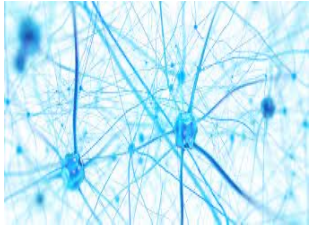


From Chain-of-Thought to LLM Powered Autonomous Agent

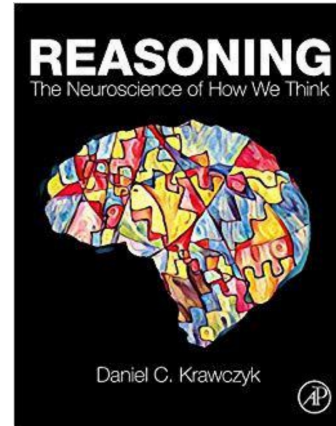
Shuofei Qiao
Zhejiang University



Cognitive Science



Negotiation



Medical Diagnosis



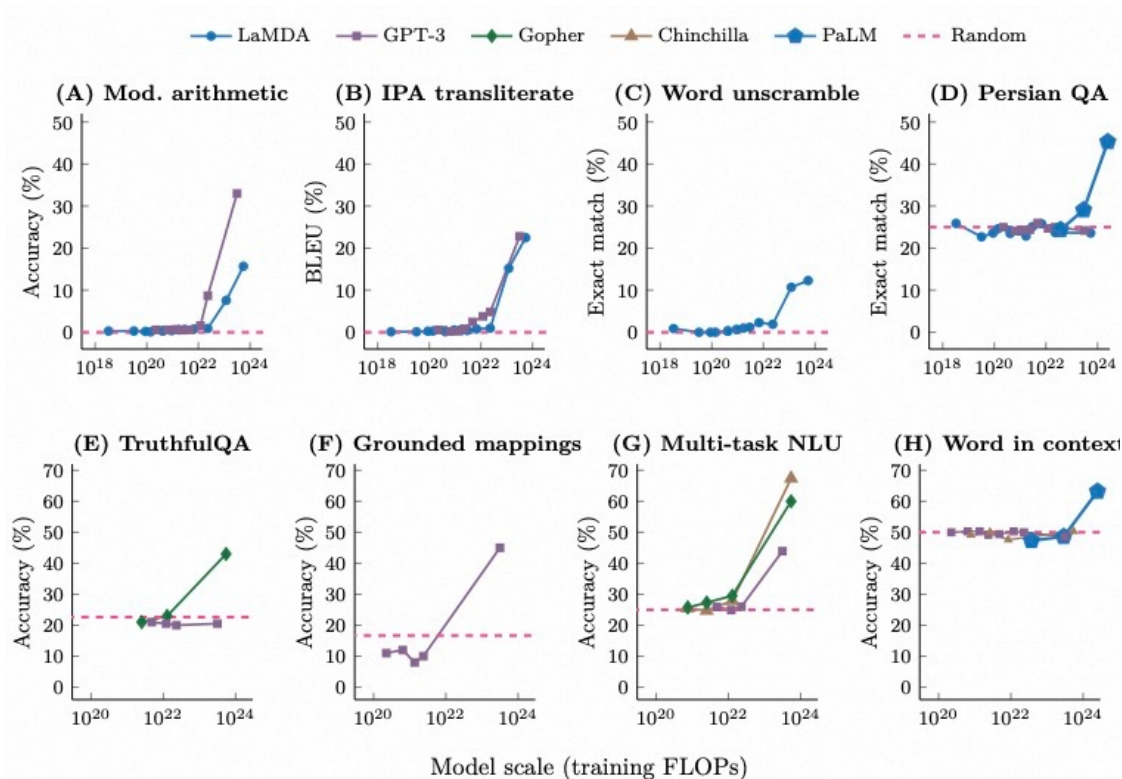
Brain Science



Reasoning is the cognitive process of drawing inferences or conclusions from observations, experiences, or information available to us. It involves the ability to analyze information, identify patterns and relationships, and make logical deductions based on those patterns and relationships.

—ChatGPT

Background Survey Agent Tool Future



Emergent Abilities

Emergent Abilities of Large Language Models, TMLR 2022

Chain of thought prompting elicits reasoning in large language models, NeurIPS 2022

Standard Prompting

Model Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The answer is 27. ❌

Chain-of-Thought Prompting

Model Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

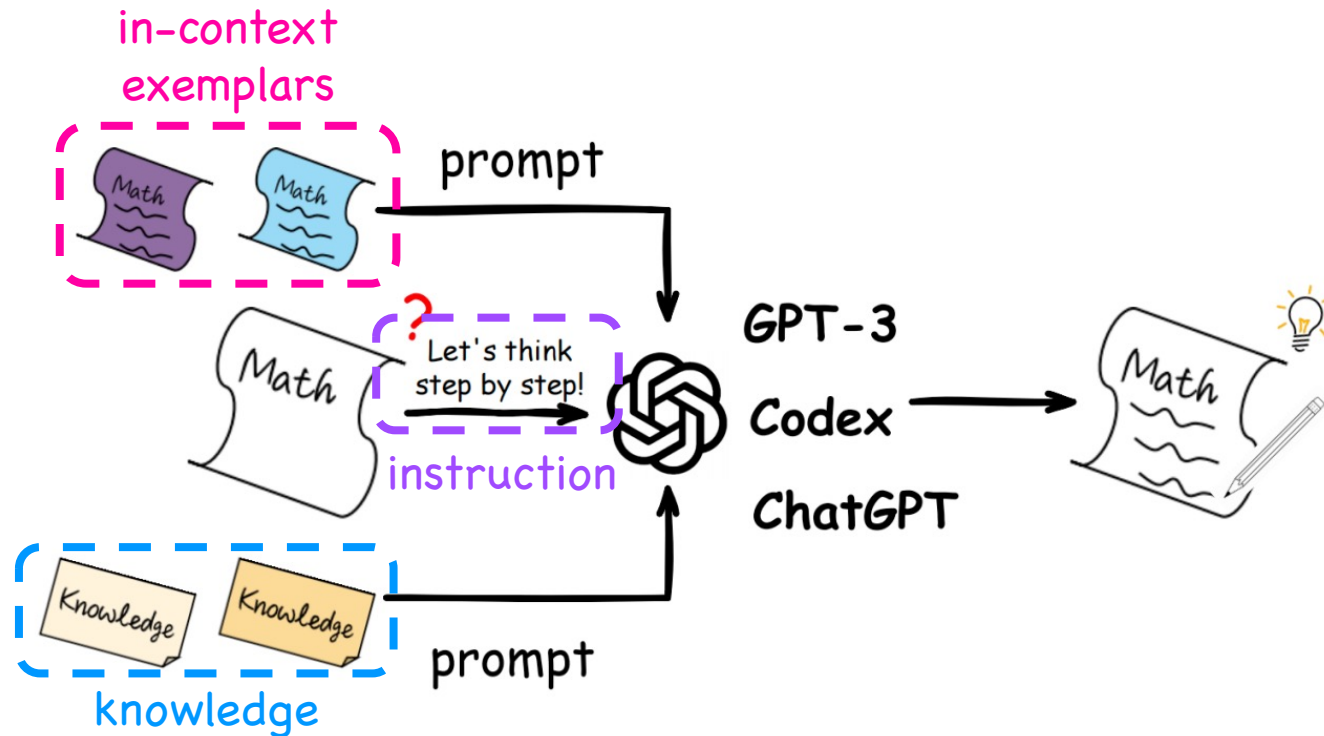
A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The cafeteria had 23 apples originally. They used 20 to make lunch. So they had $23 - 20 = 3$. They bought 6 more apples, so they have $3 + 6 = 9$. The answer is 9. ✅

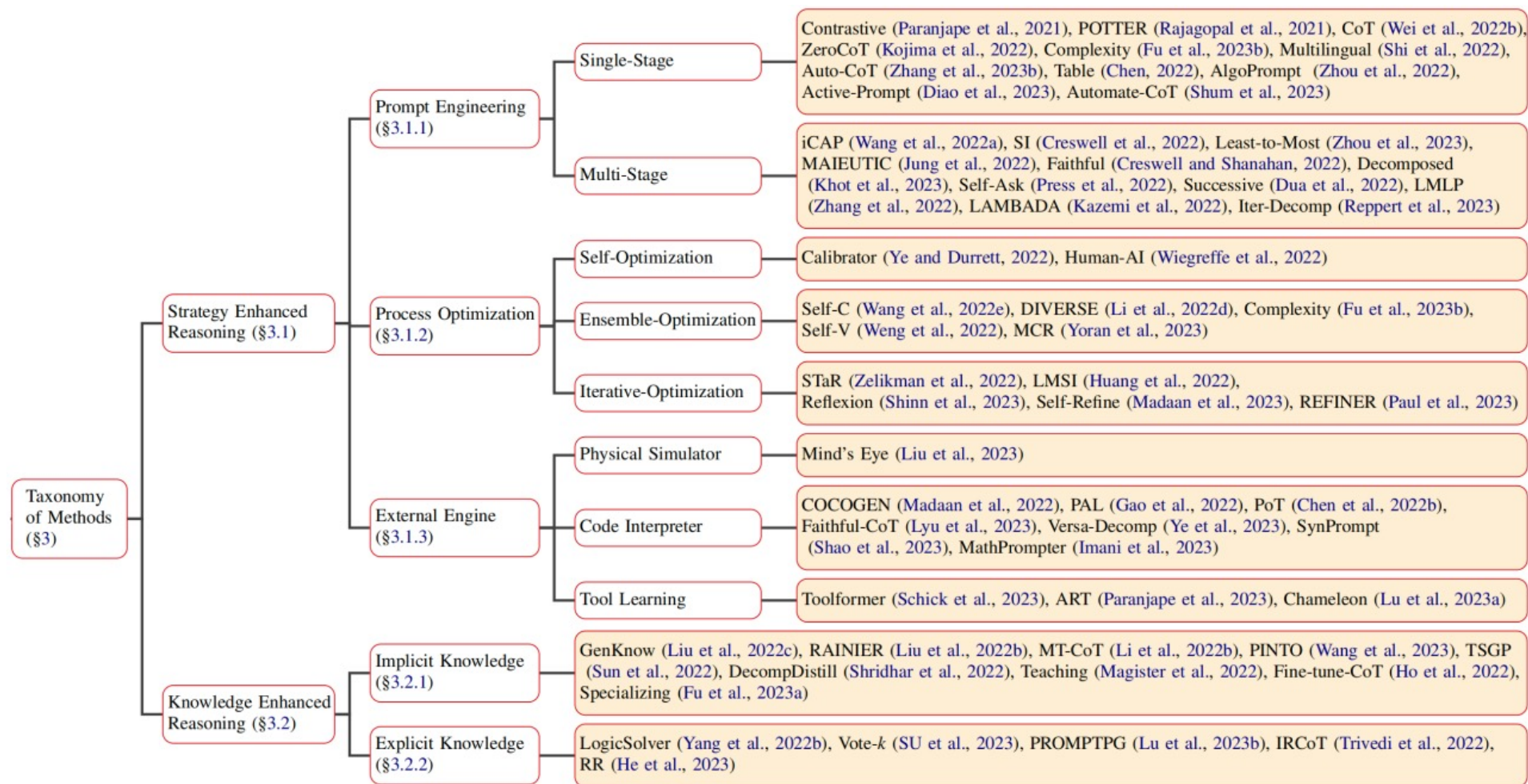
Chain of Thought

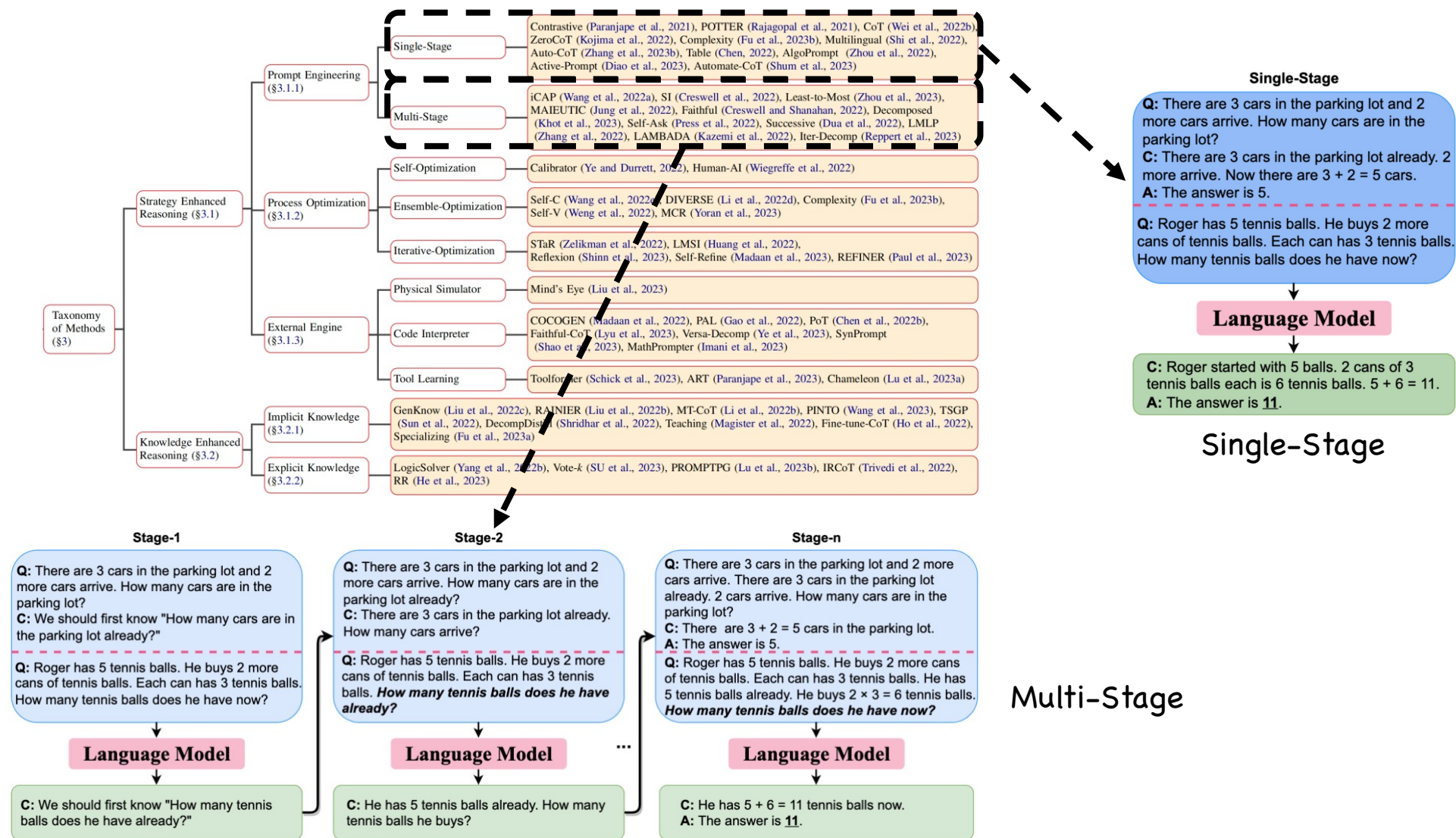


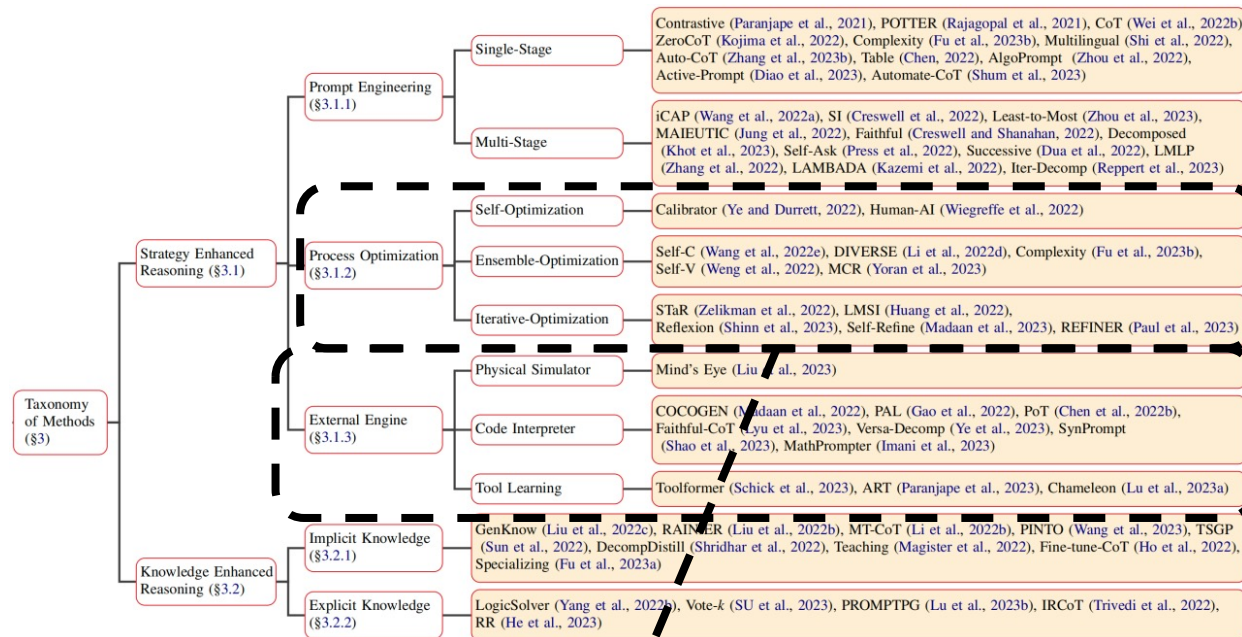
Main Contribution

- A new taxonomy of existing methods for reasoning with LM prompting
- In-depth comparisons and discussions
- Summarize benchmarks and resources for beginners
- Potential future directions

Background Survey Agent Tool Future







Simulator



→ **Prompt**

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

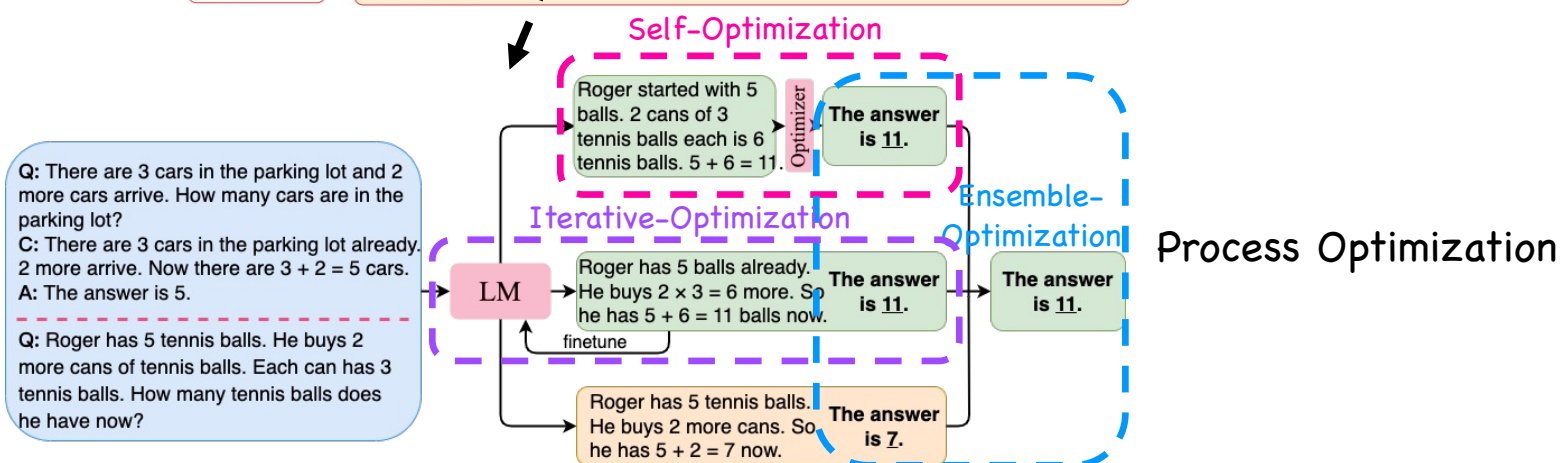
LM

Tools



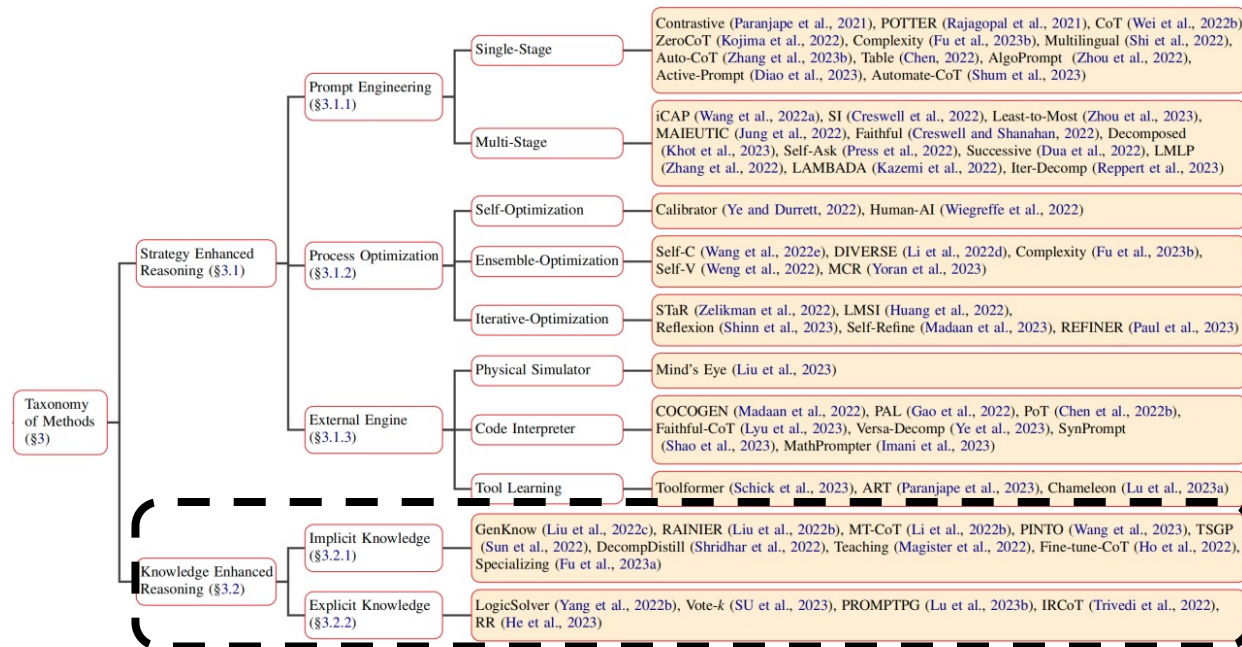
11

External Engine

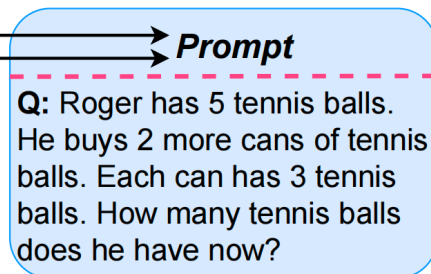


Q: There are 3 cars in the parking lot and 2 more cars arrive. How many cars are in the parking lot?
C: There are 3 cars in the parking lot already. 2 more arrive. Now there are $3 + 2 = 5$ cars.
A: The answer is 5.

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

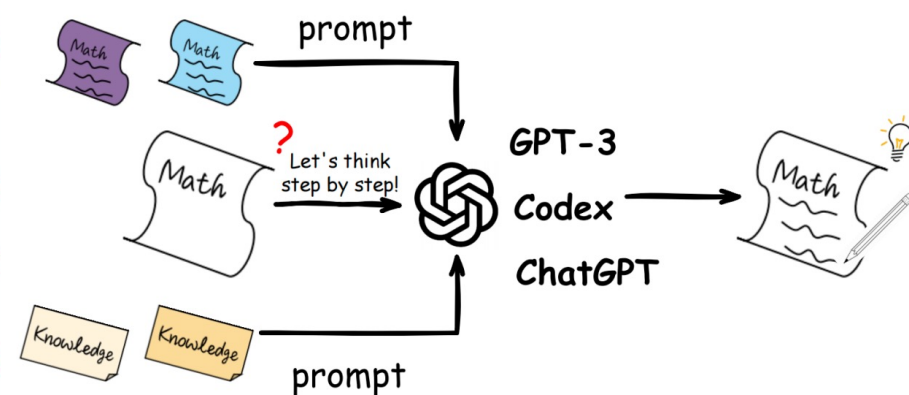


Implicit/Explicit Knowledge



LM

Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11.



- **ThoughtSource**: a central, open resource for data and tools related to chain-of-thought reasoning in LLMs.
- **LangChain**: a library designed to help developers build applications using LLMs combined with other sources of computation or knowledge.
- **LogiTorch**: a PyTorch-based library for logical reasoning on natural language.
- **λprompt**: a library that allows for building a full large LM-based prompt machines, including ones that self-edit to correct and even self-write their own execution code.
- **Promptify**: Prompt Engineering, Solve NLP Problems with LLM's & Easily generate different NLP Task prompts for popular generative models like GPT, PaLM, and more with Promptify.
- **EasyInstruct**: A package for instructing Large Language Models (LLMs) like GPT-3 in your research experiments. It is designed to be easy to use and easy to extend.



Paperlist



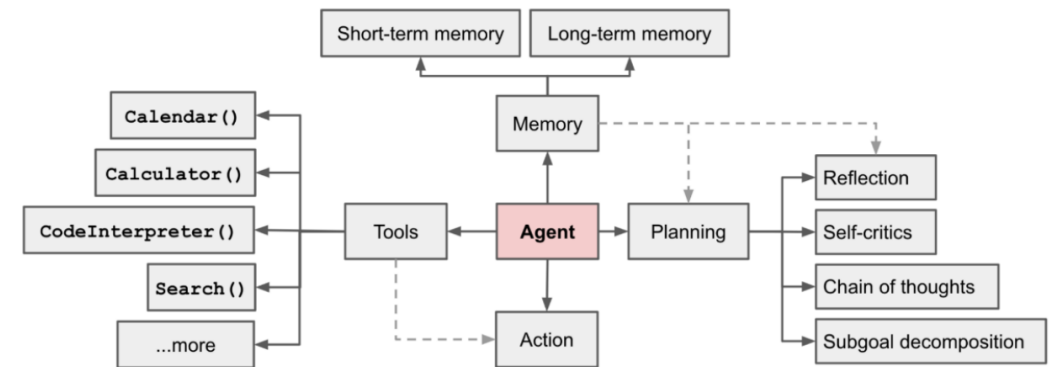
Preprint

Shortcomings of LLM

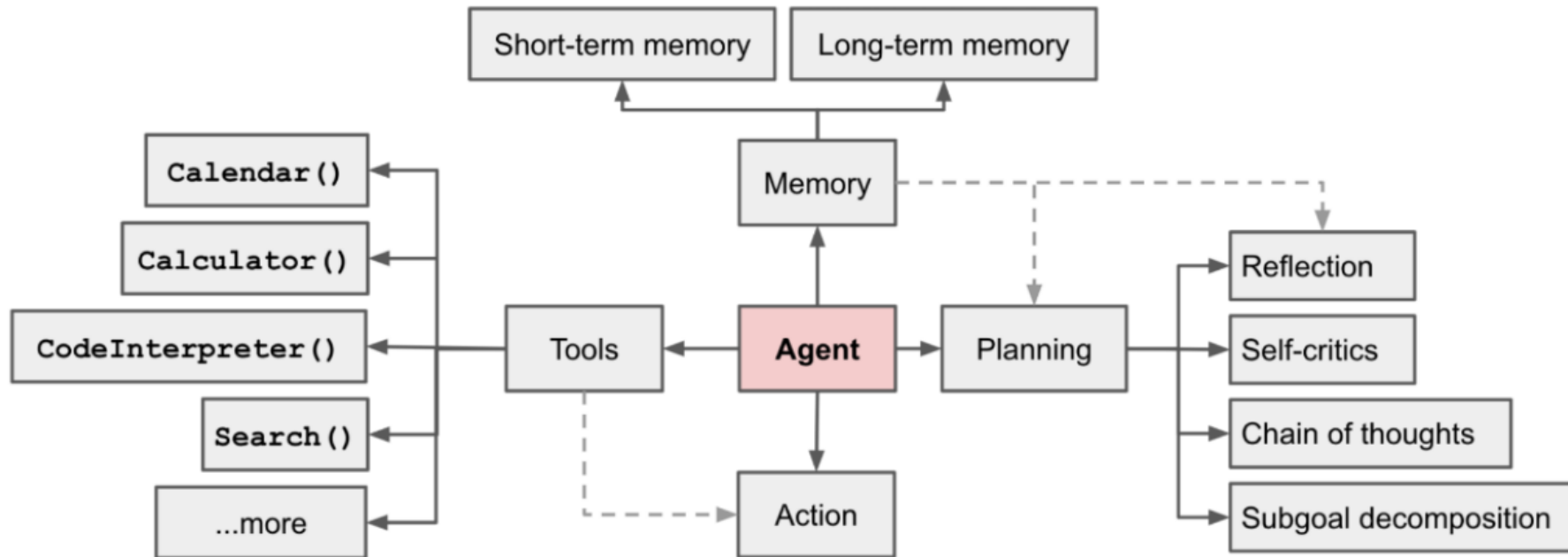
- Lack of ability to handle complex planning
- Limited contextual memory length
- Some basic abilities are weak



LLM Powered Autonomous Agents



Background Survey **Agent** Tool Future

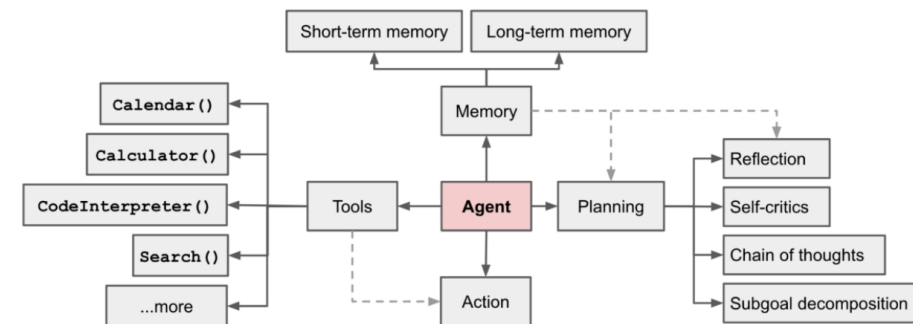
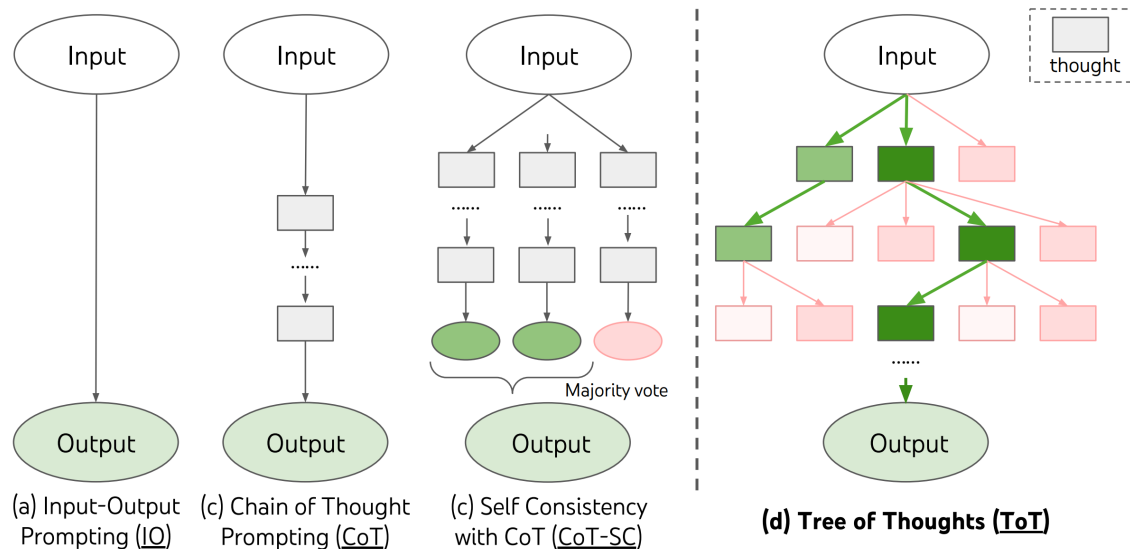


Andrew Karpathy, co-founder of OpenAI, said, “Compared to model training methods, OpenAI currently focuses more on changes in the agent field internally. **Whenever a new AI agent paper is published, the internal team is excited and discusses it seriously.**”

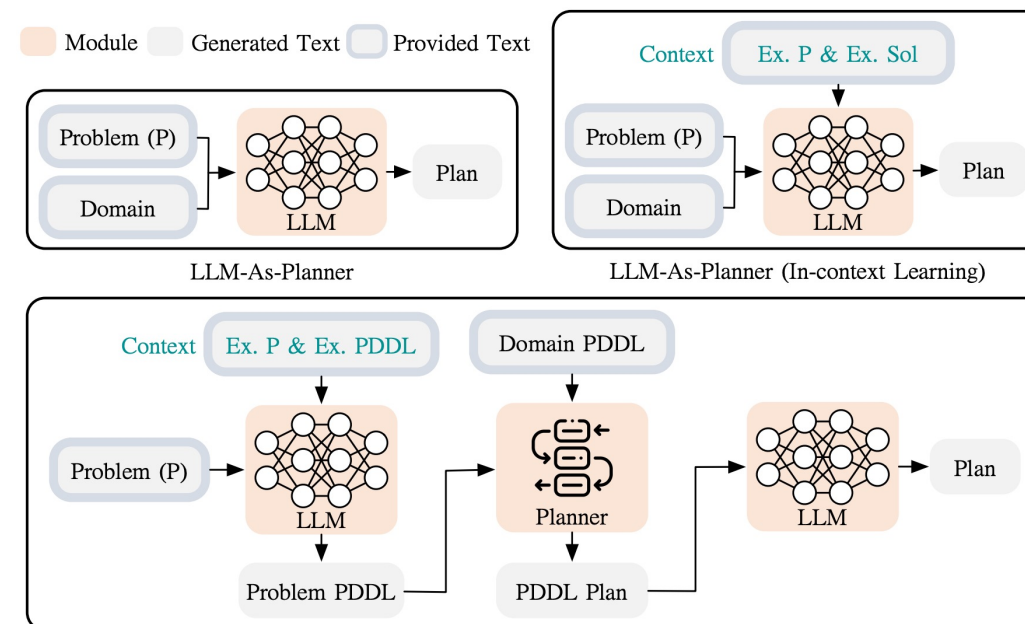
Task Decomposition

1. Chain-of-thought & Tree-of-Thoughts

- **CoT**: “think step by step” decompose hard tasks into smaller and simpler steps
- **ToT**: tree structure BFS or DFS (the search process)

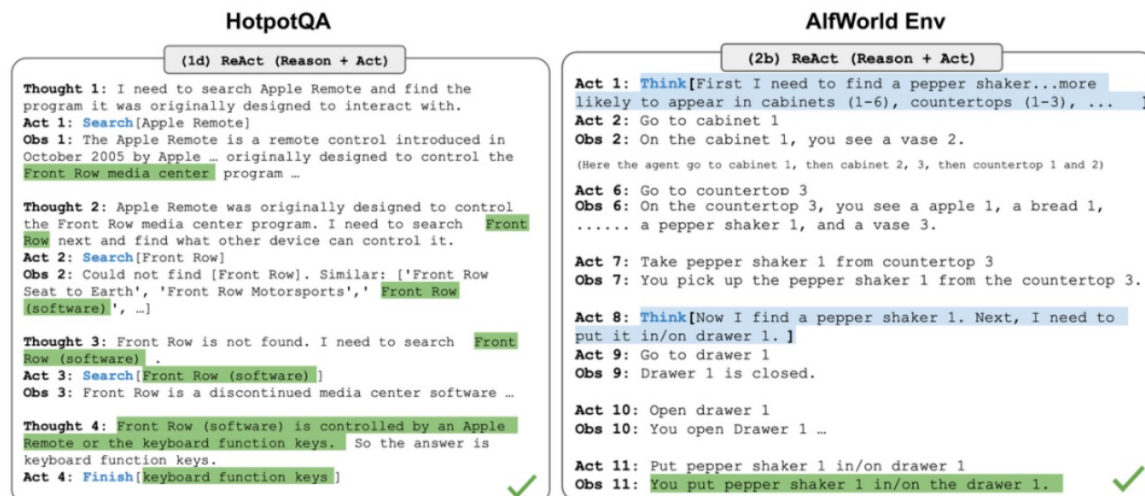


2. LLM+P

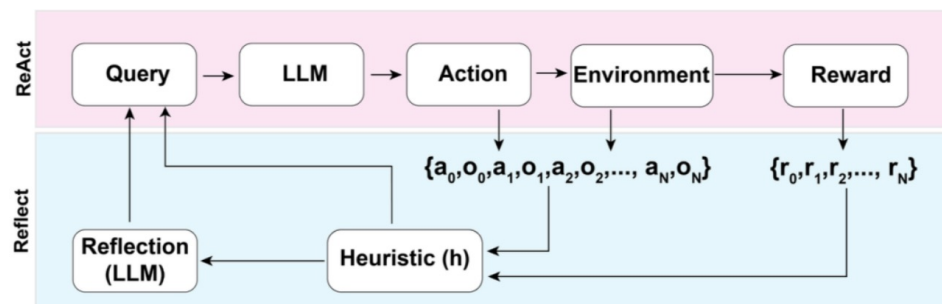


Self-Reflection

1. ReAct

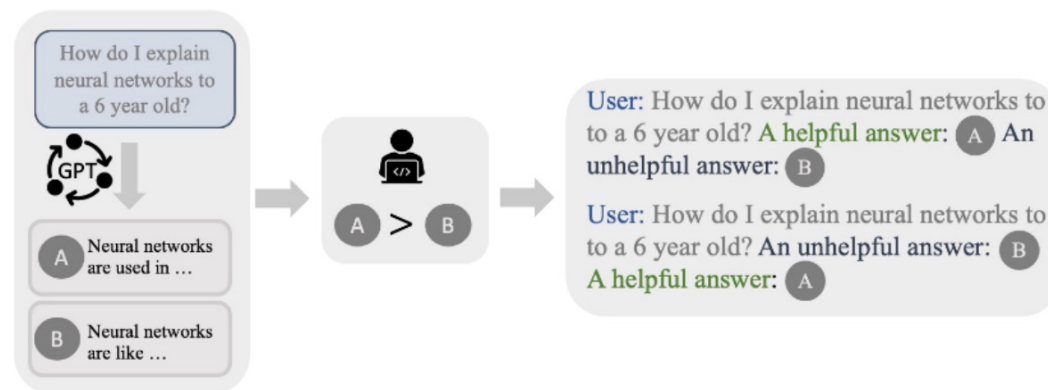


2. Reflexion



3. Chain of Hindsight

improve on its own outputs by explicitly presenting it with a sequence of past outputs, each annotated with feedback.

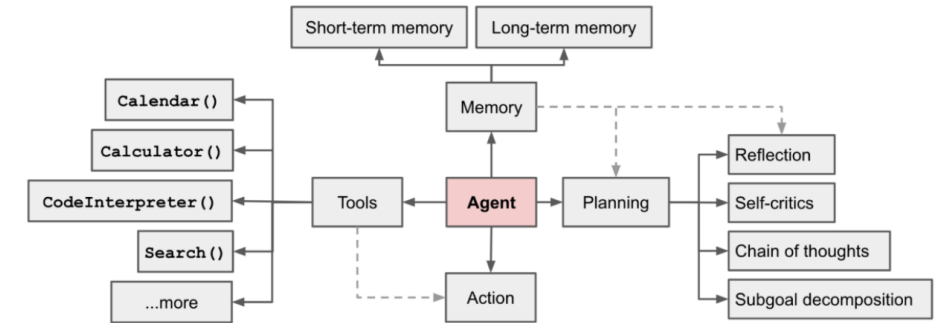
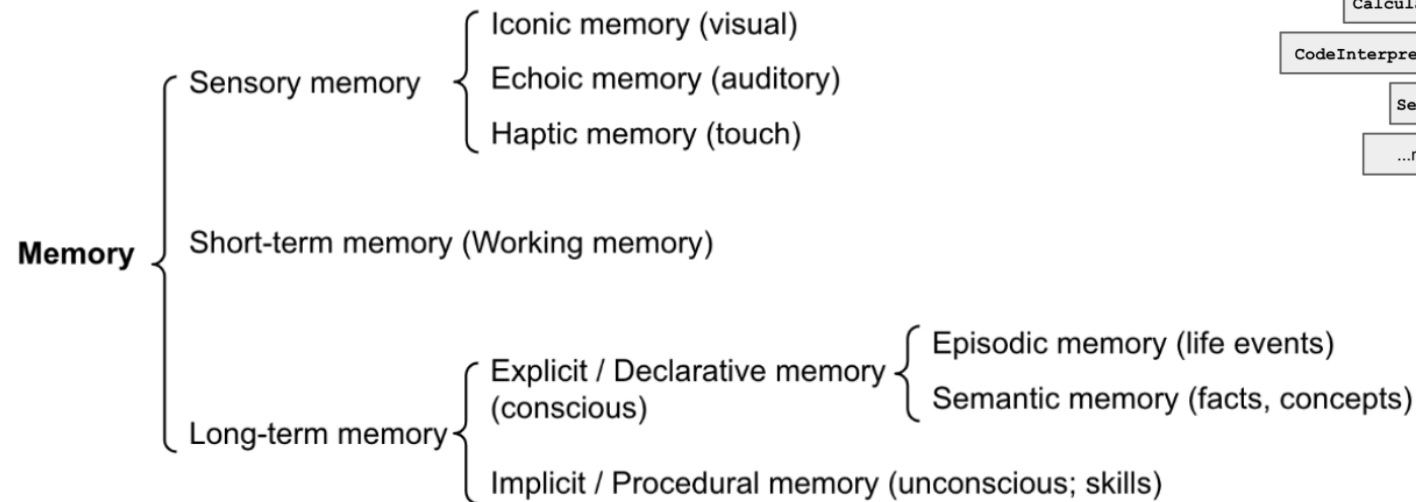


ReAct: Synergizing Reasoning and Acting in Language Models, ICLR 2023

Reflexion: Language Agents with Verbal Reinforcement Learning, 2023

Chain of Hindsight Aligns Language Models with Feedback, 2023

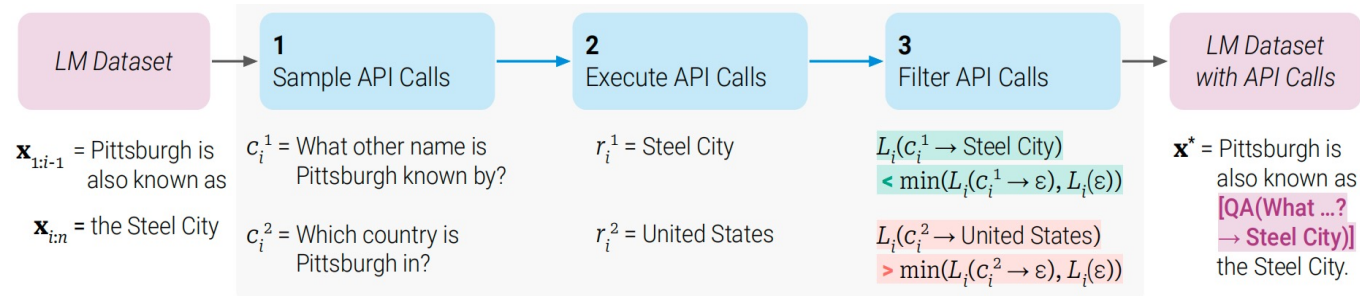
Types of Memory



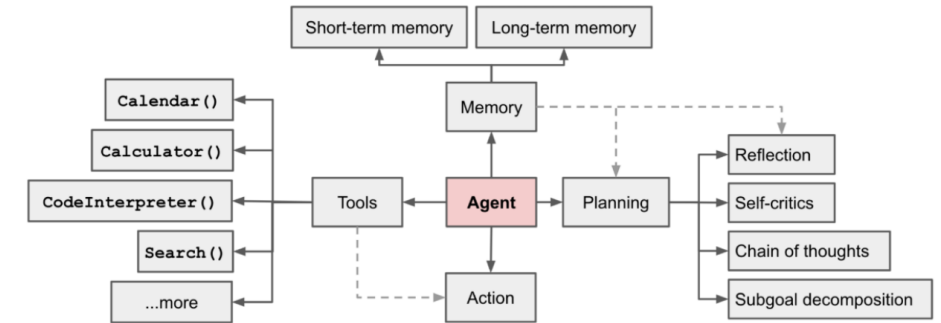
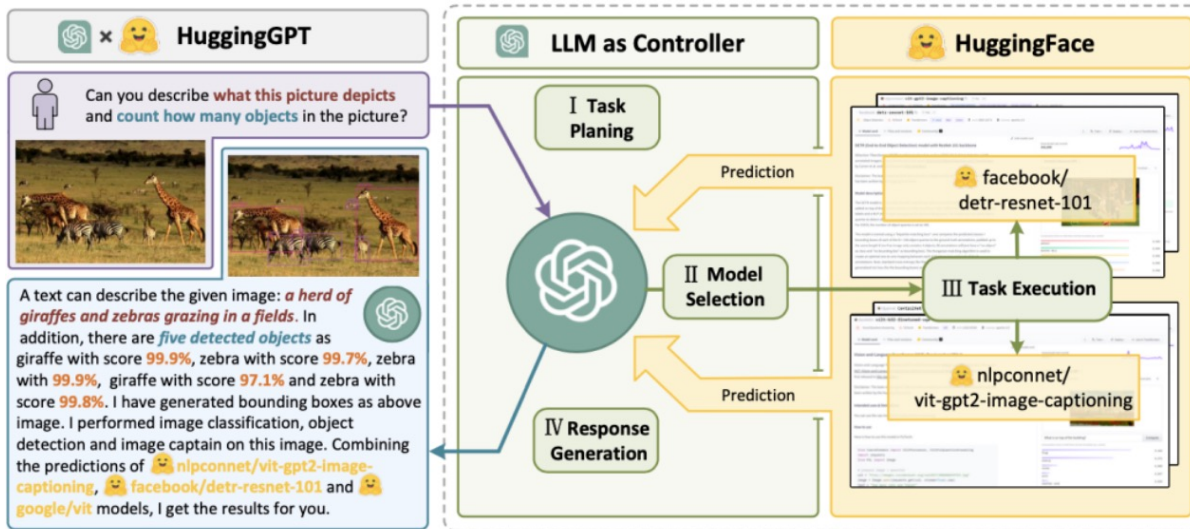
- **Sensory memory:** learning embedding representations for raw inputs, including text, image or other modalities
- **Short-term memory:** in-context learning. It is short and finite, as it is restricted by the finite context window length of Transformer.
- **Long-term memory:** the external vector store that the agent can attend to at query time, accessible via fast retrieval.

Use External Tools/API

1. Toolformer



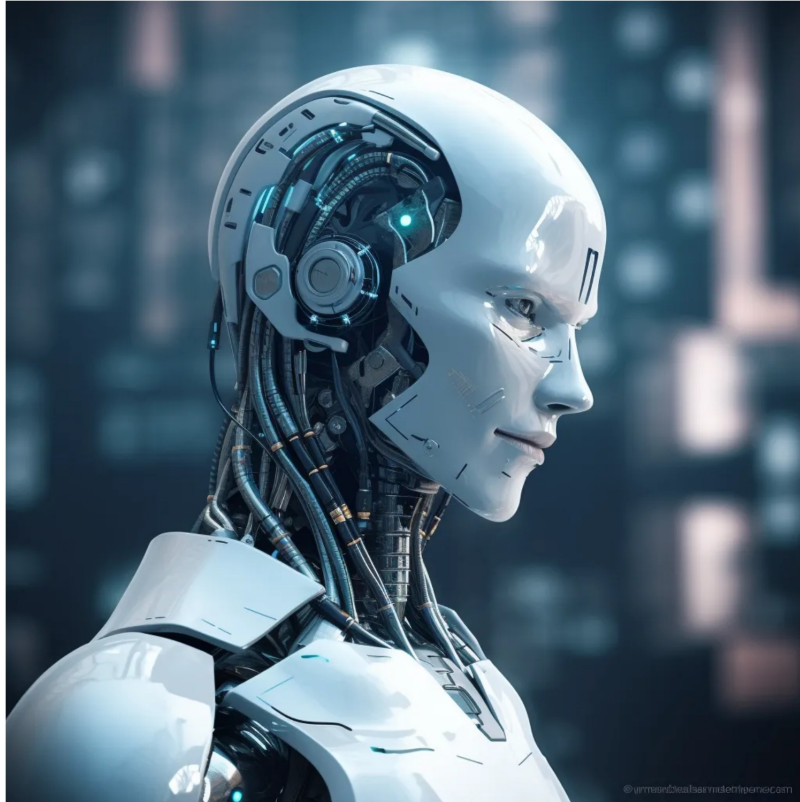
2. HuggingGPT



Toolformer: Language Models Can Teach Themselves to Use Tools, 2023

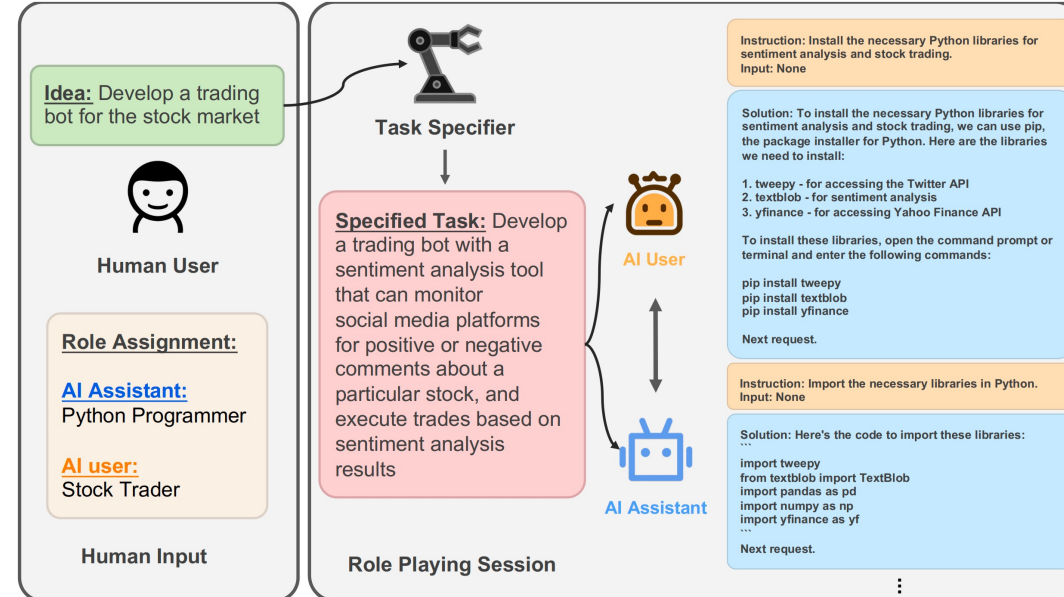
HuggingGPT: Solving AI Tasks with ChatGPT and its Friends in Hugging Face, 2023

1. Auto-GPT



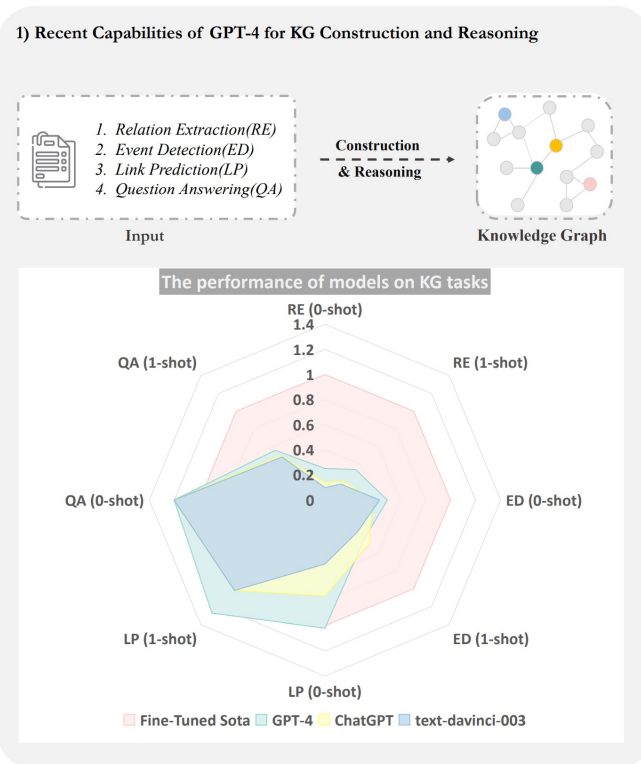
THE POWER OF AUTO-GPT: SUPERCHARGING YOUR TASKS AND PROJECTS

2. CAMEL

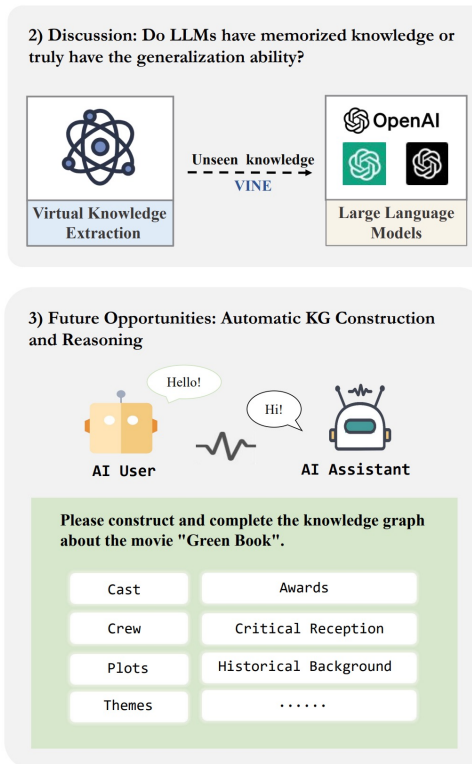


The task is made more specific using task specifier agent, leading to a well-defined task for the assistant to solve.

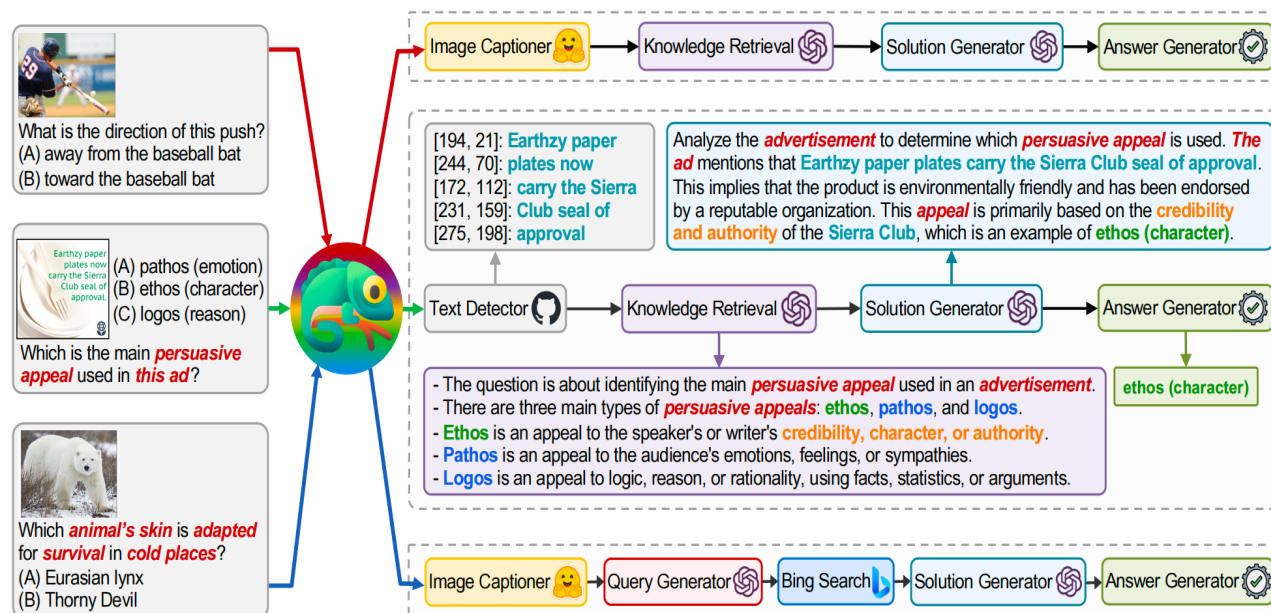
3. AutoKG



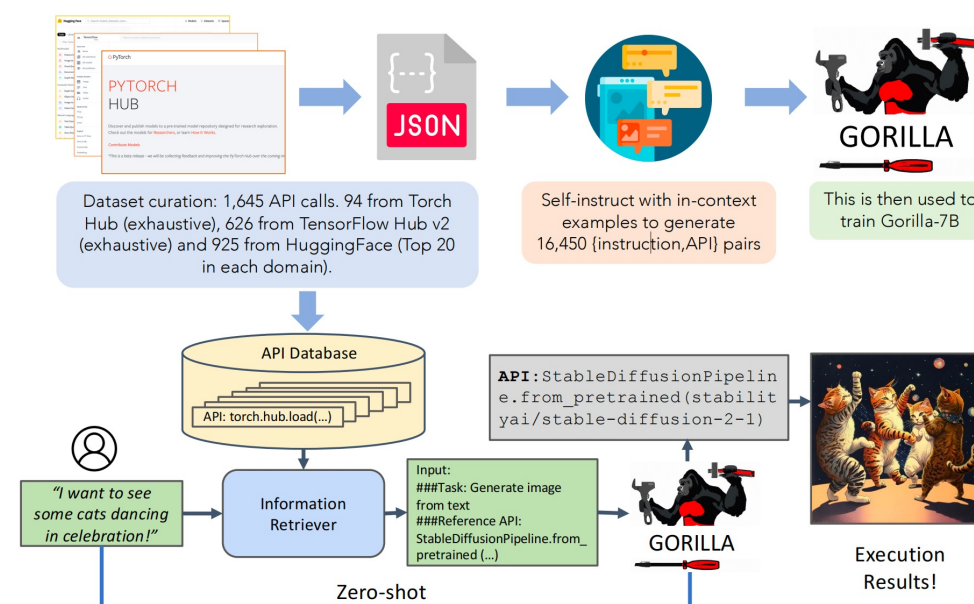
4. MachineSOM



5. Chameleon



6. Gorilla



Chameleon: Plug-and-Play Compositional Reasoning with Large Language Models, 2023

Gorilla: Large Language Model Connected with Massive APIs, 2023

The agent learns to call external APIs for extra information that is missing from the model weights (often hard to change after pre-training), including current information, code execution capability, access to proprietary information sources and more.

When and how to properly use which tools ?

General steps to use a tool:

1. Which tool to use ?
2. What information to give the tool ?
3. How to use the returned results of the tool ?



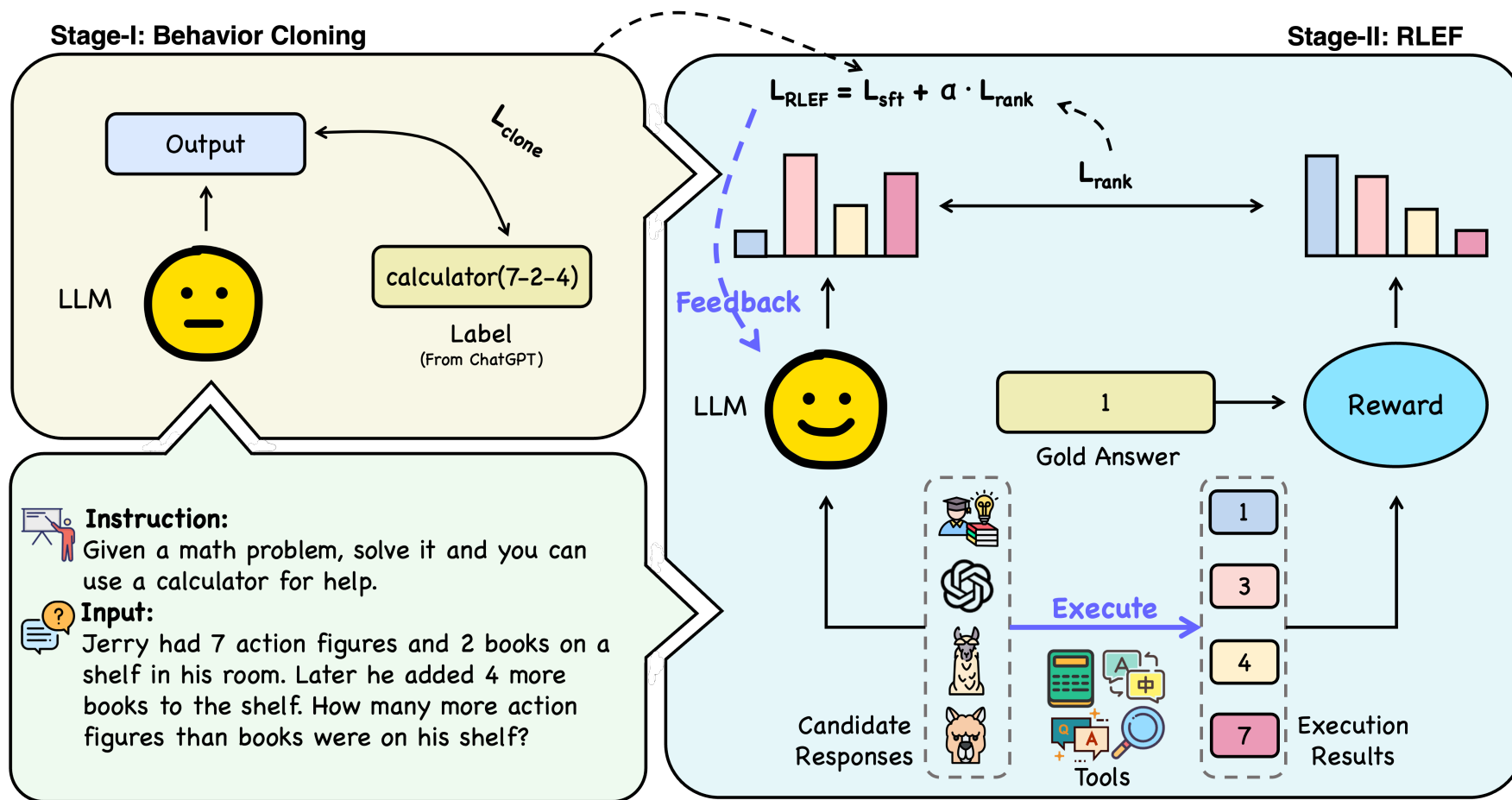
planning the complex
procedure

vs.

directly answer

label =

```
{ answer          use_tool = false
  tool_name(tool_input) use_tool = true }
```



framework of TRICE (Tool Learning with Execution Feedback)

Background Survey Agent Tool Future

main results

Model	Math Reasoning			Question Answering			LAMA	Multilingual QA	Avg.
	ASDiv	SVAMP	GSM8K	WebQ	NaturalQ	TriviaQA	T-REx	MLQA	
GPT-3.5	64.6	62.0	19.8	46.4	15.0	41.3	58.7	34.4	42.8
ChatGLM (Zero-Shot)	30.8	30.5	6.3	12.1	1.6	3.9	21.8	36.5	17.9
ChatGLM (TRICE-SPLIT)	72.9	64.0	12.4	15.2	9.6	15.2	32.7	37.3	↑14.7 32.6
ChatGLM (TRICE-MIX)	<u>75.6</u>	65.5	15.8	18.5	13.7	29.0	34.7	41.7	↑18.9 36.8
Alpaca (Zero-Shot)	31.2	22.0	3.5	32.8	5.3	15.0	39.7	37.7	23.4
Alpaca (TRICE-SPLIT)	73.4	45.0	16.3	38.2	18.6	37.8	54.6	48.2	↑18.1 41.5
Alpaca (TRICE-MIX)	75.2	58.0	<u>21.5</u>	41.4	<u>20.7</u>	<u>41.4</u>	<u>55.2</u>	52.0	↑22.3 45.7
Vicuna (Zero-Shot)	50.4	33.0	6.4	34.9	7.7	16.7	42.5	35.9	28.4
Vicuna (TRICE-SPLIT)	72.6	49.0	16.6	43.2	<u>20.7</u>	40.8	54.1	42.6	↑14.0 42.4
Vicuna (TRICE-MIX)	81.2	60.5	21.8	<u>44.1</u>	21.2	41.6	55.4	<u>49.7</u>	↑8.5 46.9

Table 3: Performance of TRICE across various tasks with different backbone models. **Zero-Shot:** models are directly evaluated without training. During this process, the model does not rely on tools. **TRICE-SPLIT:** models are trained separately for each task. **TRICE-MIX:** models are trained by combining training data from all tasks.

unseen

Model	Unseen Dataset		Unseen Tool	
	Calculator		QA Model	Retriever
	MultiArith	AddSub	SQuAD	HotpotQA
GPT-3.5	51.1	59.5	45.2	36.7
Vicuna (Zero-Shot)	42.3	44.1	28.6	19.7
Vicuna (TRICE-SPLIT)	↑20.8 <u>63.1</u>	↑31.1 <u>75.2</u>	↑2.3 30.9	—
Vicuna (TRICE-MIX)	↑24.3 66.6	↑36.4 80.5	↑7.1 <u>35.7</u>	↑7.6 <u>27.3</u>

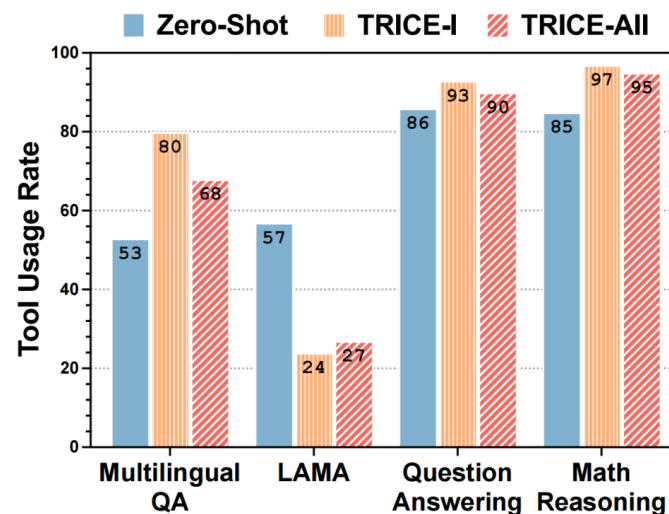
Table 5: Performance to unseen datasets and tools.

ablation study

Model	Math Reasoning			Question Answering			LAMA	Multilingual QA	Avg.
	ASDiv	SVAMP	GSM8K	WebQ	NaturalQ	TriviaQA	T-REx	MLQA	
Vicuna (Zero-Shot)	50.4	33.0	6.4	34.9	7.7	16.7	42.5	35.9	28.4
SPLIT	Vicuna (TRICE-I)	↑17.6 68.0	↑14.5 48.5	↑3.5 9.9	↑2.7 37.6	↑12.3 20.0	↑17.9 34.6	↑10.8 53.3	↑2.8 38.7
	Vicuna (TRICE-II)	↑2.6 53.0	↑3.5 36.5	↑0.7 7.1	↑6.9 41.8	↓1.0 6.7	↑4.6 21.3	↑9.4 51.9	↑3.8 32.2
	Vicuna (TRICE-ALL)	↑22.2 72.6	↑16.0 49.0	↑10.2 16.6	↑8.3 43.2	↑13.0 20.7	↑24.1 40.8	↑11.6 54.1	↑6.7 42.6
MIX	Vicuna (TRICE-I)	↑25.2 75.6	↑25.5 58.5	↑11.3 17.7	↑1.4 36.3	↑11.9 19.6	↑23.3 40.0	↑12.6 55.1	↑10.6 46.5
	Vicuna (TRICE-II)	↑2.3 52.6	↑5.0 38.0	↑10.2 16.6	↑6.0 40.9	↓0.9 6.8	↑4.6 21.3	↑8.0 50.5	↑4.2 40.1
	Vicuna (TRICE-ALL)	↑30.8 81.2	↑27.5 60.5	↑15.4 21.8	↑9.2 44.1	↑13.5 21.2	↑24.9 41.6	↑12.9 55.4	↑13.8 49.7

Table 4: Ablation Study: performance of TRICE across all tasks at different training stages. **TRICE-I**: only train by Behavior Cloning (instruction-tuning) stage. **TRICE-II**: only train by RLEF (reinforcement learning with execution feedback) stage. **TRICE-ALL**: train by both TRICE-I and TRICE-II.

selective tool usage



Background Survey Agent Tool Future

- Instruction: Given a math problem, please solve it and you can use a **Calculator** for help.
- Input: Diane wants to buy a package of cookies. The cookies cost 65 cents. Diane has 27 cents. How much more money does she need?
- Gold Answer: 38
- TRICE-I Response: Diane needs 37 cents to buy the package of cookies.
- **TRICE-I Pred:** 37 ❌
- TRICE-All Response: calculator(65-27)
- **TRICE-All Pred:** 38 ✔️

Case 1



- Instruction: Given a question, please answer it and you can use a **WikiSearch** for help.
- Input: Who plays Harley Quinn in the Lego Batman movie.
- Gold Answer: Jenny Slate
- TRICE-I Response: WikiSearch(Harley Quinn)
- **TRICE-I Pred:** Rosario Dawson ❌
- TRICE-All Response: WikiSearch(Harley Quinn Lego Batman Movie)
- **TRICE-All Pred:** Jenny Slate ✔️

Case 2



- Instruction: Given a question, please answer it and you can use a **QA model** for help.
- Input: Mongolia competed at the 1992 Winter Olympics in Albertville, ?
- Gold Answer: France
- TRICE-I Response: QA(Where were the 1992 Winter Olympics held?)
- **TRICE-I Pred:** bobsleigh ❌
- TRICE-All Response: France
- **TRICE-All Pred:** France ✔️

Case 3



- Instruction: Given a context, please answer the question in English and you can use a **Translator** for help.
- Input: There are a number of risk factors for fungal meningitis, including the use of and the loss of immunity associated with aging. 什么与丧失免疫力有关 ?
- Gold Answer: aging
- TRICE-I Response: translator(什么与丧失免疫力有关 ?)
- **TRICE-I Pred:** immunity ❌
- TRICE-All Response: translator(什么与丧失免疫力有关 ?)
- **TRICE-All Pred:** aging ✔️

Case 4



- Instruction: Given a math problem, please solve it and you can use a calculator for help.
- Input: Bryan took a look at his books and magazines. If he has 27 books and 80 magazines in each of his 23 bookshelves How many books does he have in total?
- Gold Answer: 621
- TRICE-I Response: calculator(27+80*23)
- **TRICE-I Pred:** 1867 ❌
- TRICE-All Response: calculator((27+80)*23)
- **TRICE-All Pred:** 2461 ❌

Case 5



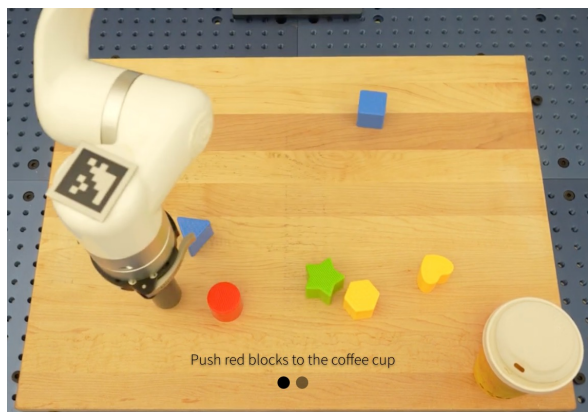
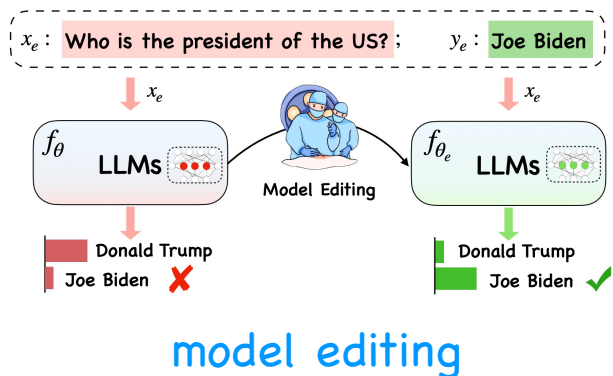
case study



Project: <https://zjunlp.github.io/project/TRICE/>

Code: <https://github.com/zjunlp/TRIC>

- We focus on addressing the challenge of **selective utilization of tools** by LLMs and propose **a two-stage end-to-end training framework dubbed TRICE** to make LLMs better tool learners with execution feedback.
- Through comprehensive experiments on various tasks and models, we have shown that our method can **achieve better performance compared to the zero-shot manner and GPT-3.5**.
- Extensive analysis illustrates that TRICE can selectively use tools by **improving the accuracy of tool usage while enhancing insufficient tool learning and mitigating excessive reliance on tools**.

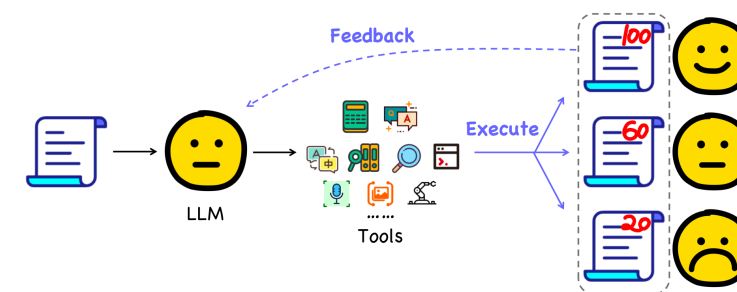


interact with environment

- Theoretical Principle of Reasoning
- Efficient Reasoning
- Robust, Faithful and Interpretable Reasoning
- Interactive Reasoning
- Generalizable (True) Reasoning



interact among multi-agent

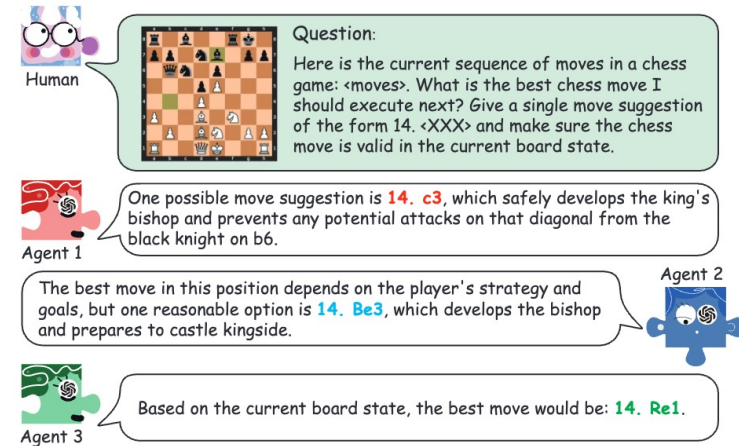
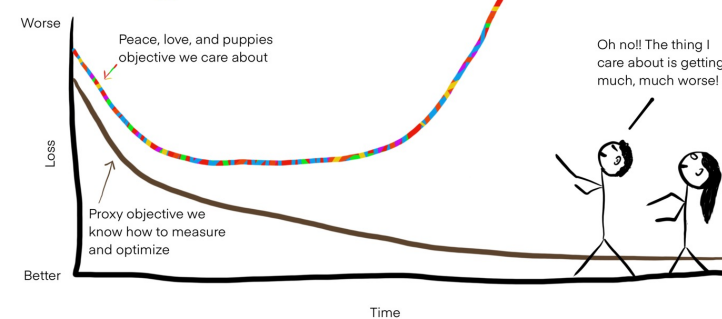


interact with tools

Reasoning with Language Model Prompting: A Survey, ACL 2023
 Editing Large Language Models: Problems, Methods, and Opportunities, 2023
 PaLM-E: An Embodied Multimodal Language Model, 2023
 Training Socially Aligned Language Models in Simulated Human Society, 2023
 Making Language Models Better Tool Learners with Execution Feedback, 2023

- Why multi-agents?
- Goodhart's Law - The better on object A, the worse on many other objects B
- What do agents interact with?
- Knowledge boundary, Brain in a Vat
- What is the preferred method of communication among agents?
- Natural language or Code
- How to communicate between agents?
- Roles, Society, Behaviors

Strong version of Goodhart's law



Training Socially Aligned Language Models in Simulated Human Society, 2023

Investigating the Factual Knowledge Boundary of Large Language Models with Retrieval Augmentation, 2023

Brain in a Vat: On Missing Pieces Towards Artificial General Intelligence in Large Language Models, 2023

PAL: Program-aided Language Models, 2023

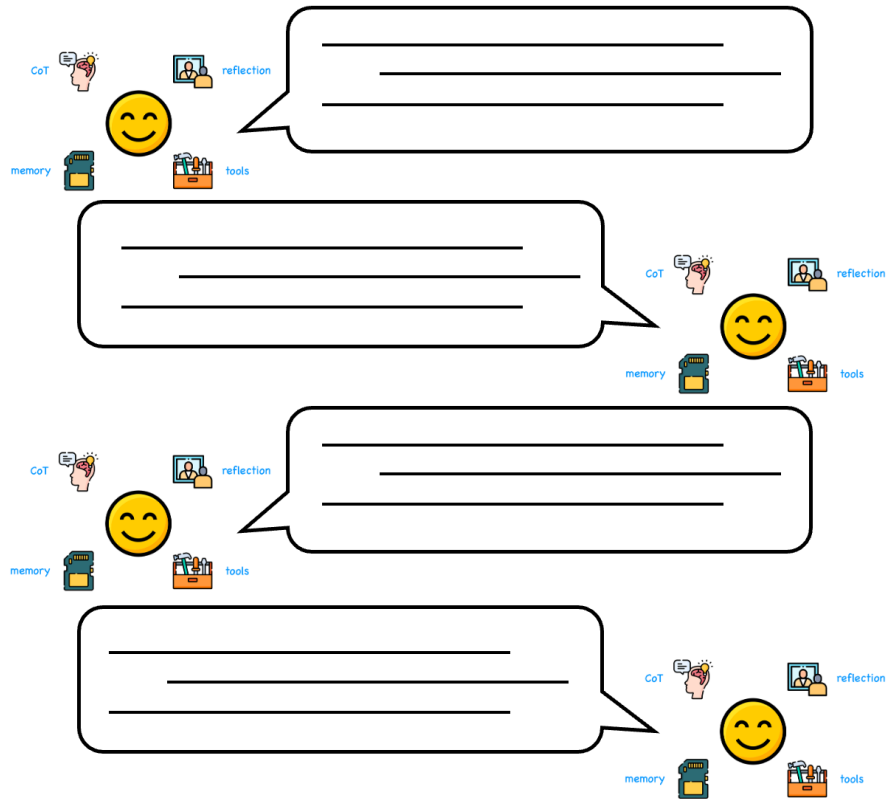
Encouraging Divergent Thinking in Large Language Models through Multi-Agent Debate, 2023



Background Survey Agent Tool Future



Background Survey Agent Tool Future



multi-agent
debate



AI+X



+

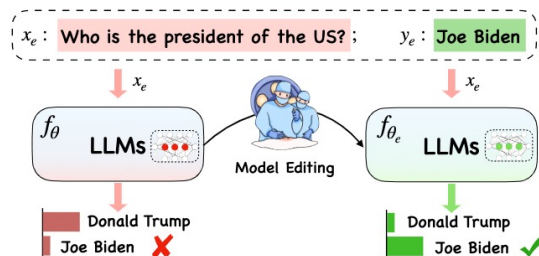


+

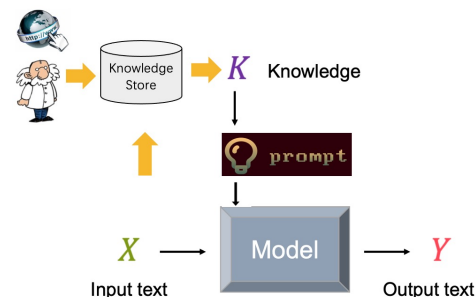


Future ?

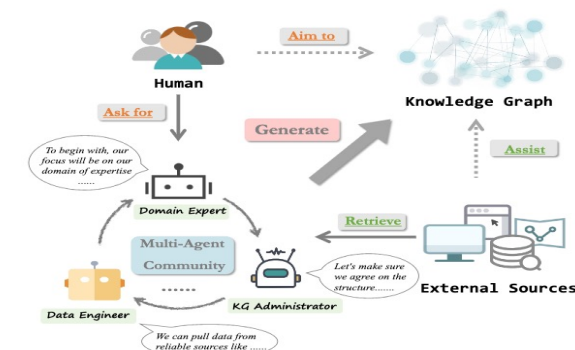
Knowledge Editing



Knowledge Prompt



Knowledge Interaction



<https://github.com/zjunlp/KnowLM>

Information
Extraction

Reasoning

Open-sourced
Pre-training

Efficient
Fine-tuning

Fast
Deployment

Thanks