Multi-agent Evaluation Stage

Agent₇ *****

Agent₆ * — CREATE

Agent₅ * — EVALUATE

Agent₄ * — ANALYZE

Agent₃ * — APPLY

Agent₂ * — UNDERSTAND

Agent₁ * — REMEMBER

Agent Feedback & Ratings

Improvement Suggestions:
1: "The new problem attempts to apply mathematical concepts related to ...",
2: "Ensure that the conditions provided in the problem lead to a solvable equation.",
3: "Consider a scenario where a student is calculating their final grade based on a new test score, with realistic averages."

*****
Agent₇

+

Performance Score : xx
Confidence Score : xx

* * * ... * *****
Agent₁ Agent₂ Agent₃ ... Agent₆ Agent₇

Succeed Problem

Q: A young chef is preparing for a cooking competition. The chef has received scores of 75, 80, and 85 in the preliminary rounds ...

A: No able to sovle it. If the fourth score is x, then (90+x) / 2 = 88. Solving for y, we get 85 + y = 164, so y = 79. Now, the average of all four ...

Passed

Checking threshold

SCORE → ← δ = 0.85

Failed

Previous Q & A

Improvement Suggestion

Conversation History

Think-Aloud "FIVE KEYS"

Flawed Math problems

Q: A student has test scores of 80, 85, 90 in math. The teacher says, the average of your last two...

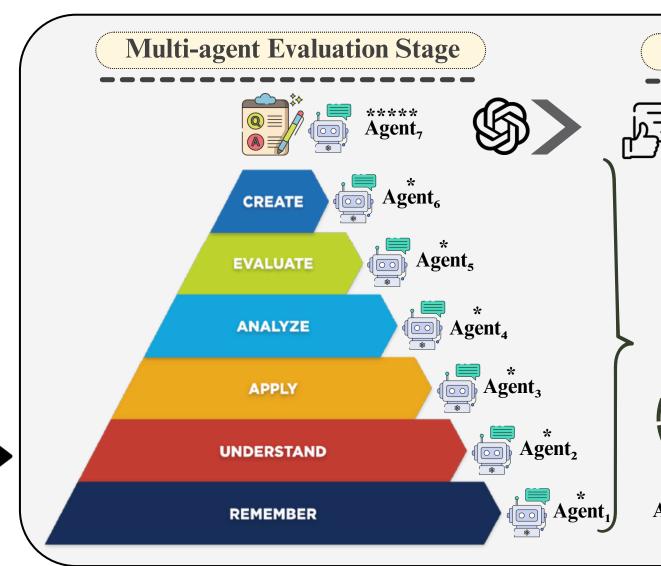**A:** No able to sovle it. If the fourth score is x, then (90+x) / 2 = 88 ...

GPT - 4o

Legends of Contents

= Waited for Evaluation

= Quality Score < δ

= Quality Score > δ

New Problem & Solution

Q: A student is preparing for a final math exam. The student has test scores of 80, 85, and 90...

A: Let the score of the final exam be x. The average of the last two tests is given by (90 + x) / 2 = 87 ...

Question Generation

LLMs