Stage1 Data collection

QA Collection

 $\begin{array}{ccc} \overset{\circ}{\rightarrow}\overset{\square}{\rightarrow}\overset{\wedge}{\triangle} & \textbf{Logic} \\ \overset{\circ}{\rightarrow}\overset{\wedge}{\rightarrow} & \textbf{Logic} & \text{two of three girls from Fang ...} \end{array}$

Math Compute 15 + 8² - 3.

Code Your are an expert Python programmer, here is your task...

Commen A: May we see the wine list please.
B: Sure. Our special wine today ...

Total 10 tasks, 17 datasets



CoTs with Error Answers

Thought: Since Tang are not selected, we knowAnswer: D

Math Thought: Firstly calculate the 15+8 = 23, then....Answer: 44

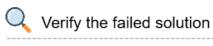
Thought: I need to understand the operation p times ... Answer: class Pair(): def_init_....

Commen Thought: The conversation id about ordering a bottle of wine, ... Answer: B

•••••

Stage2 Reflection generation

Reflection Instructions



Locate errors & diagnose potential reasons





Self-Reflection

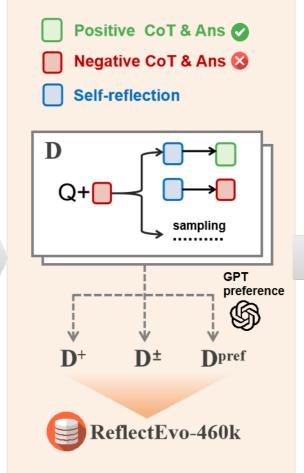
Diagnosis: A possible reason for my failure.. not follow the proper operation sequence ...

New plan: 1) Evaluate the value inside the exponentiation...2) Perform the addition operation by substitue 64 in the equation

Self-Correction

Thought: By following the plan, i need to... Answer: The result is 76.

Stage3 Reflection curation



Stage4 Reflection tuning



