, Usmrne h8er, WonderPhil, Zawersh, 43 anonymous edits

**Cuckoo hashing** *Source:* <http://en.wikipedia.org/w/index.php?oldid=506102582> *Contributors:* Alquantor, Arvindn, Bomazi, BuZZdEE.BuZZ, CRGreathouse, CesarB's unprivileged account, Cwolsfsheep, Cybercobra, David Eppstein, Dcoetzee, Ej, Headbomb, Hermel, Hetori, Jafet, LiranKatzir, Lmonson26, Mandyhan, Mark cummins, Neilc, Nyh, Pagh, Phil Boswell, Pps, Rcsprinter123, S3000, Svick, Themanian, Thore Husfeldt, Thumperward, Userask, Valentas.Kurauskas, Wjaguar, Zawersh, Zerodamage, Zr2d2, 40 anonymous edits

**Coalesced hashing** *Source:* <http://en.wikipedia.org/w/index.php?oldid=339004477> *Contributors:* Algotime, Andreas Kaufmann, Basawala, CesarB's unpriviledged account, Cic, Confuzzled, Dcoetzee, Fresheneesz, Ian Pitchford, Jafet, Jll, Oleg Alexandrov, Pmdboi, Tassedethe, Zawersh, 8 anonymous edits

**Perfect hash function** *Source:* <http://en.wikipedia.org/w/index.php?oldid=498141865> *Contributors:* 4hodmt, Arka sett, Burschik, CesarB's unpriviledged account, Cimon Avaro, Cobi, Dcoetzee, Dtrebbien, Dlugosz, E David Moyer, Fredrik, G121, Gajeam, Giftlite, Glrx, Headbomb, Itman, JMCORE, Johndburger, LOL, Maysak, Mcichelli, MegaHasher, Mudd1, Neilc, Otus, Pagh, Ruud Koot, Salrizvy, Spl, SteveT84, Wikilolo, 38 anonymous edits

**Universal hashing** *Source:* <http://en.wikipedia.org/w/index.php?oldid=507736628> *Contributors:* ArnoldReinhold, Arnstein87, Bender2k14, CharlesHBennett, ChrisGualtieri, Cleo, Copland Stalker, Copysan, Cybercobra, Cyberjoac, DanielLemire, David Eppstein, Dmharvey, Dwmalone, EmilJ, Francois.boutines, Golgofrinchian, Guruparan18, Headbomb, Jafet, Mpatrascu, Neilc, Pagh, Patmorin, RPHv, Rjwilmsi, Rswarbrick, Sdorman, SeanMack, Superm401, TPReal, Twintop, Ulamgamer, Winxa, Zolgharnein, 36 anonymous edits

**Linear hashing** *Source:* <http://en.wikipedia.org/w/index.php?oldid=498308884> *Contributors:* Az186a, BrokenSegue, CanadianLinuxUser, CesarB's unpriviledged account, David Eppstein, Dcoetzee, ENGIMa, Gail, Headbomb, Julthep, MegaHasher, MrTux, Personman, Res2216firestar, Rjwilmsi, TenPoundHammer, Zawersh, 39 anonymous edits

**Extendible hashing** *Source:* <http://en.wikipedia.org/w/index.php?oldid=477325819> *Contributors:* Alex Kapranoff, Amiceli, Boing! said Zebedee, Firsfron, Gloomy Coder, Headbomb, John of Reading, JustinWick, MegaHasher, Planetary Chaos Redux, Spamduck, Svick, Tommy2010, Treekids, Twimoki, Wolfeye90, WorldsApart, Zawersh, 65 anonymous edits

**2-choice hashing** *Source:* <http://en.wikipedia.org/w/index.php?oldid=258474388> *Contributors:* Asparagus, JaGa, Onodevo, Pichpich, Zawersh

**Pearson hashing** *Source:* <http://en.wikipedia.org/w/index.php?oldid=402967665> *Contributors:* Abednigo, Avocado, Bkonrad, Charles Matthews, Chieffhuggbear, Creative1985, Dwmalone, Dlugosz, Ed Poor, Epolk, Fredrik, Frigoris, Glrx, LittleDan, Oli Filth, Phannister, PeterPearson, Phil Boswell, Quuxplusone, The Anome, 12 anonymous edits

**Fowler–Noll–Vo hash function** *Source:* <http://en.wikipedia.org/w/index.php?oldid=499783062> *Contributors:* Apoc2400, Bachrach44, Boredzo, Bubba73, CRGreathouse, Cybercobra, Damian Yerrick, David spector, Denispir, Elf-friend, Enochlau, GabrielReid, HumphreyW, Ianbicking, Jorge Stolfi, Landon Curt Noll, Liko81, Ospalh, Phil Spectre, Raph Levien, Raymondwinn, Rbarreira, Ron Ritzman, Runtime, TruthIIPower, Uzume, Woohookitty, 43 anonymous edits

**Bitstate hashing** *Source:* <http://en.wikipedia.org/w/index.php?oldid=287405139> *Contributors:* Andreas Kaufmann, Jirislaby, Mild Bill Hiccup, Paalappoo, 1 anonymous edits

**Bloom filter** *Source:* <http://en.wikipedia.org/w/index.php?oldid=508634620> *Contributors:* 2620:149:4:401:39AC:BBCA:5443:59FA, Agl, Akryzhn, Alan Liefthing, Alexmadon, Andreas Kaufmann, AndreasBWagner, Argav, Ashish goel public, Babbage, Bblfish, Bender2k14, Bramp, Brighterorange, Buddeyp, Cconnett, CesarB's unpriviledged account, Charles Matthews, Chenopodiaceous, Chocolateboy, Chronulator, Cybercobra, David Eppstein, DavidCary, Dcoetzee, Drae, Drafiei, Drangon, Dtrebbien, Dzhim, Edward, Emorrissey, EvilGrin, Farzaneh, Flexdream, Fredrik, Furrykef, Galaad2, Galaxiaad, Gharb, Giftlite, Gtoal, Hanche, Headbomb, HereToHelp, Hilgerdenaar, Igodard, James A. Donald, Jeremy Banks, Jerz4835, Jjld, Justin W Smith, Ktj373, Kronos04, Labongo, Lesshaste, Lindsay658, Liorithiel, Marokwitz, Michael Hardy, Mindmatrix, Msikma, Naufraghi, NavenduJain, Neilc, Nhahmada, Osndok, Payrad, Pfunk42, Pnm, Quanstro, Quuxplusone, Requestion, Rlauffer, Rubicantoto, Russianspy3, Ryan Reizch, RzR, Sdorman, Ses4j, Shakeelmahate, Sharma337, ShaunMacPherson, Shreevatsa, Stev0, SteveJochen, Subrabbitt, Sun Creator, Susvolans, Svick, Tabletop, Terrycojones, The Anome, Trachten, Vsriram, Wavelength, Willpwillp, Wirthi, Wwwolf, X7q, Xiphoris, Yin, 203 anonymous edits

**Locality preserving hashing** *Source:* <http://en.wikipedia.org/w/index.php?oldid=501690233> *Contributors:* Adverick, Alok169, BenFrantzDale, Cataclysm, Elaborating, Jitse Niesen, Ligulem, Lourakis, Mdd, Phil Boswell, Zorakoid, 3 anonymous edits

**Morton number** *Source:* <http://en.wikipedia.org/w/index.php?oldid=456588713> *Contributors:* Andreas Kaufmann, Beland, BenFrantzDale, Black Falcon, Blup, CBM, Carioso, CesarB, Daniel Minor, David Eppstein, DnetSv, Edratzer, Einstein9073, Ephraim33, Fishersj, Giftlite, Hermann.tropf, Hesperian, Joriki, Kku, Kri, Lambiam, Lendorien, Lfstevens, Magioladitis, Michael Hardy, Paul Foxworthy, Pnm, Robertd, Shadowjams, Sligocki, Sverdrup, Tonsyskjellum, VivaEmilyDavies, Wpegden, Ynhockey, Zotel, 18 anonymous edits

**Zobrist hashing** *Source:* <http://en.wikipedia.org/w/index.php?oldid=471875788> *Contributors:* Asimperson, Charles Matthews, DanielLemire, Dlugosz, Evand, Giftlite, IanOsgood, Jafet, Julian Cardona, Matthew Woodcraft, Qwertys, Shirifan, Stephen Morley, WVhybrid, ZeroOne, 12 anonymous edits

**Rolling hash** *Source:* <http://en.wikipedia.org/w/index.php?oldid=502217836> *Contributors:* A5b, Bo Lindbergh, CesarB's unpriviledged account, DanielLemire, Eug, Hegariz, Jafet, Johndburger, Martlau, 36 anonymous edits

**Hash list** *Source:* <http://en.wikipedia.org/w/index.php?oldid=506987221> *Contributors:* CapitalR, CesarB's unpriviledged account, ChrisGualtieri, D climacus, DataWraith, Davidgothberg, Digitalme, Grm wnr, Harryboyles, Homerjay, IstvanWolf, Jeff3000, JonLS, K.ngdfhg, M, Ruud Koot, 14 anonymous edits

**Hash tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=502592136> *Contributors:* Arthur Rubin, Azaghal, Blaisorblade, CesarB's unpriviledged account, Ciphergoth, Cybercobra, Darthshak, DataWraith, Davidgothberg, Derek Ross, Doctorhook, GargoylMT, Gartenzweg 99, Gdr, Giftlite, Gil Schmidt, Glom2215, Gvern, Haakon, Headbomb, Htmllapps, Intgr, Jim Douglas, Jorgbrown, Kazvorpal, Knaag, Korrawit, Kroepke, M, Marudubshinki, Matt Crypto, Michael Hardy, Nonnompow, Nowhere man, Nysin, Oyigit, Pete279, Pmokeefe, Purebill, Rashless, Rconan, Rhys.rhaven, RistoLaanoja, Rmashhadi, Ruud Koot, Spexios, Tamfang, Thattommyhall, Theeldest, Thv, Tr00st, Ullner, Zack, 54 anonymous edits

**Prefix hash tree** *Source:* <http://en.wikipedia.org/w/index.php?oldid=330831311> *Contributors:* Dbenenn, Intgr, Miym, Nad, The Thing That Should Not Be, 2 anonymous edits

**Hash trie** *Source:* <http://en.wikipedia.org/w/index.php?oldid=477393784> *Contributors:* Evand, Gdr, Magioladitis, OrangeDog, Pombredanne, Tr00st, 2 anonymous edits

**Hash array mapped trie** *Source:* <http://en.wikipedia.org/w/index.php?oldid=504701279> *Contributors:* Andreas Kaufmann, Axel22, Bunnyhop11, D6, DanielWaterworth, Dgpop, Ehird, Erf5432, IMneme, JensPetersen, Jessecooke, Jrw@pobox.com, Wynand.winterbach, 13 anonymous edits

**Distributed hash table** *Source:* <http://en.wikipedia.org/w/index.php?oldid=508993828> *Contributors:* Alfio, Allquixotic, Anescient, Anthony, Apoc2400, AstReseach, Athymik, Baa, Bernhard Bauer, Bpringleimeir, Br, Bryan Derksen, Bubba hotep, Cacheonix, Cburnett, Charles Matthews, Chrismiceli, Chy168, Clsin, Cojoco, Corti, CosineKitty, Crystallina, Cwolfoosheep, CyberSkull, DJPhazer, DerHexer, Diberri, Dto, Edward, Ehn, Elauimini, Elf, EmilSit, Enzo Aquarius, Eras-mus, Eric.weigle, Erik Sandberg, F, Fcrick, FlyHigh, Frap, GPHemsley, Garas, Gary King, Godrickwok, Goofi1781, Gpierre, Greywiz, Haakon, Hadal, Happyrabbit, Harryboyles, Hbpencil, Hu12, Imeshev, Ingenth, Intgr, Irrbloss, Itai, Jamesday, Jda, Jensocan, Jnlm, John13535, Johnleach, JonHarder, KDesk, Karnesky, Kbrose, Khalid hassani, Khiladi 2010, KnightRider, Knuckles, Liamzebedee, Luqui, M.B, MMuzammils, Mabdul, Maix, Md4567, Meneth, Mgrinich, Michael miceli, Mike Rosoff, MikeSchmid111, Minority Report, Miym, Monkeyjunky, Mpias, Nad, Naff89, Nazlrag, Nealmcb, NeilN, Nethgib, Neustradamus, Nhorton, Noodles-sb, OPless, Ohnoitsjamie, Old Death, OrangeDog, Oskilian, PetrIvanov, Pkg, Phil Boswell, ReinoutS, Roger pack, Rogerdpack, Ronz, Sae1962, Scgtrp, Sean01, Search4Lancer, ShaunMacPherson, Shentino, Squalho, Ssspera, TakuyaMurata, TedPavlic, The Anome, The Belgain, TheJosh, Thingg, Thv, Tombeo, Tomdo08, UtherSRG, Wayne Slam, X1987x, Xiong Chiamiov, Xosé, Xubupt, Yramesh, 285 anonymous edits

**Consistent hashing** *Source:* <http://en.wikipedia.org/w/index.php?oldid=506795011> *Contributors:* Argv0, CesarB's unpriviledged account, ChrisGualtieri, DavidCary, Dirx, DreamOfMirrors, Edratzer, Egil, Fyrael, Gareth Jones, Harald Haugland, Harrigan, Headbomb, Hut 8.5, JonGretar, McLaughlin77, Miym, Mukkiepc, Nethgib, Pbrandao, Rene Mas, Richagandhewar, Rilak, Rohinidradthod, Sae1962, Simongdkelly, Skalet, SmallRepair, Tim Goddard, Wavelength, X7q, 36 anonymous edits

**Stable hashing** *Source:* <http://en.wikipedia.org/w/index.php?oldid=389147116> *Contributors:* Alasdair, Alex Kapranoff, CesarB's unpriviledged account, Minna Sora no Shita, Ssilverm, The Anome, TheBlueFlamingo, YUL8YYZ, 1 anonymous edits

**Koorde** *Source:* <http://en.wikipedia.org/w/index.php?oldid=415653058> *Contributors:* Bchocie, Bladegirl, D6, Fwippy, JPG-GR, Jorgeantibes, Kbrose, Kylu, Starwiz, Suhhy, Wekkolin, 5 anonymous edits

**Graph** *Source:* <http://en.wikipedia.org/w/index.php?oldid=463712910> *Contributors:* 31stCenturyMatt, A Aleg, Aaronzat, Alink, Andreas Kaufmann, Any Key, AvicAWB, Avoided, Bluebusy, Booyabazooka, Bruyninc, C4Cypher, Chochopk, Chrisholland, Cooldudefx, Cybercobra, DRAGON BOOSTER, David Eppstein, Dcoetzee, Dysprosia, Epimethius, FedericoMenaQuintero, Gallando, Giftlite, Gmelli, Graphicalx, Gvanrossum, Hariva, Hobsonlane, Jojit fb, Jon Awbrey, JonHarder, Jorge Stolfi, Juliancolton, Kate4341, Kazubon, Kbrose, KellyCoinGuy, Kendrick Hang, Klortho, KristjanJonasson, Labraun90, Liao, Max Terry, NerdyNSK, Nmz787, Obradovic Goran, Ovi 1, P0nc, Pbirmie, Pieleric, QuiteUnusual, R. S. Shaw, RG2, Rabarberski, Rborrego, Rd232, Rhanekom, Ruud Koot, Sae1962, Saimhe, Salix alba, ScNewcastle, ScaledLizard, SimonFuhrmann, Simonfairfax, SiobhanHansa, Skippydo, Sthpung, TFloto, Timwi, Tyir, UKoch, Zoicon5, ZorrolIII, ZweiOhren, 140 anonymous edits

**Adjacency list** *Source:* <http://en.wikipedia.org/w/index.php?oldid=507861707> *Contributors:* Andreas Kaufmann, Ash211, Beetstra, Bobbertface, Booyabazooka, Chmod007, Chris the speller, Cobi, Craig Pemberton, David Eppstein, Dcoetzee, Dysprosia, Fredrik, Garyzx, Giftlite, Hariva, Hobsonlane, Iridescent, Jamelan, Justin W Smith, Jwpurple, Kku, Krazelke, MathMartin, Michael Hardy, NSR, Oleg Alexandrov, Only2sea, Patmorin, Rdude, Ricardo Ferreira de Oliveira, SPTWriter, Sabrinamagers, Schneelocke, Serketan, The cRaCKEr, Twri, 26 anonymous edits

**Adjacency matrix** *Source:* <http://en.wikipedia.org/w/index.php?oldid=507736145> *Contributors:* Abdull, Abha Jain, Aleph4, Arthouse, AxelBoldt, Beetstra, BenFrantzDale, Bender235, Bender2k14, Bitsianshash, Booyabazooka, Burn, Calle, Churnett, Chenopodiaceous, David Eppstein, Dcoetzee, Debresser, Dreadstar, Dysprosia, ElBenevolente, Felix Hoffmann, Fredrik, Garyzx, Gauge, Giftlite, Headbomb, Hu12, JackSchmidt, JamesBWatson, Jean Honorio, Jogers, John of Reading, Jpbowen, Juffi, Kneufeld, Kompot 3, Lipedia, MarkSweep, MathMartin, Mbogelund, Mdrine, Michael Hardy, Miy, Morre, Natelewis, Oleg Alexandrov, Olenielsen, Only2sea, Patmorin, Paulish, Periergeia, Phils, Reina riemann, Rgdboer, Rheth, Rich Farnbrough, RobinK, SPTWriter, Salgueiro, Schneelocke, Senfo, Shilpi4560, Snowcream, Slawomir Bialy, TakuyaMurata, Tamfang, Thackstr, Tim Q. Wells, Timendum, Tomo, Ttzz, Twri, X7q, Yoav HaCohen, YuryKirkenko, Zaslav, بفسى, 72 anonymous edits

**And-inverter graph** *Source:* <http://en.wikipedia.org/w/index.php?oldid=505098142> *Contributors:* Aaron Hurst, Alan Mishchenko, Andreas Kaufmann, Andy Crews, Appraiser, Ettrig, Gregbard, Igor Markov, Jlang, Jon Awbrey, Jonathan de Boyne Pollard, Ketiltrout, Linas, Michael Hardy, Mikeblas, Nobletripe, Pigorsch, Pkkao2, Rjwilmsi, Trovatore, 5 anonymous edits

**Binary decision diagram** *Source:* <http://en.wikipedia.org/w/index.php?oldid=507503054> *Contributors:* Ajo Mama, Amccosta, Andreas Kaufmann, Andris, AshtonBenson, Bigfootbigfoot, Bluemoose, Bobke, Boute, Brighterorange, Britlak, Calabe1992, Charles Matthews, ChrisGualtieri, David Eppstein, David.Monniaux, Denim1965, Derek farn, Dirk Beyer, EmilJ, Greenrd, GregorB, Gtrmp, Happyuk, Hermel, Heron, Hitanshu D, J04n, Jason Recliner, Esq., Jay Here, Jcarroll, Jkl, KSchutte, Kakesson, Karlbrace, Karltk, Karsten Strehl, Kinu, Laudaka, LouScheffer, Matumio, McCart42, Mdd, Michael Hardy, Nouiz, Ort43v, Pce3@ij.net, Qwertys, Rohit.nadig, Rorro, Ruud Koot, Ryan Clark, Sade, SailorH, Sam Hovecar, Sirutan, Sun Creator, Taw, Tjifo098, Trivialist, Trovatore, Twri, Uli, YHoshua, YannTM, 115 anonymous edits

**Binary moment diagram** *Source:* <http://en.wikipedia.org/w/index.php?oldid=451611787> *Contributors:* Andreas Kaufmann, Headbomb, Jon Awbrey, Michael Hardy, Taw, 2 anonymous edits

**Zero-suppressed decision diagram** *Source:* <http://en.wikipedia.org/w/index.php?oldid=491002125> *Contributors:* Andreas Kaufmann, Bkell, Doradus, Eep?, Esbenrune, Graham87, Jon Awbrey, Qwertys, Taw, TubularWorld, Waggars, 5 anonymous edits

**Propositional directed acyclic graph** *Source:* <http://en.wikipedia.org/w/index.php?oldid=491665847> *Contributors:* Aagtbdfoua, Andreas Kaufmann, BD2412, Brighterorange, DRap, Dennis714, Nbarth, RUN, Selket, Tjifo098, Trovatore, 4 anonymous edits

**Graph-structured stack** *Source:* <http://en.wikipedia.org/w/index.php?oldid=399640593> *Contributors:* Andreas Kaufmann, Charles Matthews, Djspiewak, Fredrik, Jaredwf, Sadads, SpaceMoose, TimBentley, 2 anonymous edits

**Scene graph** *Source:* <http://en.wikipedia.org/w/index.php?oldid=498123690> *Contributors:* Altenmann, Amatheny, Andreas Kaufmann, Anita.havele, Ap, Arun Geo John, BAxelrod, C4Cypher, CALR, CTho, CamTarn, Cbraga, Cmdrjameson, Cpu111, Cybercobra, Davepape, Docu, DzzD, EmanWilm, Engwar, Fredrik, Furrykef, GregorB, J04n, JLaTondre, Jfmantis, Jmdyck, KickAssClown, Kmote, Lancect, Leroyluc, Lockeownj00, Lowellian, Martin Kraus, Mav, Mayalld, Mecanismo, Michael Hardy, Mortene, Mushin, Nknight, Polyglut, PeterV1510, Reinyday, Rivelozzo, Rufous, Sae1962, Stephanie40, SynergyBlades, Tommstein, Tony Sidaway, WillC2 45220, Wsriley, Zundark, 75 anonymous edits

**Big O notation** *Source:* <http://en.wikipedia.org/w/index.php?oldid=508703038> *Contributors:* 132.204.25.xxx, 4v4l0n42, A-Ge0, A. Pichler, ABCD, Abdull, Adashiel, Addps4cat, Aelvin, Ahmad Faridi, Ahoerstermeier, Alan smithee, Alex Selby, Algoman101, Alksentrs, AllanBz, Altenmann, AnOddName, Andre Engels, Andreas Kaufmann, Ankit Maity, Anonymous Dissident, Anthony Appleyard, Apanag, Arno Matthias, Arthur Rubin, Arunmoezhi, Arvindn, Ascánder, AvicAWB, AxelBoldt, B4hand, BMB, Bagsc, Barak Sh, Baronjonas, Ben pcc, BenFrantzDale, Bergstra, Bird of paradox, Bkell, Bomazi, Booyabazooka, Borgx, Brad7777, Brion VIBBER, Btyner, Bubba73, Buster79, C45207, CBKAAtopsails, CRGreathouse, Calculuslover, Charlow3, Charles Matthews, CharlesGillingham, ChazBeckett, ChrisForno, Colfulus, Compotatoj, Connelly, Conversion script, Cookie4869, CosineKitty, Curb Chain, Curps, Cybercobra, CyborgTosser, D4g0thur, DFS454, Dachshund, Dadudadu, Damian Yerrick, Danakil, Danny, Dark Charles, David Eppstein, Davidwt, Deljr, Dcoetzee, Deeparnab, Den fjättrade ankan, Derlay, Dhuss, Diberrri, Diego diaz espinoza, Dionyziz, DniQ, Donfbreed, Doradus, Dr. Universe, Draco flavus, Drpaule, Duagloth, Dysprosia, EconoPhysicist, Efnar, El C, Elephant in a tornado, Eleveneleven, Elias, EmilJ, Enochlau, Epachamo, Eric119, Ernie shoemaker, Eus Kevin, FauxFaux, Fayenatic london, Fede Reghe, Felix Wiemann, Fennec, FiachraByrne, Fibonacci, FilipeS, Flouran, Foxjwill, Fredrik, Fvw, Gadig, Gene Ward Smith, GeorgieMcBain, Giftlite, Gilliam, Gjd001, Glassmage, Gracenotes, Graham87, Gremagor, Gutza, H.ehsaan, Haham hanuka, Hans Adler, Hdante, Head, Headbomb, HenningThielemann, Henrygb, Hermel, Hlg, Ichernev, Intrg, InverseHypercube, Isis, Ixf64, JHMM13, JIP, Jacobolus, James.S, Jaredwf, Javit, Jeronimo, Jim1138, Jleedev, JoeKearney, JoergenB, JohnWStockwell, Jolsfa123, Jonathanzung, Josephjeevan, JoshuaZ, Jowan2005, Jpkotta, Jthillik, Justin W Smith, Jwh335, Kan8eDie, Katsushi, KneeLess, KoenDelaere, Koffeyahoo, Kri, LC, LOL, Lambiam, Lamro, Leithp, Leonard G., LeonardoGregianin, Leyceec, Linas, Ling.Nut, Lugnad, Luqui, MFH, MIT Trekkie, Macrakis, Mad Jaqk, Maksim-e, Marc van Leeuwen, MarkOlah, MathMartin, Matiasholte, Mattbuck, McKay, Mcstrother, Melcombe, Michael Hardy, Michael Rogers, Michael Slone, Miguel, Mike Schwartz, Mindmatrix, Mitchoyoshitaka, Miy, Mobius, Modeha, Mpagano, Mrypsilon, Mstuomel, Mxn, Nbarth, NephestTheFirst, Neile, Nejko, NeoUrfahrner, Netheri96, Nils Grimsso, NovaDog, O Pavlos, Oleg Alexandrov, Oliphaunt, Opelio, Optikos, Ott2, PGWG, PL290, Patrick, Patrick Lucas, Paul August, PaulTanenbaum, Paxcoder, Pcuff, Pete4512, PhilKnight, Plutor, Poor Yorick, Prumpf, Quendus, Qwfp, R'n'B, R.e.b., R3m0t, Raknarf44, Rebroad, Reinderien, RexNL, Rfl, Riceplaytexas, Rjwilmsi, RobertBorgersen, RobinK, Rockingravi, Rovenhot, Royote, Rschwieb, Ruud Koot, Sabalka, Sameer0s, Sapphorain, SchfiftyThree, Sciurinae, ScotsmanRS, Shalom Yechiel, Shellgirl, Shizhao, Shoessss, Shreevatsa, Simetrical, Simon Fenney, Skaraoke, Sligocki, Smjg, Sophus Bie, Spitzak, Stefan.karpinski, Stephen Compall, Stevenj, Stevertigo, Stimpy, Sydbarrett74, Syncategoremata, Szepl, TNARasslin, Taemyr, TakuyaMurata, Tardis, Tarotcards, Taw, The Anome, TheBiggestFootballFan, TheSeven, Thenub314, Tide rolls, Timwi, Tony Fox, Tosha, Tritium6, Tyco.skinner, Ultimus, Universalss, User A1, Vanisheduser12a67, Vector, Vedant, VictorAnyakin, WavePart, Whosyourjudas, Whouk, Wikibuki, Writer on wiki, Wtmitchell, Yarin Kaul, ZAB, Zack, Zero sharp, ZeroOne, ZiggyMo, Zowch, Zundark, Zvika, Île flottante, 661 anonymous edits

**Amortized analysis** *Source:* <http://en.wikipedia.org/w/index.php?oldid=503638334> *Contributors:* Allens, Altenmann, Andreas Kaufmann, BarretB, Brazzy, Brona, CWenger, Caesura, Dcoetzee, EconoPhysicist, Giftlite, Jimp, Joriki, Laubrau, Mantipula, Michael Hardy, MrBlok, Nneonneo, Oravec, Pierre de Lyon, Poor Yorick, PrologFan, Qutezuze, R'n'B, Rbonvall, Safek, Stevo2001, Svick, TakuyaMurata, Torzsmokus, User A1, Vramasub, Worch, 33 anonymous edits

**Locality of reference** *Source:* <http://en.wikipedia.org/w/index.php?oldid=508808027> *Contributors:* 16@r, 18.94, A5b, Adrianwn, Andreas Kaufmann, BMF81, Ben Wraith, BenFrantzDale, Brianhe, Charles Matthews, Chasrmartin, Cic, Costaluzi, Cpiral, Cyberjoac, DTM, Derek farn, Dfarrell07, Dinno, Ecb29, Ee00224, Einstein9073, Felix Andrews, Firsfron, Fredrik, Headbomb, Helwr, Intrg, Ixf64, JPG-GR, John, JonHarder, Jruderman, KatelynJohann, Kbdank71, Kenyon, Kurt Jansson, Maverick1715, Mboverload, Mirror Vax, Naroza, NocNokNeo, Not-just-yeti, O!English, Phils, Piet Delpport, Pnm, Prohlep, R'n'B, Radagast83, Randalomalous, ShakataGaNaI, Stephen Morley, TakuyaMurata, The Anome, Themusicgod1, Uncle G, Uttar, Wernher, Zawersh, 50 anonymous edits

**Standard Template Library** *Source:* <http://en.wikipedia.org/w/index.php?oldid=506545410> *Contributors:* 1exec1, A5b, Ahannani, Alex.mccarthy, AleyCZ, Alfio, Alksentrs, Alvin-cs, AnOddName, Andres, Andrew Helwer, Aprock, Aragorn2, Avinashm, Barfooz, BenFrantzDale, Benhoyt, Bobbyi, Brainfck, Bwagstaff, CWenger, Chemisor, ChessKnight, Chris-gore, Cibu, Crasshopper, Cubbi, Curps, DanielKO, David Eppstein, Decltype, Deineka, Demonkoryu, Deryck Chan, Discospinster, Dreftymac, Drnrgvry, Emuka, Enerjazzzer, Flex, Frecklefoot, Furrykef, Gaius Cornelius, Geregenn2, Glenn, GoldenMedian, Gustavo.mori, Hadas583, Happyuk, Hydrogen Iodide, Ian Pitchford, Ipsign, J04n, JTN, Jason Quinn, Jesse Viviano, Jibijibib, Jcolotti, Jkt, Jorend, Josh Cherry, Joule36e5, K3rb, Ken, Kutulu, Lawrencegold, LiDaobing, Lihaas, LilHelpa, Mankarse, MarSch, Marc Mongenet, Martin Moene, Merphant, Minghong, Mirror Vax, Mmerex, Modster, Moritz, Mrjeff, Mushroom, Muthudesigner, NapoliRoma, Neile, Nneonneo, Norm mit, Nwbeeson, Ollydbg, Omnicog, Orderud, Öguz Ergin, Pavel Vozenilek, Pedram.salehpoor, Pieleric, Piet Delpport, Pomoxis, Pt, Pulseczar, R2100, Rijkbenik, Rjwilmsi, Rookkey, RoySmith, Sae1962, Sairahulreddy, Saziel, Sdorrance, Sebor, Shoaler, Smalljim, Smyth, Soumyasch, Spoon!, Spurrmoses, Stoph, Strcat, Stretraider, Tacvek, Tavianator, TheGeomaster, Thumperward, TingChong Ma, Tkgaui, Tobias Bergemann, Toffile, Tony Sidaway, Torc2, Tyler Oderkirk, Uiweo, Vald, Vincenzo.romano, Whaa?, Wiki0709, Wmbolle, Wolkykim, Xerxesnine, Zigger, Zundark, Zzuuzz, 159 anonymous edits

# Image Sources, Licenses and Contributors

**Image:Hash table 3 1 1 0 1 0 0 SP.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Hash\\_table\\_3\\_1\\_1\\_0\\_1\\_0\\_0\\_SP.svg](http://en.wikipedia.org/w/index.php?title=File:Hash_table_3_1_1_0_1_0_0_SP.svg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Jorge Stolfi

**Image:Singly-linked-list.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Singly-linked-list.svg> *License:* Public Domain *Contributors:* Lasindi

**File: single\_node1.jpg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Single\\_node1.jpg](http://en.wikipedia.org/w/index.php?title=File:Single_node1.jpg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:IshtitaMundada

**Image:purely functional list before.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Purely\\_functional\\_list\\_before.svg](http://en.wikipedia.org/w/index.php?title=File:Purely_functional_list_before.svg) *License:* GNU Free Documentation License *Contributors:* Original uploader was VineetKumar at en.wikipedia

**Image:purely functional list after.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Purely\\_functional\\_list\\_after.svg](http://en.wikipedia.org/w/index.php?title=File:Purely_functional_list_after.svg) *License:* GNU Free Documentation License *Contributors:* Original uploader was VineetKumar at en.wikipedia

**Image:purely functional tree before.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Purely\\_functional\\_tree\\_before.svg](http://en.wikipedia.org/w/index.php?title=File:Purely_functional_tree_before.svg) *License:* GNU Free Documentation License *Contributors:* Original uploader was VineetKumar at en.wikipedia

**Image:purely functional tree after.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Purely\\_functional\\_tree\\_after.svg](http://en.wikipedia.org/w/index.php?title=File:Purely_functional_tree_after.svg) *License:* GNU Free Documentation License *Contributors:* Original uploader was VineetKumar at en.wikipedia

**Image:Singly linked list.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Singly\\_linked\\_list.png](http://en.wikipedia.org/w/index.php?title=File:Singly_linked_list.png) *License:* Public Domain *Contributors:* User:Dcoetzee

**Image:Data stack.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Data\\_stack.svg](http://en.wikipedia.org/w/index.php?title=File:Data_stack.svg) *License:* Public Domain *Contributors:* User:Boivie

**Image:ProgramCallStack2.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:ProgramCallStack2.png> *License:* Public Domain *Contributors:* Agateller, DragonLord, LobStoR, Obradovic Goran

**File:Decimaltobinary.pdf** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Decimaltobinary.pdf> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Nipunbayas

**Image:Tower of Hanoi.jpeg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Tower\\_of\\_Hanoi.jpeg](http://en.wikipedia.org/w/index.php?title=File:Tower_of_Hanoi.jpeg) *License:* GNU Free Documentation License *Contributors:* Ies, Ævar Arnþjórf Bjarmason

**File:Towersofhanoi1.pdf** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Towersofhanoi1.pdf> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Nipunbayas

**File:Towersofhanoi2.pdf** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Towersofhanoi2.pdf> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Nipunbayas

**File:Towersofhanoi3.pdf** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Towersofhanoi3.pdf> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Nipunbayas

**File:Towersofhanoi4.pdf** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Towersofhanoi4.pdf> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Nipunbayas

**File:Towerofhanoi.jpg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Towerofhanoi.jpg> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:PranavAmbhore

**File:Railroadcars2.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Railroadcars2.png> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:PranavAmbhore

**File:Quicksort1.pdf** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Quicksort1.pdf> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Nipunbayas

**File:Quicksort2.pdf** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Quicksort2.pdf> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Nipunbayas

**File:Quicksort3.pdf** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Quicksort3.pdf> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Nipunbayas

**File:Quicksort4.pdf** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Quicksort4.pdf> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Nipunbayas

**File:Quicksort5.pdf** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Quicksort5.pdf> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Nipunbayas

**File:Quicksort6.pdf** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Quicksort6.pdf> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Nipunbayas

**File:Quicksort7.pdf** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Quicksort7.pdf> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Nipunbayas

**File:Stockspan.pdf** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Stockspan.pdf> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Nipunbayas

**Image:Data Queue.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Data\\_Queue.svg](http://en.wikipedia.org/w/index.php?title=File:Data_Queue.svg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Foroa, Hazmat2, Martynas Patašius, Vegpuiff

**File:Loudspeaker.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Loudspeaker.svg> *License:* Public Domain *Contributors:* Bayo, Gmaxwell, Husky, Iamunknown, Mirithing, Myself488, Nethac DIU, Omegatron, Rocket000, The Evil IP address, Wouterhagens, 20 anonymous edits

**File:binary tree.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Binary\\_tree.svg](http://en.wikipedia.org/w/index.php?title=File:Binary_tree.svg) *License:* Public Domain *Contributors:* User:Dcoetzee

**Image:Array of array storage.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Array\\_of\\_array\\_storage.svg](http://en.wikipedia.org/w/index.php?title=File:Array_of_array_storage.svg) *License:* Public Domain *Contributors:* User:Dcoetzee

**File:Dynamic array.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Dynamic\\_array.svg](http://en.wikipedia.org/w/index.php?title=File:Dynamic_array.svg) *License:* Creative Commons Zero *Contributors:* User:Dcoetzee

**Image:HashedArrayTree16.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:HashedArrayTree16.svg> *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* Surachit at en.wikipedia

**Image:Circular buffer.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Circular\\_buffer.svg](http://en.wikipedia.org/w/index.php?title=File:Circular_buffer.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* en:User:Cburnett

**Image:Circular buffer - empty.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Circular\\_buffer\\_-\\_empty.svg](http://en.wikipedia.org/w/index.php?title=File:Circular_buffer_-_empty.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* en:User:Cburnett

**Image:Circular buffer - XX1XXXXX.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Circular\\_buffer\\_-\\_XX1XXXXX.svg](http://en.wikipedia.org/w/index.php?title=File:Circular_buffer_-_XX1XXXXX.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* en:User:Cburnett

**Image:Circular buffer - XX123XX.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Circular\\_buffer\\_-\\_XX123XX.svg](http://en.wikipedia.org/w/index.php?title=File:Circular_buffer_-_XX123XX.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* en:User:Cburnett

**Image:Circular buffer - XXXX3XX.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Circular\\_buffer\\_-\\_XXXX3XX.svg](http://en.wikipedia.org/w/index.php?title=File:Circular_buffer_-_XXXX3XX.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* en:User:Cburnett

**Image:Circular buffer - 6789345.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Circular\\_buffer\\_-\\_6789345.svg](http://en.wikipedia.org/w/index.php?title=File:Circular_buffer_-_6789345.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* en:User:Cburnett

**Image:Circular buffer - 6789AB5.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Circular\\_buffer\\_-\\_6789AB5.svg](http://en.wikipedia.org/w/index.php?title=File:Circular_buffer_-_6789AB5.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* en:User:Cburnett

**Image:Circular buffer - X789ABX.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Circular\\_buffer\\_-\\_X789ABX.svg](http://en.wikipedia.org/w/index.php?title=File:Circular_buffer_-_X789ABX.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* en:User:Cburnett

**Image:Circular buffer - XX123XX with pointers.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Circular\\_buffer\\_-\\_XX123XX\\_with\\_pointers.svg](http://en.wikipedia.org/w/index.php?title=File:Circular_buffer_-_XX123XX_with_pointers.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* en:User:Cburnett

**Image:Circular buffer - 6789AB5 with pointers.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Circular\\_buffer\\_-\\_6789AB5\\_with\\_pointers.svg](http://en.wikipedia.org/w/index.php?title=File:Circular_buffer_-_6789AB5_with_pointers.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* en:User:Cburnett

**Image:Circular buffer - 6789AB5 full.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Circular\\_buffer\\_-\\_6789AB5\\_full.svg](http://en.wikipedia.org/w/index.php?title=File:Circular_buffer_-_6789AB5_full.svg) *License:* GNU Free Documentation License *Contributors:* en:User:Cburnett, modifications de:User:DrZoom

**Image:Circular buffer - 6789AB5 empty.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Circular\\_buffer\\_-\\_6789AB5\\_empty.svg](http://en.wikipedia.org/w/index.php?title=File:Circular_buffer_-_6789AB5_empty.svg) *License:* GNU Free Documentation License *Contributors:* en:User:Cburnett, modifications de:User:DrZoom

**File:SCD algebraic notation.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:SCD\\_algebraic\\_notation.svg](http://en.wikipedia.org/w/index.php?title=File:SCD_algebraic_notation.svg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Chess\_board\_blank.svg: \*AAA\_SVG\_Chessboard\_and\_chess\_pieces\_04.svg: ILA-boy SCD\_algebraic\_notation.png: Klin derivative work: Beao SCD\_algebraic\_notation.png: Klin derivative work: Beao

**Image:Abramowitz&Stegun.page97.agr.jpg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Abramowitz&Stegun.page97.agr.jpg> *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* User:ArnoldReinhold, User:Trojancowboy

**Image:Interpolation example linear.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Interpolation\\_example\\_linear.svg](http://en.wikipedia.org/w/index.php?title=File:Interpolation_example_linear.svg) *License:* Public Domain *Contributors:* Berland

**File:Singly-linked-list.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Singly-linked-list.svg> *License:* Public Domain *Contributors:* Lasindi



**File:Linked list post office box analogy.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Linked\\_list\\_post\\_office\\_box\\_analogy.svg](http://en.wikipedia.org/w/index.php?title=File:Linked_list_post_office_box_analogy.svg) *License:* Creative Commons Zero *Contributors:* User:Dcoetzee

**Image:Circularly-linked-list.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Circularly-linked-list.svg> *License:* Public Domain *Contributors:* Lasindi

**Image:Doubly-linked-list.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Doubly-linked-list.svg> *License:* Public Domain *Contributors:* Lasindi

**File:CPT-LinkedLists-addingnode.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:CPT-LinkedLists-addingnode.svg> *License:* Public domain *Contributors:* Singly\_linked\_list\_insert\_after.png: Derrick Coetzee derivative work: Pluke (talk)

**File:CPT-LinkedLists-deletingnode.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:CPT-LinkedLists-deletingnode.svg> *License:* Public domain *Contributors:* Singly\_linked\_list\_delete\_after.png: Derrick Coetzee derivative work: Pluke (talk)

**File:Unrolled linked lists (1-8).PNG** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Unrolled\\_linked\\_lists\\_\(1-8\).PNG](http://en.wikipedia.org/w/index.php?title=File:Unrolled_linked_lists_(1-8).PNG) *License:* GNU Free Documentation License *Contributors:* Shigeru23

**File:VList example diagram.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:VList\\_example\\_diagram.png](http://en.wikipedia.org/w/index.php?title=File:VList_example_diagram.png) *License:* Public Domain *Contributors:* User:Dcoetzee

**Image:Skip list.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Skip\\_list.svg](http://en.wikipedia.org/w/index.php?title=File:Skip_list.svg) *License:* Public Domain *Contributors:* Wojciech Mula

**File:MTF Algorithm.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:MTF\\_Algorithm.png](http://en.wikipedia.org/w/index.php?title=File:MTF_Algorithm.png) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:SaurabhKB

**File:CountAlgorithm.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:CountAlgorithm.png> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:SaurabhKB

**File:Transpose Algorithm.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Transpose\\_Algorithm.png](http://en.wikipedia.org/w/index.php?title=File:Transpose_Algorithm.png) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:SaurabhKB

**Image:binary tree.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Binary\\_tree.svg](http://en.wikipedia.org/w/index.php?title=File:Binary_tree.svg) *License:* Public Domain *Contributors:* User:Dcoetzee

**File:Insertion of binary tree node.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Insertion\\_of\\_binary\\_tree\\_node.svg](http://en.wikipedia.org/w/index.php?title=File:Insertion_of_binary_tree_node.svg) *License:* Creative Commons Zero *Contributors:* Hazmat2

**File:Deletion of internal binary tree node.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Deletion\\_of\\_internal\\_binary\\_tree\\_node.svg](http://en.wikipedia.org/w/index.php?title=File:Deletion_of_internal_binary_tree_node.svg) *License:* Creative Commons Zero *Contributors:* Hazmat2

**Image:Binary tree in array.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Binary\\_tree\\_in\\_array.svg](http://en.wikipedia.org/w/index.php?title=File:Binary_tree_in_array.svg) *License:* Public Domain *Contributors:* User:Dcoetzee

**Image:N-ary to binary.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:N-ary\\_to\\_binary.svg](http://en.wikipedia.org/w/index.php?title=File:N-ary_to_binary.svg) *License:* Public Domain *Contributors:* CyHawk

**File:Binary search tree.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Binary\\_search\\_tree.svg](http://en.wikipedia.org/w/index.php?title=File:Binary_search_tree.svg) *License:* Public Domain *Contributors:* User:Booyabazooka, User:Dcoetzee

**File:binary search tree delete.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Binary\\_search\\_tree\\_delete.svg](http://en.wikipedia.org/w/index.php?title=File:Binary_search_tree_delete.svg) *License:* Public Domain *Contributors:* User:Dcoetzee

**Image:Unbalanced binary tree.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Unbalanced\\_binary\\_tree.svg](http://en.wikipedia.org/w/index.php?title=File:Unbalanced_binary_tree.svg) *License:* Public Domain *Contributors:* Me (Intgr)

**Image:AVLtreef.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:AVLtreef.svg> *License:* Public Domain *Contributors:* User:Mikm

**Image:Tree rotation.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Tree\\_rotation.png](http://en.wikipedia.org/w/index.php?title=File:Tree_rotation.png) *License:* GNU Free Documentation License *Contributors:* User:Ramasamy

**Image:Tree Rotations.gif** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Tree\\_Rotations.gif](http://en.wikipedia.org/w/index.php?title=File:Tree_Rotations.gif) *License:* Creative Commons Attribution-Sharealike 2.5 *Contributors:* C. A. Russell, Mtanti, 1 anonymous edits

**Image:Tree Rebalancing.gif** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Tree\\_Rebalancing.gif](http://en.wikipedia.org/w/index.php?title=File:Tree_Rebalancing.gif) *License:* Creative Commons Attribution-Sharealike 2.5 *Contributors:* Original uploader was Mtanti at en.wikipedia

**Image:Weight\_balanced\_tree2.jpg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Weight\\_balanced\\_tree2.jpg](http://en.wikipedia.org/w/index.php?title=File:Weight_balanced_tree2.jpg) *License:* Public Domain *Contributors:* KenE

**Image:Threaded tree.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Threaded\\_tree.svg](http://en.wikipedia.org/w/index.php?title=File:Threaded_tree.svg) *License:* Public Domain *Contributors:* R. S. Shaw

**File:ThreadTree Inorder Array.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:ThreadTree\\_Inorder\\_Array.png](http://en.wikipedia.org/w/index.php?title=File:ThreadTree_Inorder_Array.png) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Yeskw

**File:ThreadTree Inorder Array123456789.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:ThreadTree\\_Inorder\\_Array123456789.png](http://en.wikipedia.org/w/index.php?title=File:ThreadTree_Inorder_Array123456789.png) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Yeskw

**File:Normal Binary Tree.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Normal\\_Binary\\_Tree.png](http://en.wikipedia.org/w/index.php?title=File:Normal_Binary_Tree.png) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Vaibhavvc1092

**File:Threaded Binary Tree.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Threaded\\_Binary\\_Tree.png](http://en.wikipedia.org/w/index.php?title=File:Threaded_Binary_Tree.png) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Vaibhavvc1092

**File: AVL Tree Rebalancing.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:AVL\\_Tree\\_Rebalancing.svg](http://en.wikipedia.org/w/index.php?title=File:AVL_Tree_Rebalancing.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* CyHawk

**Image:binary search tree delete.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Binary\\_search\\_tree\\_delete.svg](http://en.wikipedia.org/w/index.php?title=File:Binary_search_tree_delete.svg) *License:* Public Domain *Contributors:* User:Dcoetzee

**Image:Red-black tree example.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Red-black\\_tree\\_example.svg](http://en.wikipedia.org/w/index.php?title=File:Red-black_tree_example.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* en>User:Cburnett

**Image:Red-black tree example (B-tree analogy).svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Red-black\\_tree\\_example\\_\(B-tree\\_analogy\).svg](http://en.wikipedia.org/w/index.php?title=File:Red-black_tree_example_(B-tree_analogy).svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* fr:Utilisateur:Verdy\_p

**Image:Red-black tree insert case 3.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Red-black\\_tree\\_insert\\_case\\_3.png](http://en.wikipedia.org/w/index.php?title=File:Red-black_tree_insert_case_3.png) *License:* Public Domain *Contributors:* Users Cintrom, Deco, Deelkar on en.wikipedia

**Image:Red-black tree insert case 4.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Red-black\\_tree\\_insert\\_case\\_4.png](http://en.wikipedia.org/w/index.php?title=File:Red-black_tree_insert_case_4.png) *License:* Public Domain *Contributors:* User Deco on en.wikipedia

**Image:Red-black tree insert case 5.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Red-black\\_tree\\_insert\\_case\\_5.png](http://en.wikipedia.org/w/index.php?title=File:Red-black_tree_insert_case_5.png) *License:* Public Domain *Contributors:* User Deco on en.wikipedia

**Image:Red-black tree delete case 2.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Red-black\\_tree\\_delete\\_case\\_2.png](http://en.wikipedia.org/w/index.php?title=File:Red-black_tree_delete_case_2.png) *License:* Public Domain *Contributors:* Users Deelkar, Deco on en.wikipedia

**Image:Red-black tree delete case 3.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Red-black\\_tree\\_delete\\_case\\_3.png](http://en.wikipedia.org/w/index.php?title=File:Red-black_tree_delete_case_3.png) *License:* Public Domain *Contributors:* User Deco on en.wikipedia

**Image:Red-black tree delete case 4.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Red-black\\_tree\\_delete\\_case\\_4.png](http://en.wikipedia.org/w/index.php?title=File:Red-black_tree_delete_case_4.png) *License:* Public Domain *Contributors:* User Deco on en.wikipedia

**Image:Red-black tree delete case 5.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Red-black\\_tree\\_delete\\_case\\_5.png](http://en.wikipedia.org/w/index.php?title=File:Red-black_tree_delete_case_5.png) *License:* Public Domain *Contributors:* User Deco on en.wikipedia

**Image:Red-black tree delete case 6.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Red-black\\_tree\\_delete\\_case\\_6.png](http://en.wikipedia.org/w/index.php?title=File:Red-black_tree_delete_case_6.png) *License:* Public Domain *Contributors:* User Deco on en.wikipedia

**Image:Red Black Shape Cases.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Red\\_Black\\_Shape\\_Cases.svg](http://en.wikipedia.org/w/index.php?title=File:Red_Black_Shape_Cases.svg) *License:* GNU Free Documentation License *Contributors:* Why Not A Duck

**Image:AA Tree Shape Cases.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:AA\\_Tree\\_Shape\\_Cases.svg](http://en.wikipedia.org/w/index.php?title=File:AA_Tree_Shape_Cases.svg) *License:* GNU Free Documentation License *Contributors:* Why Not A Duck

**Image:AA Tree Skew2.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:AA\\_Tree\\_Skew2.svg](http://en.wikipedia.org/w/index.php?title=File:AA_Tree_Skew2.svg) *License:* GNU Free Documentation License *Contributors:* User:Why Not A Duck based on work by

**Image:AA Tree Split2.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:AA\\_Tree\\_Split2.svg](http://en.wikipedia.org/w/index.php?title=File:AA_Tree_Split2.svg) *License:* GNU Free Documentation License *Contributors:* User:Why Not A Duck based on work by

**File:splay\_tree\_zig.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Splay\\_tree\\_zig.svg](http://en.wikipedia.org/w/index.php?title=File:Splay_tree_zig.svg) *License:* GNU Free Documentation License *Contributors:* Zig.gif: User:Regnar derivative work: Winniehell (talk)

**Image:Zigzig.gif** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Zigzig.gif> *License:* GNU Free Documentation License *Contributors:* Dcoetzee, Regnar, Remember the dot

**Image:Zigzag.gif** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Zigzag.gif> *License:* GNU Free Documentation License *Contributors:* Dcoetzee, Regnar, Remember the dot

**Image:T-tree-1.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:T-tree-1.png> *License:* GNU Free Documentation License *Contributors:* Original uploader was Hholzgra at en.wikipedia

**Image:T-tree-2.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:T-tree-2.png> *License:* GNU Free Documentation License *Contributors:* Hholzgra

**Image:Rope example.jpg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Rope\\_example.jpg](http://en.wikipedia.org/w/index.php?title=File:Rope_example.jpg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Meng Yao

**Image:Search rope.jpg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Search\\_rope.jpg](http://en.wikipedia.org/w/index.php?title=File:Search_rope.jpg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Meng Yao

**Image:Split1 rope.jpg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Split1\\_rope.jpg](http://en.wikipedia.org/w/index.php?title=File:Split1_rope.jpg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Meng Yao

**Image:Split2 rope.jpg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Split2\\_rope.jpg](http://en.wikipedia.org/w/index.php?title=File:Split2_rope.jpg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Meng Yao

**Image:Split3 rope.jpg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Split3\\_rope.jpg](http://en.wikipedia.org/w/index.php?title=File:Split3_rope.jpg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Meng Yao

**Image:Top tree.jpg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Top\\_tree.jpg](http://en.wikipedia.org/w/index.php?title=File:Top_tree.jpg) *License:* Public Domain *Contributors:* Afrozenator, Skier Dude

**File:Flg80.PNG** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Flg80.PNG> *License:* Public Domain *Contributors:* Tango tree

**File:Flg70.PNG** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Flg70.PNG> *License:* Public Domain *Contributors:* Tango tree

**File:Flg30.PNG** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Flg30.PNG> *License:* Public Domain *Contributors:* User:Tango Tree

**File:Flg90.PNG** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Flg90.PNG> *License:* Public Domain *Contributors:* Tango tree

**File:Flg100.PNG** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Flg100.PNG> *License:* Public Domain *Contributors:* Tango tree

**File:Flg110.PNG** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Flg110.PNG> *License:* Public Domain *Contributors:* Tango tree

**File:Flg120.PNG** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Flg120.PNG> *License:* Public Domain *Contributors:* Tango tree

**File:Flg130.PNG** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Flg130.PNG> *License:* Public Domain *Contributors:* Bkell, Tango tree

**File:Flg140.PNG** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Flg140.PNG> *License:* Public Domain *Contributors:* Tango tree

**File:Flg150.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Flg150.png> *License:* Creative Commons Attribution 3.0 *Contributors:* Tango tree

**File:Flg16.PNG** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Flg16.PNG> *License:* Public Domain *Contributors:* Tango tree

**File:Wilber10.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Wilber10.png> *License:* Creative Commons Attribution 3.0 *Contributors:* Tango tree

**Image:VebDiagram.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:VebDiagram.svg> *License:* Public Domain *Contributors:* Gailcarmichael

**File:Cartesian tree.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Cartesian\\_tree.svg](http://en.wikipedia.org/w/index.php?title=File:Cartesian_tree.svg) *License:* Public Domain *Contributors:* David Eppstein

**File:Cartesian tree range searching.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Cartesian\\_tree\\_range\\_searching.svg](http://en.wikipedia.org/w/index.php?title=File:Cartesian_tree_range_searching.svg) *License:* Public Domain *Contributors:* David Eppstein

**File:Bracketing pairs.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Bracketing\\_pairs.svg](http://en.wikipedia.org/w/index.php?title=File:Bracketing_pairs.svg) *License:* Public Domain *Contributors:* David Eppstein

**Image:TreapAlphaKey.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:TreapAlphaKey.svg> *License:* Creative Commons Zero *Contributors:* Qef

**File:B-tree.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:B-tree.svg> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* CyHawk

**Image:B tree insertion example.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:B\\_tree\\_insertion\\_example.png](http://en.wikipedia.org/w/index.php?title=File:B_tree_insertion_example.png) *License:* Public Domain *Contributors:* User:Maxtremus

**File:Bplustree.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Bplustree.png> *License:* Creative Commons Attribution 3.0 *Contributors:* Grundprinzip

**Image:2-3-4 tree 2-node.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:2-3-4\\_tree\\_2-node.svg](http://en.wikipedia.org/w/index.php?title=File:2-3-4_tree_2-node.svg) *License:* Public Domain *Contributors:* Chrismiceli

**Image:2-3-4 tree 3-node.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:2-3-4-tree\\_3-node.svg](http://en.wikipedia.org/w/index.php?title=File:2-3-4-tree_3-node.svg) *License:* Public Domain *Contributors:* Chrismiceli

**Image:2-3-4 tree 4-node.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:2-3-4\\_tree\\_4-node.png](http://en.wikipedia.org/w/index.php?title=File:2-3-4_tree_4-node.png) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* Ghakko, Haui

**Image:2-3-4 tree insert 1.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:2-3-4\\_tree\\_insert\\_1.svg](http://en.wikipedia.org/w/index.php?title=File:2-3-4_tree_insert_1.svg) *License:* Public Domain *Contributors:* Chrismiceli

**File:Queap.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Queap.svg> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Patrice Arruda

**File:FusionTreeSketch.gif** *Source:* <http://en.wikipedia.org/w/index.php?title=File:FusionTreeSketch.gif> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Vladfi

**Image:Bxtree.PNG** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Bxtree.PNG> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Bxtree

**Image:Max-Heap.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Max-Heap.svg> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Ermishin

**File:Max-heap.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Max-heap.png> *License:* Public Domain *Contributors:* Created by Onar Vikingstad 2005.

**File:Min-heap.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Min-heap.png> *License:* Public Domain *Contributors:* Original uploader was Vikingstad at en.wikipedia

**File:Heap add step1.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Heap\\_add\\_step1.svg](http://en.wikipedia.org/w/index.php?title=File:Heap_add_step1.svg) *License:* Public Domain *Contributors:* Ilmari Karonen

**File:Heap add step2.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Heap\\_add\\_step2.svg](http://en.wikipedia.org/w/index.php?title=File:Heap_add_step2.svg) *License:* Public Domain *Contributors:* Ilmari Karonen

**File:Heap add step3.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Heap\\_add\\_step3.svg](http://en.wikipedia.org/w/index.php?title=File:Heap_add_step3.svg) *License:* Public Domain *Contributors:* Ilmari Karonen

**File:Heap remove step1.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Heap\\_remove\\_step1.svg](http://en.wikipedia.org/w/index.php?title=File:Heap_remove_step1.svg) *License:* Public Domain *Contributors:* Ilmari Karonen

**File:Heap remove step2.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Heap\\_remove\\_step2.svg](http://en.wikipedia.org/w/index.php?title=File:Heap_remove_step2.svg) *License:* Public Domain *Contributors:* Ilmari Karonen

**File:Binary tree in array.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Binary\\_tree\\_in\\_array.svg](http://en.wikipedia.org/w/index.php?title=File:Binary_tree_in_array.svg) *License:* Public Domain *Contributors:* User:Dcoetzee

**File:Binary Heap with Array Implementation.JPG** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Binary\\_Heap\\_with\\_Array\\_Implementation.JPG](http://en.wikipedia.org/w/index.php?title=File:Binary_Heap_with_Array_Implementation.JPG) *License:* Creative Commons Zero *Contributors:* Bobmath

**Image:Binomial\_Trees.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Binomial\\_Trees.svg](http://en.wikipedia.org/w/index.php?title=File:Binomial_Trees.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* Kilom691, Lemontea, Ma-Lik

**Image:Binomial-heap-13.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Binomial-heap-13.svg> *License:* GNU Free Documentation License *Contributors:* User:D0ktorz

**Image:Binomial heap merge1.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Binomial\\_heap\\_merge1.svg](http://en.wikipedia.org/w/index.php?title=File:Binomial_heap_merge1.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* Lemontea

**Image:Binomial heap merge2.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Binomial\\_heap\\_merge2.svg](http://en.wikipedia.org/w/index.php?title=File:Binomial_heap_merge2.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* Lemontea

**Image:Fibonacci heap.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Fibonacci\\_heap.png](http://en.wikipedia.org/w/index.php?title=File:Fibonacci_heap.png) *License:* GNU Free Documentation License *Contributors:* User:Brona, User:Brona/Images/fibonacci\_heap.tex

**Image:Fibonacci heap extractmin1.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Fibonacci\\_heap\\_extractmin1.png](http://en.wikipedia.org/w/index.php?title=File:Fibonacci_heap_extractmin1.png) *License:* GNU Free Documentation License *Contributors:* User:Brona, User:Brona/Images/fibonacci\_heap.tex

**Image:Fibonacci heap extractmin2.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Fibonacci\\_heap\\_extractmin2.png](http://en.wikipedia.org/w/index.php?title=File:Fibonacci_heap_extractmin2.png) *License:* GNU Free Documentation License *Contributors:* User:Brona, User:Brona/Images/fibonacci\_heap.tex

**Image:Fibonacci heap-decreasekey.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Fibonacci\\_heap-decreasekey.png](http://en.wikipedia.org/w/index.php?title=File:Fibonacci_heap-decreasekey.png) *License:* GNU Free Documentation License *Contributors:* User:Brona, User:Brona/Images/fibonacci\_heap.tex

**Image:beap.jpg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Beap.jpg> *License:* GNU Free Documentation License *Contributors:* User:OrphanBot

**image:Leftist-trees-S-value.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Leftist-trees-S-value.svg> *License:* Public Domain *Contributors:* Computergeeksjw (talk)

**Image:Min-height-biased-leftist-tree-initialization-part1.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Min-height-biased-leftist-tree-initialization-part1.png> *License:* Public Domain *Contributors:* Buss, Qef

**Image:Min-height-biased-leftist-tree-initialization-part2.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Min-height-biased-leftist-tree-initialization-part2.png> *License:* Public Domain *Contributors:* Buss, Qef

**Image:Min-height-biased-leftist-tree-initialization-part3.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Min-height-biased-leftist-tree-initialization-part3.png> *License:* Public Domain *Contributors:* Buss, Qef

**Image:SkewHeapMerge1.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:SkewHeapMerge1.svg> *License:* Public Domain *Contributors:* Quinntaylor

**Image:SkewHeapMerge2.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:SkewHeapMerge2.svg> *License:* Public Domain *Contributors:* Quinntaylor

**Image:SkewHeapMerge3.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:SkewHeapMerge3.svg> *License:* Public Domain *Contributors:* Quinntaylor

**Image:SkewHeapMerge4.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:SkewHeapMerge4.svg> *License:* Public Domain *Contributors:* Quinntaylor

**Image:SkewHeapMerge5.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:SkewHeapMerge5.svg> *License:* Public Domain *Contributors:* Quinntaylor

**Image:SkewHeapMerge6.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:SkewHeapMerge6.svg> *License:* Public Domain *Contributors:* Quinntaylor

**Image:SkewHeapMerge7.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:SkewHeapMerge7.svg> *License:* Public Domain *Contributors:* Quinntaylor

**Image:trie example.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Trie\\_example.svg](http://en.wikipedia.org/w/index.php?title=File:Trie_example.svg) *License:* Public Domain *Contributors:* Booyabazooka (based on PNG image by Deco). Modifications by Superm401.

**File:BitwiseTreesScaling.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:BitwiseTreesScaling.png> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Ned14

**File:RedBlackTreesScaling.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:RedBlackTreesScaling.png> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Ned14

**File:HashTableScaling.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:HashTableScaling.png> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Ned14

**File:Patricia trie.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Patricia\\_trie.svg](http://en.wikipedia.org/w/index.php?title=File:Patricia_trie.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* Claudio Rocchini

**File:An example of how to find a string in a Patricia trie.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:An\\_example\\_of\\_how\\_to\\_find\\_a\\_string\\_in\\_a\\_Patricia\\_trie.png](http://en.wikipedia.org/w/index.php?title=File:An_example_of_how_to_find_a_string_in_a_Patricia_trie.png) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Saffles

**File:Inserting the string 'water' into a Patricia trie.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Inserting\\_the\\_string\\_'water'\\_into\\_a\\_Patricia\\_trie.png](http://en.wikipedia.org/w/index.php?title=File:Inserting_the_string_'water'_into_a_Patricia_trie.png) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Saffles

**File:Insert 'slower' with a null node into a Patricia trie.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Insert\\_'slower'\\_with\\_a\\_null\\_node\\_into\\_a\\_Patricia\\_trie.png](http://en.wikipedia.org/w/index.php?title=File:Insert_'slower'_with_a_null_node_into_a_Patricia_trie.png) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Saffles

**File:Insert 'test' into a Patricia trie when 'tester' exists.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Insert\\_'test'\\_into\\_a\\_Patricia\\_trie\\_when\\_'tester'\\_exists.png](http://en.wikipedia.org/w/index.php?title=File:Insert_'test'_into_a_Patricia_trie_when_'tester'_exists.png) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Saffles

**File:Inserting the word 'team' into a Patricia trie with a split.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Inserting\\_the\\_word\\_'team'\\_into\\_a\\_Patricia\\_trie\\_with\\_a\\_split.png](http://en.wikipedia.org/w/index.php?title=File:Inserting_the_word_'team'_into_a_Patricia_trie_with_a_split.png) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Saffles

**File:Insert 'toast' into a Patricia trie with a split and a move.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Insert\\_'toast'\\_into\\_a\\_Patricia\\_trie\\_with\\_a\\_split\\_and\\_a\\_move.png](http://en.wikipedia.org/w/index.php?title=File:Insert_'toast'_into_a_Patricia_trie_with_a_split_and_a_move.png) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Saffles

**Image:Suffix tree BANANA.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Suffix\\_tree\\_BANANA.svg](http://en.wikipedia.org/w/index.php?title=File:Suffix_tree_BANANA.svg) *License:* Public Domain *Contributors:* Maciej Jaros (commons: Nux, wiki-pl: Nux) (PNG version by Nils Grimsmo)

**Image:Suffix tree ABAB BABA.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Suffix\\_tree\\_ABAB\\_BABA.svg](http://en.wikipedia.org/w/index.php?title=File:Suffix_tree_ABAB_BABA.svg) *License:* Public Domain *Contributors:* derivative work: Johannes Rössel (talk) Suffix\_tree\_ABAB\_BABA.png: User:Nils Grimsmo

**Image:Trie-dawg.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Trie-dawg.svg> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Chkno

**Image:Andortree.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Andortree.png> *License:* Public Domain *Contributors:* User:saumaun

**File:SPQR tree.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:SPQR\\_tree.svg](http://en.wikipedia.org/w/index.php?title=File:SPQR_tree.svg) *License:* Public Domain *Contributors:* David Eppstein

**File:spaghettistack.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Spaghettistack.svg> *License:* Public Domain *Contributors:* Ealf

**File:Example of BSP tree construction - step 1.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Example\\_of\\_BSP\\_tree\\_construction\\_-\\_step\\_1.svg](http://en.wikipedia.org/w/index.php?title=File:Example_of_BSP_tree_construction_-_step_1.svg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Zahnradzacken

**File:Example of BSP tree construction - step 2.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Example\\_of\\_BSP\\_tree\\_construction\\_-\\_step\\_2.svg](http://en.wikipedia.org/w/index.php?title=File:Example_of_BSP_tree_construction_-_step_2.svg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Zahnradzacken

**File:Example of BSP tree construction - step 3.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Example\\_of\\_BSP\\_tree\\_construction\\_-\\_step\\_3.svg](http://en.wikipedia.org/w/index.php?title=File:Example_of_BSP_tree_construction_-_step_3.svg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Zahnradzacken

**File:Example of BSP tree construction - step 4.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Example\\_of\\_BSP\\_tree\\_construction\\_-\\_step\\_4.svg](http://en.wikipedia.org/w/index.php?title=File:Example_of_BSP_tree_construction_-_step_4.svg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Zahnradzacken

**File:Example of BSP tree construction - step 5.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Example\\_of\\_BSP\\_tree\\_construction\\_-\\_step\\_5.svg](http://en.wikipedia.org/w/index.php?title=File:Example_of_BSP_tree_construction_-_step_5.svg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Zahnradzacken

**File:Example of BSP tree construction - step 6.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Example\\_of\\_BSP\\_tree\\_construction\\_-\\_step\\_6.svg](http://en.wikipedia.org/w/index.php?title=File:Example_of_BSP_tree_construction_-_step_6.svg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Zahnradzacken

**File:Example of BSP tree construction - step 7.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Example\\_of\\_BSP\\_tree\\_construction\\_-\\_step\\_7.svg](http://en.wikipedia.org/w/index.php?title=File:Example_of_BSP_tree_construction_-_step_7.svg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Zahnradzacken

**File:Example of BSP tree construction - step 8.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Example\\_of\\_BSP\\_tree\\_construction\\_-\\_step\\_8.svg](http://en.wikipedia.org/w/index.php?title=File:Example_of_BSP_tree_construction_-_step_8.svg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Zahnradzacken

**File:Example of BSP tree construction - step 9.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Example\\_of\\_BSP\\_tree\\_construction\\_-\\_step\\_9.svg](http://en.wikipedia.org/w/index.php?title=File:Example_of_BSP_tree_construction_-_step_9.svg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Zahnradzacken

**File:Example of BSP tree traversal.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Example\\_of\\_BSP\\_tree\\_traversal.svg](http://en.wikipedia.org/w/index.php?title=File:Example_of_BSP_tree_traversal.svg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Zahnradzacken

**Image:Segment tree instance.gif** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Segment\\_tree\\_instance.gif](http://en.wikipedia.org/w/index.php?title=File:Segment_tree_instance.gif) *License:* Public Domain *Contributors:* Alfredo J. Herrera Lago

**File:1-dimensional-range-tree.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:1-dimensional-range-tree.svg> *License:* Creative Commons Zero *Contributors:* Matthew Eastman (talk)

**File:1-dimensional-range-query.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:1-dimensional-range-query.svg> *License:* Creative Commons Zero *Contributors:* Matthew Eastman (talk)

**Image:bin computational geometry.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Bin\\_computational\\_geometry.png](http://en.wikipedia.org/w/index.php?title=File:Bin_computational_geometry.png) *License:* Public Domain *Contributors:* McLoaf, Qef, Yumf

**File:3dtree.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:3dtree.png> *License:* GNU General Public License *Contributors:* Btyner, 2 anonymous edits

**File:Kdtree 2d.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Kdtree\\_2d.svg](http://en.wikipedia.org/w/index.php?title=File:Kdtree_2d.svg) *License:* GNU Free Documentation License *Contributors:* Original uploader was KiwiSunset at en.wikipedia

**File:Tree 0001.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Tree\\_0001.svg](http://en.wikipedia.org/w/index.php?title=File:Tree_0001.svg) *License:* Public Domain *Contributors:* MYguel

**File:KDTree-animation.gif** *Source:* <http://en.wikipedia.org/w/index.php?title=File:KDTree-animation.gif> *License:* Creative Commons Attribution-Sharealike 3.0,2.5,2.0,1.0 *Contributors:* User\_A1. Original uploader was User A1 at en.wikipedia

**Image:Implicitmaxkdtree.gif** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Implicitmaxkdtree.gif> *License:* Public Domain *Contributors:* Genieser

**Image:Point quadtree.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Point\\_quadtree.svg](http://en.wikipedia.org/w/index.php?title=File:Point_quadtree.svg) *License:* Public Domain *Contributors:* David Eppstein

**Image:Quad tree bitmap.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Quad\\_tree\\_bitmap.svg](http://en.wikipedia.org/w/index.php?title=File:Quad_tree_bitmap.svg) *License:* Public Domain *Contributors:* Wojciech Mula

**File:Octree2.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Octree2.svg> *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* WhiteTimberwolf, PNG version: Nü

**File:Four-level Z.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Four-level\\_Z.svg](http://en.wikipedia.org/w/index.php?title=File:Four-level_Z.svg) *License:* GNU Free Documentation License *Contributors:* David Eppstein, based on a image by Hesperian.

**File:Lebesgue-3d-step2.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Lebesgue-3d-step2.png> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Robert Dickau

**File:Lebesgue-3d-step3.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Lebesgue-3d-step3.png> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Robert Dickau

**File:Z-curve.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Z-curve.svg> *License:* Public Domain *Contributors:* David Eppstein

**File:Bigmin.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Bigmin.svg> *License:* Creative Commons Attribution-Sharealike 2.5 *Contributors:* User:DnetSvq

**Image:R-tree.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:R-tree.svg> *License:* Public Domain *Contributors:* Skinkie,

**Image:RTree-Visualization-3D.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:RTree-Visualization-3D.svg> *License:* Public Domain *Contributors:* Chire

**File:Zipcodes-Germany-GuttmanRTree.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Zipcodes-Germany-GuttmanRTree.svg> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Chire

**File:Zipcodes-Germany-AngTanSplit.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Zipcodes-Germany-AngTanSplit.svg> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Chire

**File:Zipcodes-Germany-RStarTree.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Zipcodes-Germany-RStarTree.svg> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Chire

**Image:RTree 2D.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:RTree\\_2D.svg](http://en.wikipedia.org/w/index.php?title=File:RTree_2D.svg) *License:* Public Domain *Contributors:* Chire

**Image:figure1 left.gif** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Figure1\\_left.gif](http://en.wikipedia.org/w/index.php?title=File:Figure1_left.gif) *License:* Public Domain *Contributors:* Okoky

**Image:figure1 right.gif** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Figure1\\_right.gif](http://en.wikipedia.org/w/index.php?title=File:Figure1_right.gif) *License:* Public Domain *Contributors:* Okoky

**Image:figure2 Hilbert.gif** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Figure2\\_Hilbert.gif](http://en.wikipedia.org/w/index.php?title=File:Figure2_Hilbert.gif) *License:* Public Domain *Contributors:* Okoky

**Image:figure3 data rects.gif** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Figure3\\_data\\_rects.gif](http://en.wikipedia.org/w/index.php?title=File:Figure3_data_rects.gif) *License:* Public Domain *Contributors:* Okoky

**Image:figure4 file structure.gif** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Figure4\\_file\\_structure.gif](http://en.wikipedia.org/w/index.php?title=File:Figure4_file_structure.gif) *License:* Public Domain *Contributors:* Okoky

**File:Hash table 3 1 1 0 0 SP.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Hash\\_table\\_3\\_1\\_1\\_0\\_0\\_SP.svg](http://en.wikipedia.org/w/index.php?title=File:Hash_table_3_1_1_0_0_SP.svg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Jorge Stolfi

**File:Hash table 5 0 1 1 1 1 LL.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Hash\\_table\\_5\\_0\\_1\\_1\\_1\\_1\\_LL.svg](http://en.wikipedia.org/w/index.php?title=File:Hash_table_5_0_1_1_1_1_LL.svg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Jorge Stolfi

**File:Hash table 5 0 1 1 1 1 0 LL.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Hash\\_table\\_5\\_0\\_1\\_1\\_1\\_1\\_0\\_LL.svg](http://en.wikipedia.org/w/index.php?title=File:Hash_table_5_0_1_1_1_1_0_LL.svg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Jorge Stolfi

**File:Hash table 5 0 1 1 1 1 0 SP.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Hash\\_table\\_5\\_0\\_1\\_1\\_1\\_1\\_0\\_SP.svg](http://en.wikipedia.org/w/index.php?title=File:Hash_table_5_0_1_1_1_1_0_SP.svg) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Jorge Stolfi

**File:Hash table average insertion time.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Hash\\_table\\_average\\_insertion\\_time.png](http://en.wikipedia.org/w/index.php?title=File:Hash_table_average_insertion_time.png) *License:* Public Domain *Contributors:* User:Dcoetzee

**Image:Hash table 4 1 1 0 0 1 0 LL.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Hash\\_table\\_4\\_1\\_1\\_0\\_0\\_1\\_0\\_LL.svg](http://en.wikipedia.org/w/index.php?title=File:Hash_table_4_1_1_0_0_1_0_LL.svg) *License:* Public Domain *Contributors:* Jorge Stolfi

**Image:Hash table 4 1 1 0 0 0 0 LL.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Hash\\_table\\_4\\_1\\_1\\_0\\_0\\_0\\_0\\_LL.svg](http://en.wikipedia.org/w/index.php?title=File:Hash_table_4_1_1_0_0_0_0_LL.svg) *License:* Public Domain *Contributors:* Jorge Stolfi

**Image:Hash table 4 1 0 0 0 0 0 LL.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Hash\\_table\\_4\\_1\\_0\\_0\\_0\\_0\\_0\\_LL.svg](http://en.wikipedia.org/w/index.php?title=File:Hash_table_4_1_0_0_0_0_0_LL.svg) *License:* Public Domain *Contributors:* Jorge Stolfi

**Image:HASHTB12.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:HASHTB12.svg> *License:* Public Domain *Contributors:* Helix84, Simeon87, Velociostrich, Xhienne

**File:Quadratic probing.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Quadratic\\_probing.png](http://en.wikipedia.org/w/index.php?title=File:Quadratic_probing.png) *License:* Creative Commons Zero *Contributors:* User:vaibhav1992

**Image:cuckoo.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Cuckoo.svg> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Zerodamage

**Image:CoalescedHash.jpg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:CoalescedHash.jpg> *License:* Public Domain *Contributors:* Confuzzled, Nv8200p

**Image:Extendible hashing 1.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Extendible\\_hashing\\_1.svg](http://en.wikipedia.org/w/index.php?title=File:Extendible_hashing_1.svg) *License:* Creative Commons Attribution 3.0 *Contributors:* Svick

**Image:Extendible hashing 2.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Extendible\\_hashing\\_2.svg](http://en.wikipedia.org/w/index.php?title=File:Extendible_hashing_2.svg) *License:* Creative Commons Attribution 3.0 *Contributors:* Svick

**Image:Extendible hashing 3.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Extendible\\_hashing\\_3.svg](http://en.wikipedia.org/w/index.php?title=File:Extendible_hashing_3.svg) *License:* Creative Commons Attribution 3.0 *Contributors:* Svick

**Image:Extendible hashing 4.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Extendible\\_hashing\\_4.svg](http://en.wikipedia.org/w/index.php?title=File:Extendible_hashing_4.svg) *License:* Creative Commons Attribution 3.0 *Contributors:* Svick

**Image:Extendible hashing 5.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Extendible\\_hashing\\_5.svg](http://en.wikipedia.org/w/index.php?title=File:Extendible_hashing_5.svg) *License:* Creative Commons Attribution 3.0 *Contributors:* Svick

**Image:Extendible hashing 6.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Extendible\\_hashing\\_6.svg](http://en.wikipedia.org/w/index.php?title=File:Extendible_hashing_6.svg) *License:* Creative Commons Attribution 3.0 *Contributors:* Svick

**File:Bloom filter.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Bloom\\_filter.svg](http://en.wikipedia.org/w/index.php?title=File:Bloom_filter.svg) *License:* Public Domain *Contributors:* David Epstein

**File:Bloom filter speed.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Bloom\\_filter\\_speed.svg](http://en.wikipedia.org/w/index.php?title=File:Bloom_filter_speed.svg) *License:* Public Domain *Contributors:* RMcPhillip, 1 anonymous edits

**File:Bloom filter fp probability.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Bloom\\_filter\\_fp\\_probability.svg](http://en.wikipedia.org/w/index.php?title=File:Bloom_filter_fp_probability.svg) *License:* Creative Commons Attribution 3.0 *Contributors:* Jerz4835

**File:AttenuatedBloomFilter.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:AttenuatedBloomFilter.png> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Akryzhn

**Image:Hash list.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Hash\\_list.svg](http://en.wikipedia.org/w/index.php?title=File:Hash_list.svg) *License:* Public Domain *Contributors:* Hash\_list.png: Original uploader was Davidgothberg at en.wikipedia conversion to SVG: DataWraith (talk)

**Image:Hash Tree.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Hash\\_Tree.svg](http://en.wikipedia.org/w/index.php?title=File:Hash_Tree.svg) *License:* Creative Commons Zero *Contributors:* User:Azaghal

**File:DHT en.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:DHT\\_en.svg](http://en.wikipedia.org/w/index.php?title=File:DHT_en.svg) *License:* Public Domain *Contributors:* Jinlin

**File:De\_bruijn\_graph-for\_binary\_sequence\_of\_order\_4.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:De\\_bruijn\\_graph-for\\_binary\\_sequence\\_of\\_order\\_4.svg](http://en.wikipedia.org/w/index.php?title=File:De_bruijn_graph-for_binary_sequence_of_order_4.svg) *License:* Public domain *Contributors:* SVG version: Wekkolin, Original version: english wikipedia user Michael\_Hardy

**Image:Koorde lookup routing.JPG** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Koorde\\_lookup\\_routing.JPG](http://en.wikipedia.org/w/index.php?title=File:Koorde_lookup_routing.JPG) *License:* Public Domain *Contributors:* Suhhy

**Image:Koorde lookup code.JPG** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Koorde\\_lookup\\_code.JPG](http://en.wikipedia.org/w/index.php?title=File:Koorde_lookup_code.JPG) *License:* Public Domain *Contributors:* Suhhy

**Image:6n-graf.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:6n-graf.svg> *License:* Public Domain *Contributors:* User:AzaToth

**Image:Simple cycle graph.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Simple\\_cycle\\_graph.svg](http://en.wikipedia.org/w/index.php?title=File:Simple_cycle_graph.svg) *License:* Public Domain *Contributors:* Booyabazooka at en.wikipedia

**Image:6n-graph2.svg** *Source:* <http://en.wikipedia.org/w/index.php?title=File:6n-graph2.svg> *License:* Public Domain *Contributors:* Booyabazooka, Dcoetzee, 1 anonymous edits

**File:Symmetric group 4; Cayley graph 1,5,21 (Nauru Petersen); numbers.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Symmetric\\_group\\_4;\\_Cayley\\_graph\\_1,5,21\\_\(Nauru\\_Petersen\);\\_numbers.svg](http://en.wikipedia.org/w/index.php?title=File:Symmetric_group_4;_Cayley_graph_1,5,21_(Nauru_Petersen);_numbers.svg) *License:* Public Domain *Contributors:* Lipedia

**File:Symmetric group 4; Cayley graph 1,5,21 (adjacency matrix).svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Symmetric\\_group\\_4;\\_Cayley\\_graph\\_1,5,21\\_\(adjacency\\_matrix\).svg](http://en.wikipedia.org/w/index.php?title=File:Symmetric_group_4;_Cayley_graph_1,5,21_(adjacency_matrix).svg) *License:* Public Domain *Contributors:* Lipedia

**File:Symmetric group 4; Cayley graph 4,9; numbers.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Symmetric\\_group\\_4;\\_Cayley\\_graph\\_4,9;\\_numbers.svg](http://en.wikipedia.org/w/index.php?title=File:Symmetric_group_4;_Cayley_graph_4,9;_numbers.svg) *License:* Public Domain *Contributors:* GrapheCayley-S4-Plan.svg: Fool (talk) derivative work: Lipedia (talk)

**File:Symmetric group 4; Cayley graph 4,9 (adjacency matrix).svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Symmetric\\_group\\_4;\\_Cayley\\_graph\\_4,9\\_\(adjacency\\_matrix\).svg](http://en.wikipedia.org/w/index.php?title=File:Symmetric_group_4;_Cayley_graph_4,9_(adjacency_matrix).svg) *License:* Public Domain *Contributors:* Lipedia

**Image:And-inverter-graph.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:And-inverter-graph.png> *License:* GNU Free Documentation License *Contributors:* Florian Pigorsch.

**File:BDD.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:BDD.png> *License:* GNU Free Documentation License *Contributors:* Original uploader was IMeowbot at en.wikipedia

**File:BDD simple.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:BDD\\_simple.svg](http://en.wikipedia.org/w/index.php?title=File:BDD_simple.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* Dirk Beyer

**File:BDD Variable Ordering Bad.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:BDD\\_Variable\\_Ordering\\_Bad.svg](http://en.wikipedia.org/w/index.php?title=File:BDD_Variable_Ordering_Bad.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* Dirk Beyer

**File:BDD Variable Ordering Good.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:BDD\\_Variable\\_Ordering\\_Good.svg](http://en.wikipedia.org/w/index.php?title=File:BDD_Variable_Ordering_Good.svg) *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* Dirk Beyer

**File:BDD2pdag.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:BDD2pdag.png> *License:* GNU Free Documentation License *Contributors:* Original uploader was RUN at en.wikipedia

**File:BDD2pdag simple.svg** *Source:* [http://en.wikipedia.org/w/index.php?title=File:BDD2pdag\\_simple.svg](http://en.wikipedia.org/w/index.php?title=File:BDD2pdag_simple.svg) *License:* GNU Free Documentation License *Contributors:* User:Selket and User:RUN (original)

**Image:Graph-structured stack 1 - jaredwf.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Graph-structured\\_stack\\_1\\_-\\_jaredwf.png](http://en.wikipedia.org/w/index.php?title=File:Graph-structured_stack_1_-_jaredwf.png) *License:* Public Domain *Contributors:* Original uploader was Jaredwf at en.wikipedia

**Image:Stacks jaredwf.png** *Source:* [http://en.wikipedia.org/w/index.php?title=File:Stacks\\_jaredwf.png](http://en.wikipedia.org/w/index.php?title=File:Stacks_jaredwf.png) *License:* GNU Free Documentation License *Contributors:* jaredwf

**File:Big-O notation.png** *Source:* <http://en.wikipedia.org/w/index.php?title=File:Big-O notation.png> *License:* GNU General Public License *Contributors:* Fede Reghe, Razorbliss

# License

---

Creative Commons Attribution-Share Alike 3.0 Unported  
[//creativecommons.org/licenses/by-sa/3.0/](https://creativecommons.org/licenses/by-sa/3.0/)

---