

Data-Collection Framework for LoGU

Query q: Could you tell me what you know about Annabel Pitcher?

Revision Operator Γ_{op}

Γ_{pos} (Positive operation)

For **refuted** atomic claims:

1. Select a part of claims
2. Revise the selected claims to uncertainty.

Remain the **supported** atomic claims unchanged

Γ_{neg} (Negative operation)

For **supported** atomic claims:

1. Select a part of claims
2. Revise the selected claims to uncertainty

Remain the **refuted** atomic claims unchanged

Model Response R : Annabel Pitcher is a British author. She was born on August 15, 1974, in London, England. [...]

Fact-Checking



Atomic Claim 1: Annabel was born in London, England.



Atomic Claim 2: Annabel was born on August 15, 1974.

Revision

Γ_{op} (Atomic Claims), $op = \{pos, neg\}$

Assembling

R_{pos} : Annabel Pitcher is a British author. She was born in London, England, **but I am not sure when she was born.** [...]

R_{neg} : Annabel Pitcher is a British author. She was born on August 15, 1974, **but I am not sure where she was born.** [...]

Sub-Challenges

Uncertainty Suppression

Annabel Pitcher [...] **best known for her young adult and children's novels.** **She was born on August 15, 1974,** in London, England. [...]

Uncertainty Misalignment

Annabel Pitcher [...] I am not sure what she was best known for. And I am not sure when or **where she was born.** [...]

Target

Two-Stage Training

Stage I: LoGU-SFT

☆ Using $\{q, R_{pos}\}$

Stage II: LoGU-DPO

☆ Using preference pairs

- $\{q, R_{pos}, R\}, \{q, R_{pos}, R_{neg}\}, \{q, R, R_{neg}\}.$