Self-Reflection in LLM Agents:
Effects on Problem-Solving Performance

Matthew Renze and Erhan Guven

Johns Hopkins University

Can an LLM reflect on its own chain of thought?

Problems with LLMs

Problems with LLMs

Limited knowledge

Problems with LLMs

Limited knowledge

Errors in reasoning

Problems with LLMs

Limited knowledge

Errors in reasoning

Model hallucination

Problems with LLMs

Limited knowledge

Errors in reasoning

Model hallucination

Unproductive loops

Potential Solutions

Self-reflection

Potential Solutions

Self-reflection

Use errors as feedback

Potential Solutions

Self-reflection

Use errors as feedback

Detect and correct CoT errors

Self-Reflection in Large Language Model Agents: Effects on Problem-Solving Performance

1st Matthew Renze
Johns Hopkins University
Baltimore, MD, USA
mrenze1@jhu.edu

2nd Erhan Guven
Johns Hopkins University
Baltimore, MD, USA
eguven2@jhu.edu

***Abstract*—In this study, we investigated the effects of self-reflection in large language models (LLMs) on problem-solving performance. We instructed nine popular LLMs to answer a series of multiple-choice questions to provide a performance baseline. For each incorrectly answered question, we instructed eight types of self-reflecting LLM agents to reflect on their mistakes and provide themselves with guidance to improve problem-solving. Then, using this guidance, each self-reflecting agent attempted to re-answer the same questions. Our results indicate that LLM agents are able to significantly improve their problem-solving performance through self-reflection ($p < 0.001$). In addition, we compared the various types of self-reflection to determine their individual contribution to performance. All code and data are available on GitHub at https://github.com/matthewrenze/self-reflection**

***Index Terms*—self-reflection, large language model, LLM, agent**

observed successfully using tools, including web browsers, search engines, code interpreters, etc. [10], [19]–[21].

However, these LLM agents have several limitations. They have limited knowledge, make errors in reasoning, hallucinate output, and get stuck in unproductive loops [4]–[9].

To improve their performance, we can provide them with a series of cognitive capabilities. For example, we can provide them with a CoT [1]–[3], access to external memory [22]–[25], and the ability to learn from feedback [10], [18], [19].

Learning from feedback can be decomposed into several components. These components include the source of the feedback, the type of feedback, and the strategy used to learn from feedback [11]. There are two sources of feedback (i.e., internal or external feedback) and two main types of feedback

Background

Background

Humans can reflect on thoughts

Background

Humans can reflect on thoughts

LLMs generate chain of thought

Background

Humans can reflect on thoughts

LLMs generate chain of thought

Can an LLM reflect on its CoT?

Background

Humans can reflect on thoughts

LLMs generate chain of thought

Can an LLM reflect on its CoT?

What types of reflections help?

Prior Literature

Prior Literature

Learning from feedback

Sources:

Madaan, et al. (2023)
Self-refine: Iterative refinement with self-feedback

Pan, et al. (2023)
Automatically correcting LLMs: Surveying the landscape of diverse self-correction strategies

Shinn, et al. (2023)
Reflexion: Language agents with verbal reinforcement learning

Yao, et al. (2022)
React: Synergizing reasoning and acting in language models

Prior Literature

Learning from feedback

Multi-model debate

Sources:

Pan, et al. (2023)
Automatically correcting LLMs: Surveying the landscape of diverse self-correction strategies

Prior Literature

Learning from feedback

Multi-model debate

Self-correction

Sources:

Pan, et al. (2023)
Automatically correcting LLMs: Surveying the landscape of diverse self-correction strategies

Prior Literature

Learning from feedback

Multi-model debate

Self-correction

Introspection

Ji, et al. (2023)
Towards mitigating hallucination in large language models via self-reflection

Madaan, et al. (2023)
Self-refine: Iterative refinement with self-feedback

Shinn, et al. (2023)
Reflexion: Language agents with verbal reinforcement learning

Toy, et al. (2024)
Metacognition is all you need? using introspection in generative agents to improve goal-directed behavior

Wang and Zhao (2023)
Metacognitive prompting improves understanding in large language models

Methods

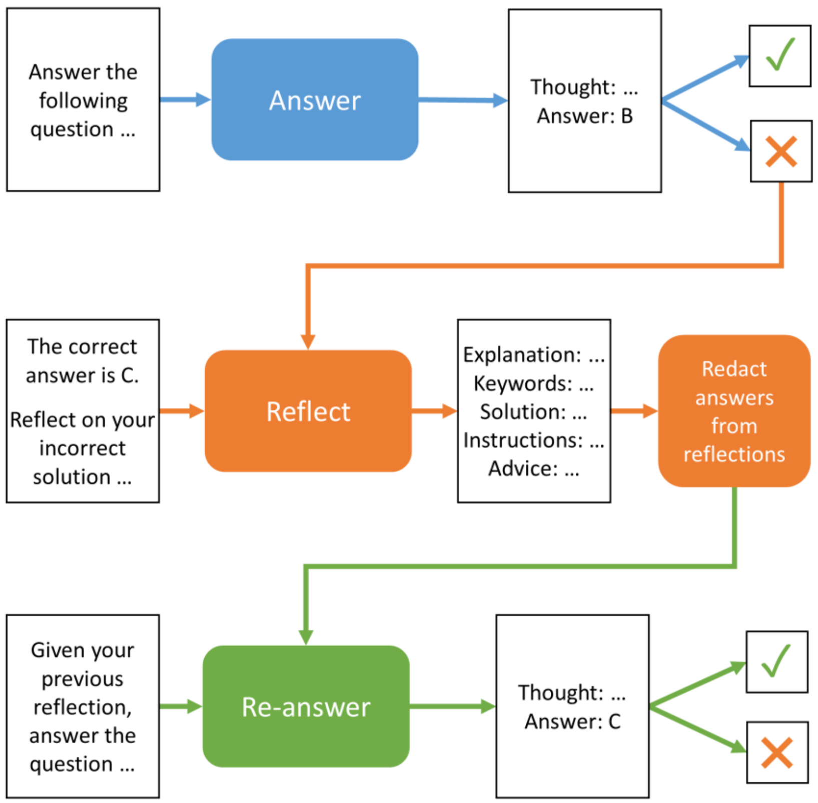
Experiment

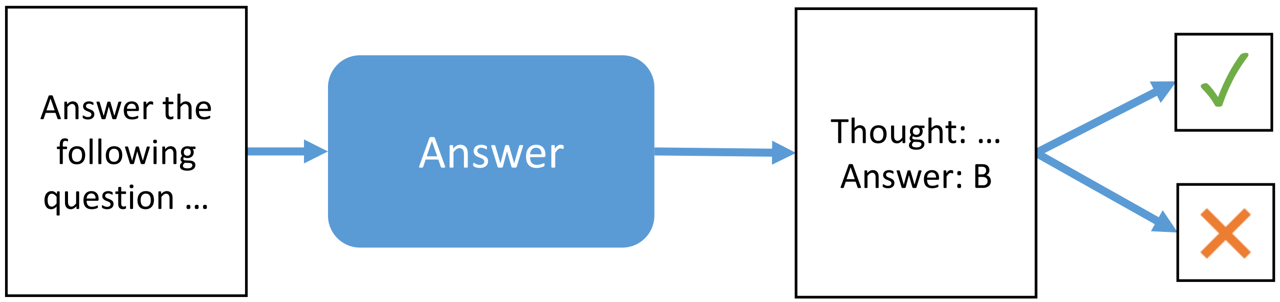
9 models

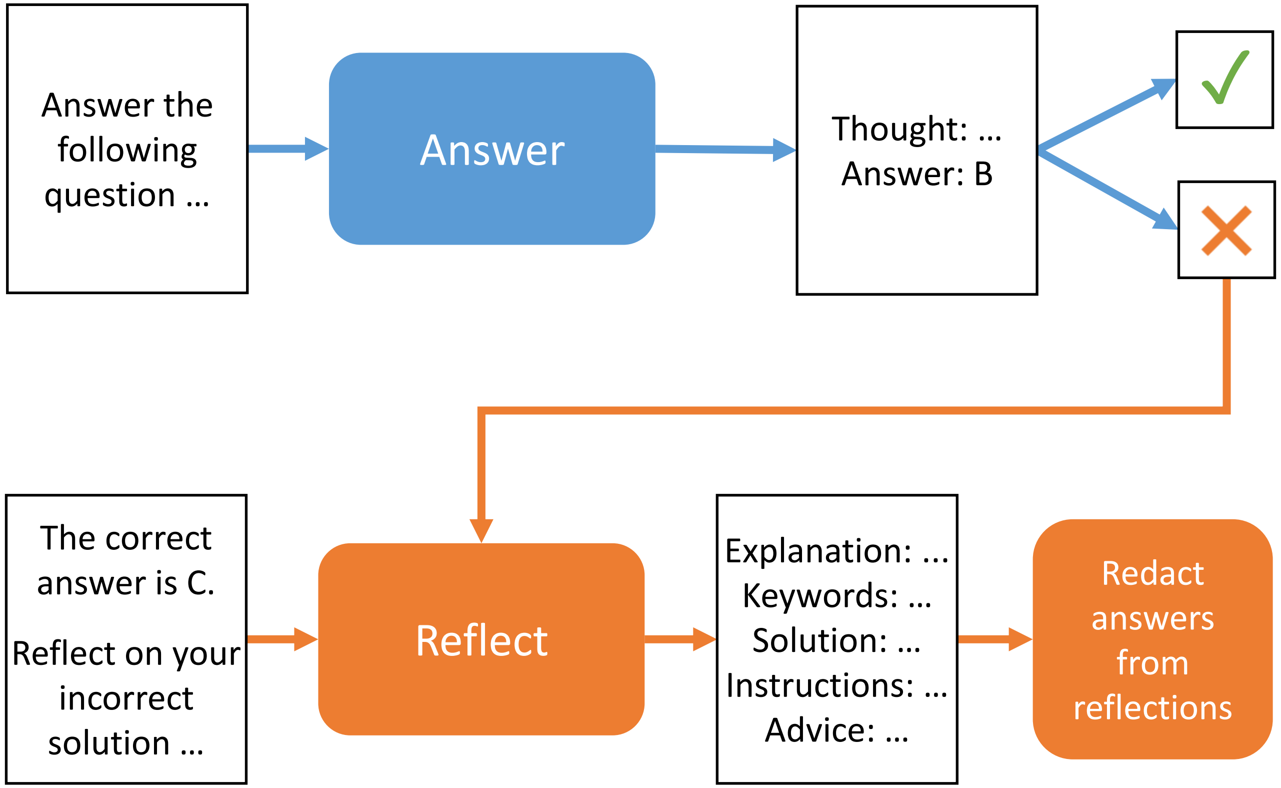
9 agents

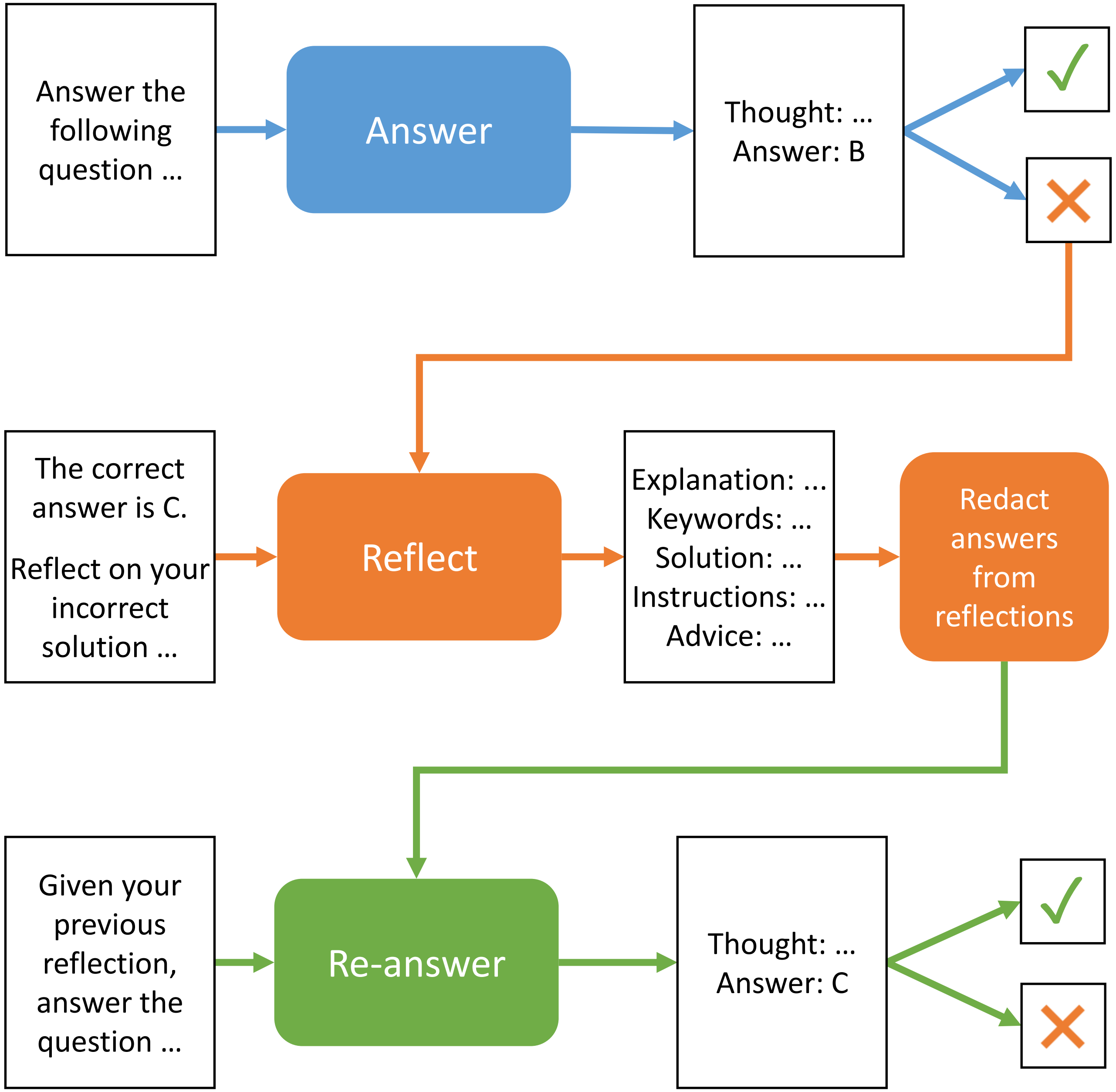
10 exams

1,000 questions




Answer the following question …
Answer
Thought: …
Answer: B


Answer the following question ...
Answer
Thought: ...
Answer: B
The correct answer is C.
Reflect on your incorrect solution ...
Reflect
Explanation: ...
Keywords: ...
Solution: ...
Instructions: ...
Advice: ...
Redact answers from reflections



Answer the following question ...
Answer
Thought: ...
Answer: B
The correct answer is C.
Reflect on your incorrect solution ...
Reflect
Explanation: ...
Keywords: ...
Solution: ...
Instructions: ...
Advice: ...
Redact answers from reflections
Given your previous reflection, answer the question ...
Re-answer
Thought: ...
Answer: C

Models

Name	Vendor	Released	License	Source
Claude 3 Opus	Anthropic	2024-03-04	Closed	Anthropic (2024)
Command R+	Cohere	2024-04-04	Open	Cohere (2024)
Gemini 1.0 Pro	Google	2023-12-06	Closed	Gemini Team (2023)
Gemini 1.5 Pro (Preview)	Google	2024-02-15	Closed	Gemini Team (2024)
GPT-3.5 Turbo	OpenAI	2022-11-30	Closed	OpenAI (2022)
GPT-4	OpenAI	2023-03-14	Closed	OpenAI (2023)
Llama 2 7B Chat	Meta	2023-07-18	Open	Meta (2023)
Llama 2 70B Chat	Meta	2023-07-18	Open	Meta (2023)
Mistral Large	Mistral AI	2024-02-26	Open	Mistral AI (2024)

Agents

Baseline – no self-reflection

Retry – informed and retries

Keywords – list of error keywords

Advice – advice for improvement

Explanation – what caused error

Instructions – steps for solving

Solution – step-by-step solution

Composite – all six reflections

Unredacted – no redactions

Exams

Problem Set	Benchmark	Domain	Questions	License	Source
ARC Challenge Test	ARC	Science	1,173	CC BY-SA	Clark (2018)
AQUA-RAT	AGI Eval	Math	254	Apache v2.0	Zhong (2023)
Hellaswag Val	Hellaswag	Common Sense Reasoning	10,042	MIT	Zellers (2019)
LogiQA (English)	AGI Eval	Logic	651	GitHub	Liu (2020)
LSAT-AR	AGI Eval	Law (Analytic Reasoning)	230	MIT	Wang (2021)
LSAT-LR	AGI Eval	Law (Logical Reasoning)	510	MIT	Wang (2021)
LSAT-RC	AGI Eval	Law (Reading Comprehension)	260	MIT	Wang (2021)
MedMCQA Valid	MedMCQA	Medicine	6,150	MIT	Pal (2022)
SAT-English	AGI Eval	English	206	MIT	Zhong (2023)
SAT-Math	AGI Eval	Math	220	MIT	Zhong (2023)

[System Prompt]

You are an expert in {{topic}}.

Your task is to answer the following multiple-choice questions.

Think step-by-step to ensure you have the correct answer.

Then, answer the question using the following format ‘Action: Answer("[choice]")’

The parameter [choice] is the letter or number of the answer you want to select (e.g. "A", "B", "C", or "D")

For example, ‘Answer("C")’ will select the choice "C" as the best answer.

You MUST select one of the available choices;
the answer CANNOT be "None of the Above".

Be concise in your response but include any essential information.

…

[Example Problem]

Topic: Geography

Question: What is the capital of the state where Johns Hopkins University is located?

Choices:
  A: Baltimore
  B: Annapolis
  C: Des Moines
  D: Las Vegas

[Example Solution]

Thought:

  Johns Hopkins University is located in Baltimore, Maryland.

  The capital of Maryland is Annapolis.

Action: Answer("B")

Topic: Science

Question: Red-clay bricks have a density of approximately 2000 kg/m^3. Air has a density of 1 kg/m^3. Which of the following has the lowest mass?

Choices:

A: 2 m^3 of bricks

B: 4 m^3 of bricks

C: 6000 m^3 of air

D: 10,000 m^3 of air

[System Prompt]

You are an expert in {{topic}}.

You have incorrectly answered the following multiple-choice question.

Your task is to reflect on the problem, your solution, and the correct answer.

You will then use this information help you answer the same question in the future.

First, explain why you answered the question incorrectly.

Second, list the keywords that describe the type of your errors
from most general to most specific.

Third, solve the problem again, step-by-step, based on your knowledge
of the correct answer.

Fourth, create a list of detailed instructions to help you correctly
solve this problem in the future.

Finally, create a list of general advice to help you solve similar
types of problems in the future.

Explanation:

I incorrectly calculated the mass of each option.

I made a calculation error in option [REDACTED].

Explanation:

I incorrectly calculated the mass of each option.

I made a calculation error in option [REDACTED].

Error Keywords:

- Calculation error
- Multiplication error
- Unit error

Solution:

To find the lowest mass, we need to calculate the mass of each option.

Mass = Density x Volume

Option [REDACTED]: Mass = 2000 kg/m^3 x 2 m^3 = 4000 kg

Option [REDACTED]: Mass = 2000 kg/m^3 x 4 m^3 = 8000 kg

Option [REDACTED]: Mass = 1 kg/m^3 x 6000 m^3 = 6000 kg

Option [REDACTED]: Mass = 1 kg/m^3 x 10000 m^3 = 10000 kg

The option with the lowest mass is option [REDACTED], with a mass of 4000 kg.

Instructions:

1. Identify the density of each material.
2. Identify the volume of each option.
3. Calculate the mass of each option using the formula Mass = Density x Volume.
4. Double-check your calculations to ensure accuracy.
5. Choose the option with the lowest mass.

Advice:

- Always read the question carefully and understand the problem.
- Always decompose complex problems into multiple simple steps.
- Always think through each subproblem step-by-step.
- Never skip any steps; be explicit in each step of your reasoning.
- Always double-check your calculations and final answer.
- Remember to use the correct units in your calculations.

Results

Questions: 100

Correct: 86

Incorrect: 14

Accuracy: 0.8600

Errors: 0

Input Tokens: 30072

Output Tokens: 4925

Total Tokens: 34997

…

Analysis

McNemar test

$$\chi^2 = \frac{(b-c)^2}{b+c}$$

where *b* and *c* are discordant pairs in

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Analysis

McNemar test

Binary outcomes

$$\chi^2 = \frac{(b-c)^2}{b+c}$$

where *b* and *c* are discordant pairs in

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Analysis

McNemar test

Binary outcomes

Paired by question

$$\chi^2 = \frac{(b-c)^2}{b+c}$$

where *b* and *c* are discordant pairs in

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Analysis

McNemar test

Binary outcomes

Paired by question

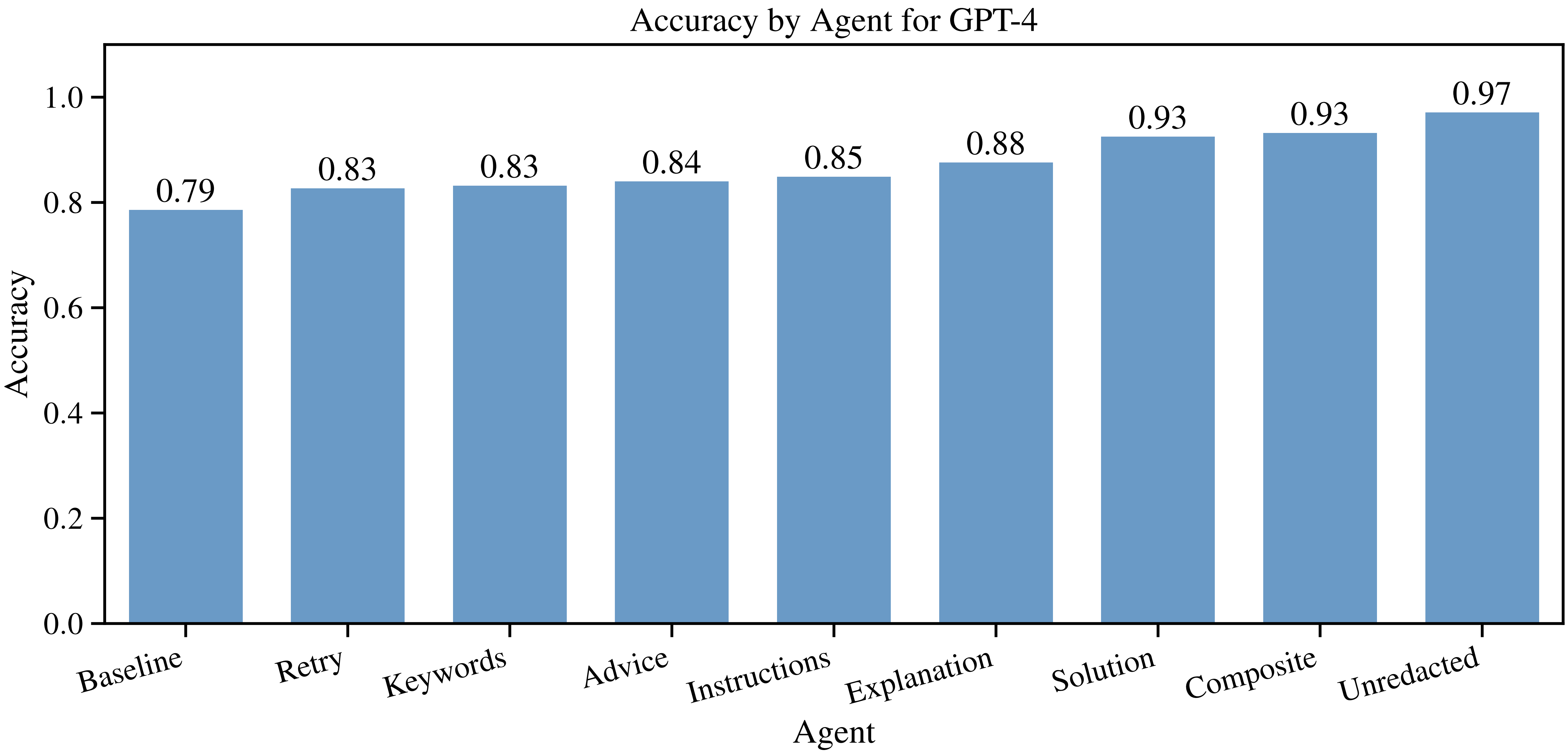
Compares discordant pairs

$$\chi^2 = \frac{(b-c)^2}{b+c}$$

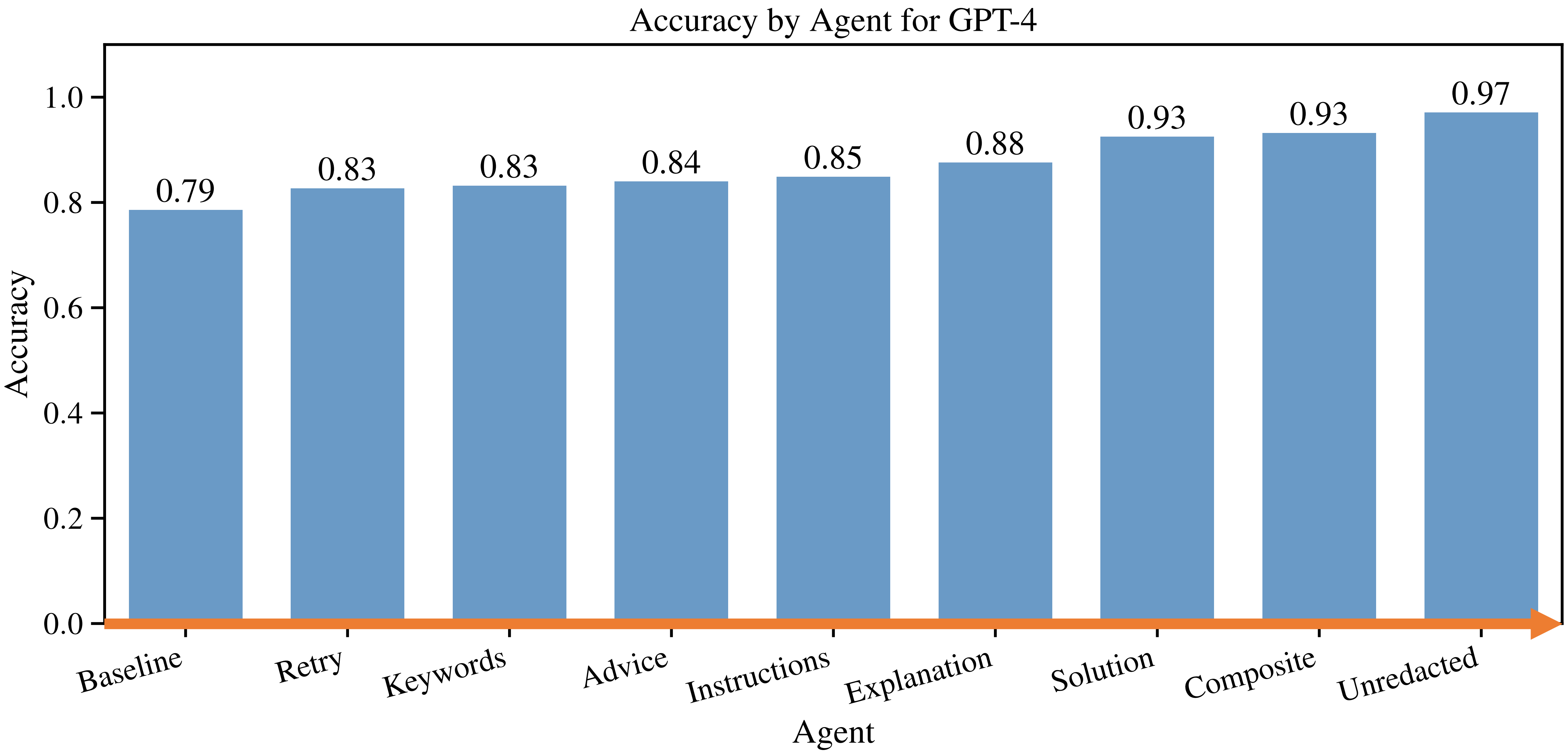
where *b* and *c* are discordant pairs in

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

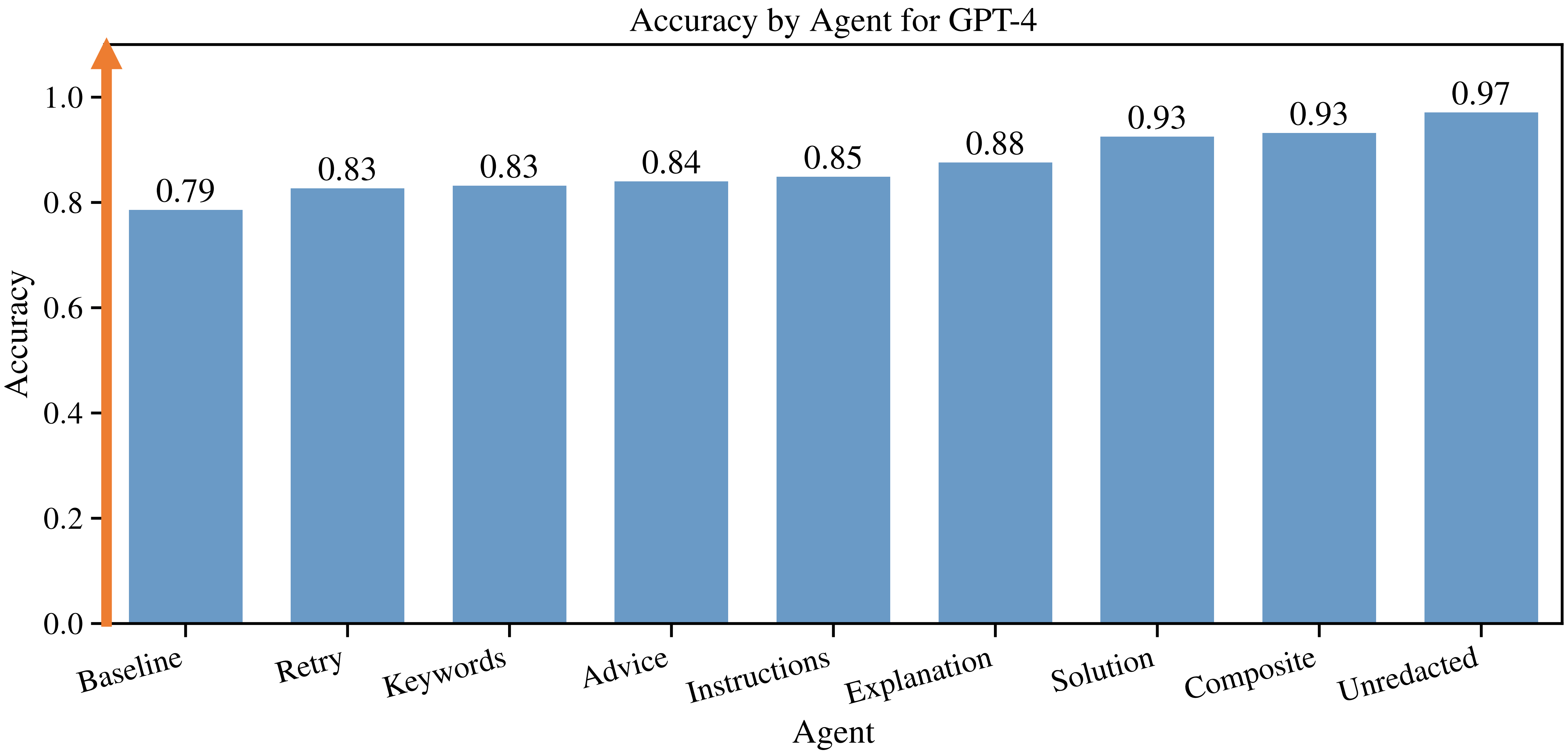
Results

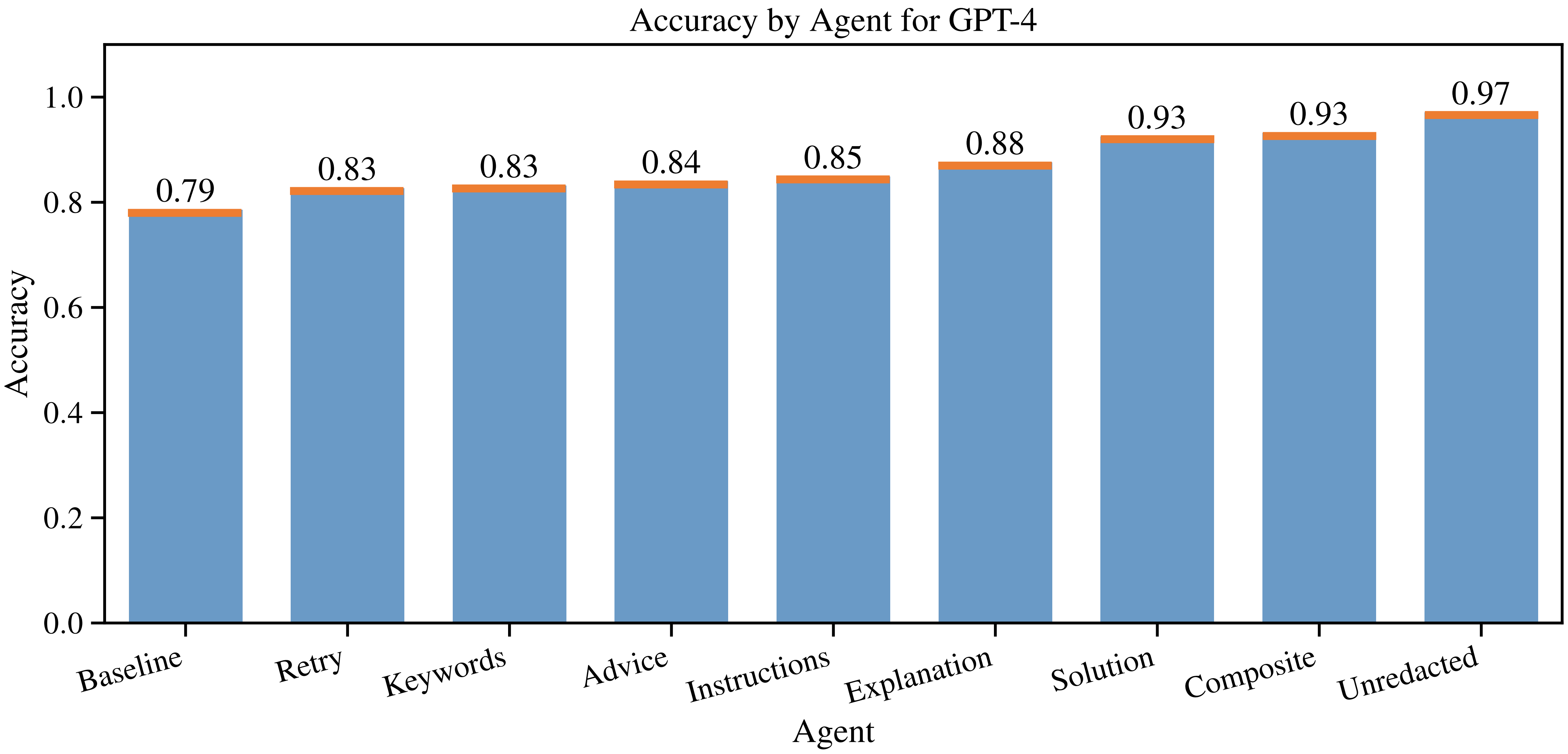


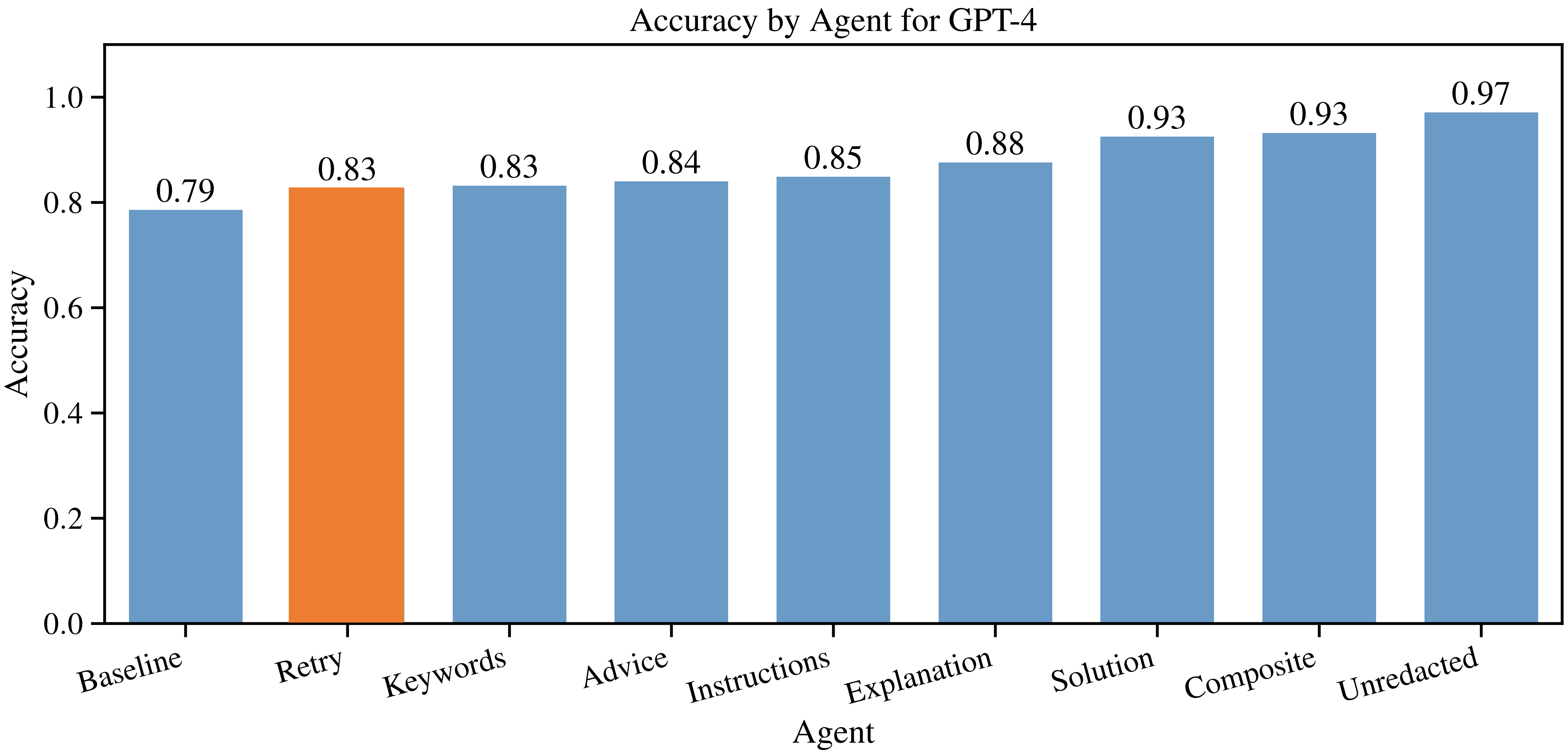
Accuracy by Agent for GPT-4
Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
0.79
0.83
0.83
0.84
0.85
0.88
0.93
0.93
0.97
Baseline
Retry
Keywords
Advice
Instructions
Explanation
Solution
Composite
Unredacted
Agent

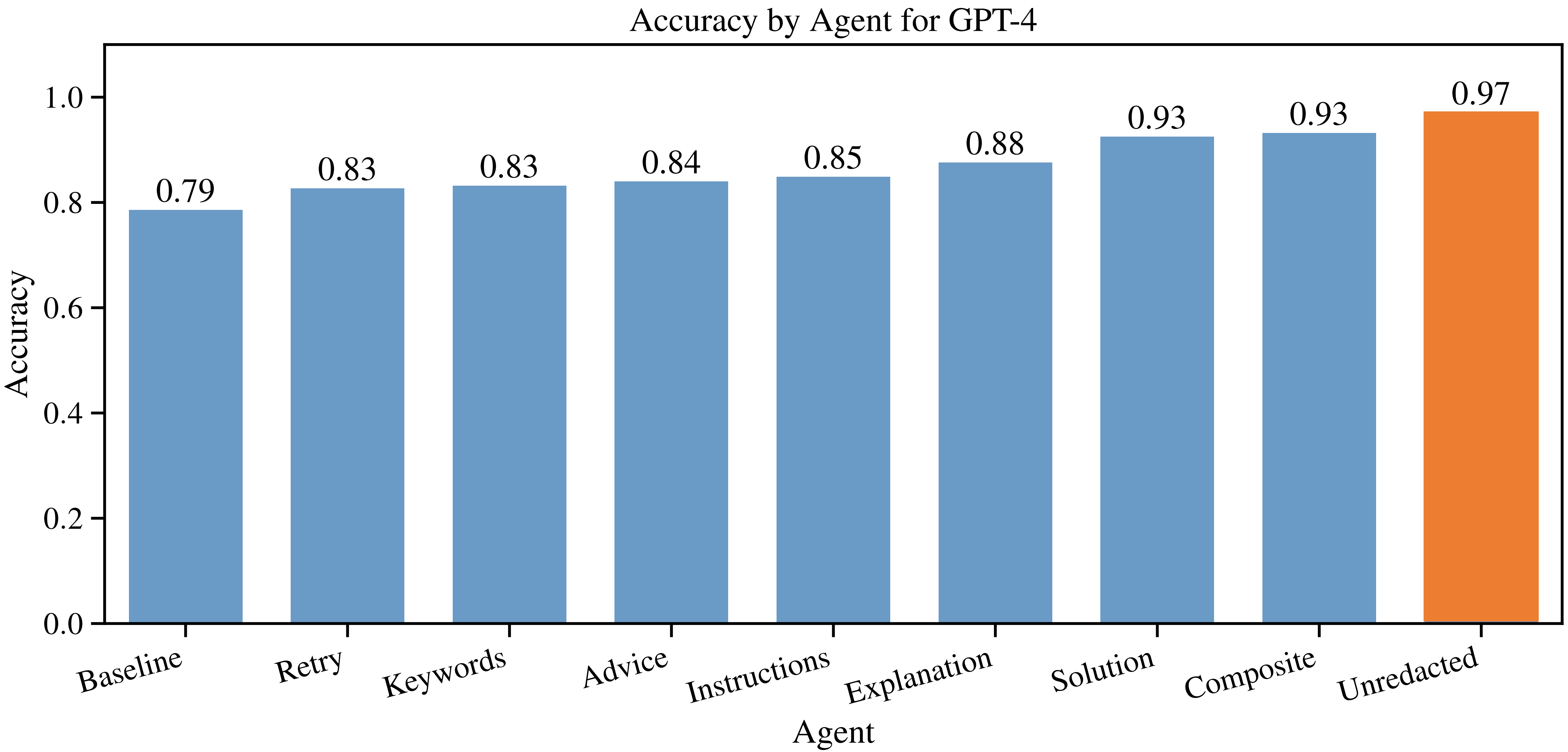


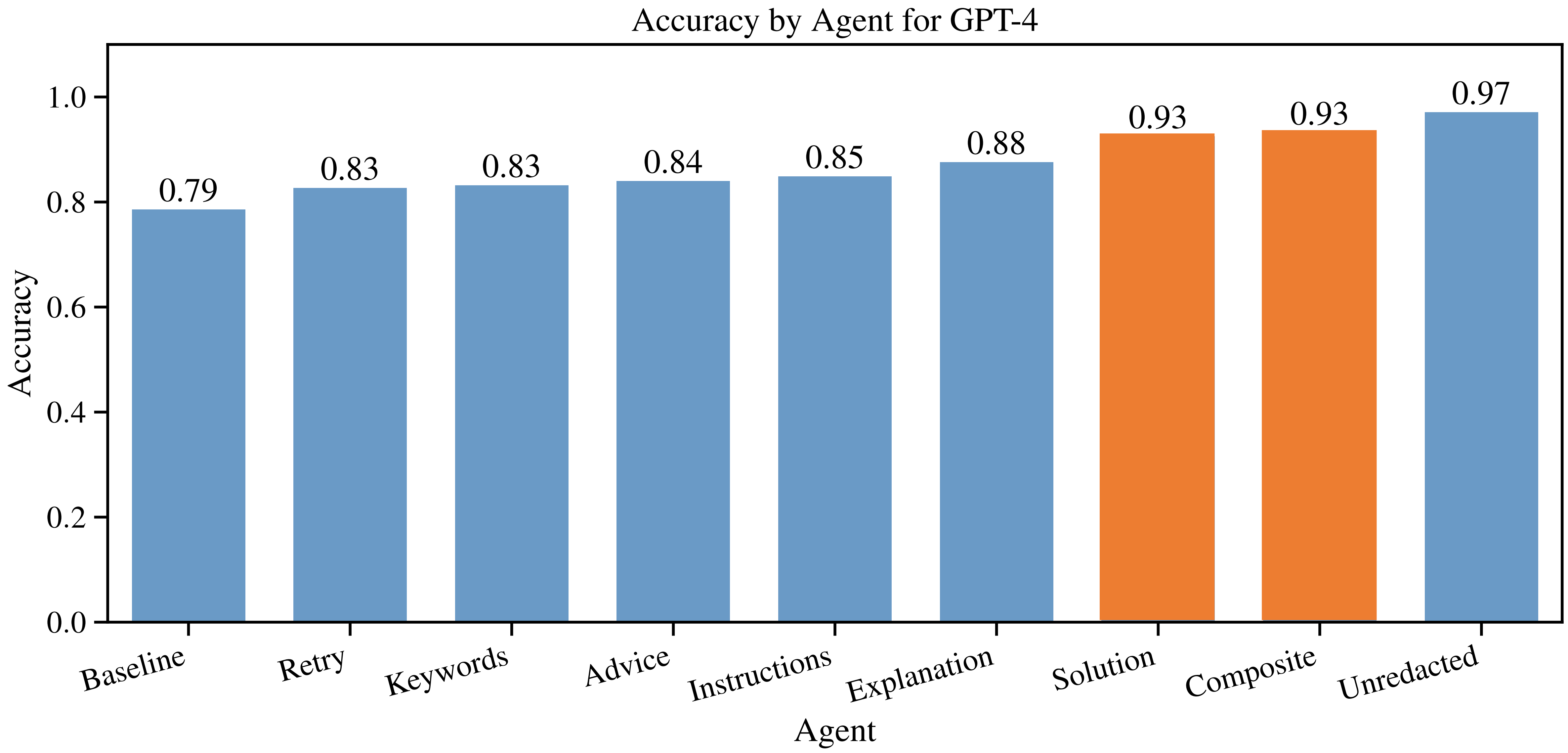
Accuracy by Agent for GPT-4
Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
0.79
0.83
0.83
0.84
0.85
0.88
0.93
0.93
0.97
Baseline
Retry
Keywords
Advice
Instructions
Explanation
Solution
Composite
Unredacted
Agent


Accuracy by Agent for GPT-4
Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
0.79
0.83
0.83
0.84
0.85
0.88
0.93
0.93
0.97
Baseline
Retry
Keywords
Advice
Instructions
Explanation
Solution
Composite
Unredacted
Agent

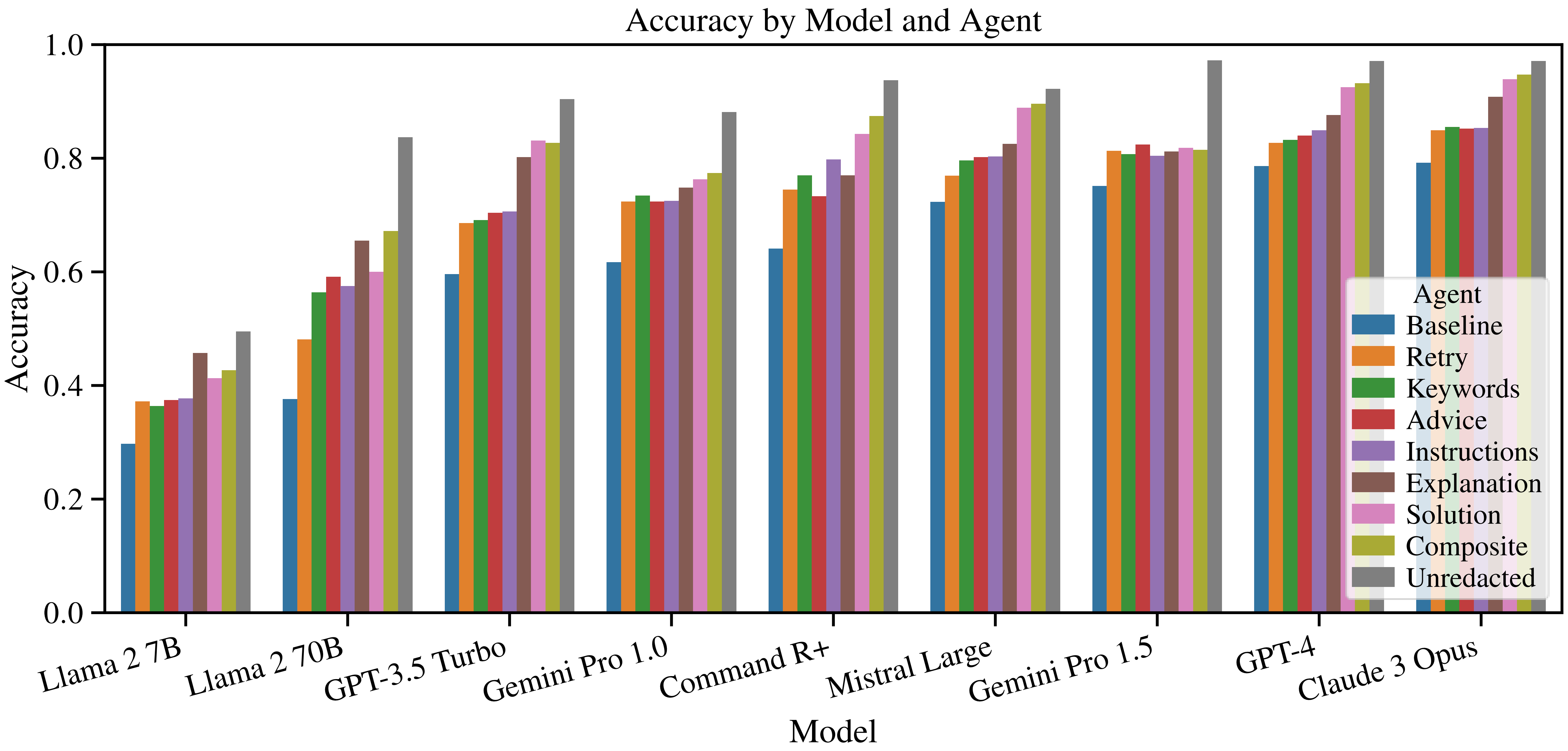

Accuracy by Agent for GPT-4
Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
0.79
0.83
0.83
0.84
0.85
0.88
0.93
0.93
0.97
Baseline
Retry
Keywords
Advice
Instructions
Explanation
Solution
Composite
Unredacted
Agent


Accuracy by Agent for GPT-4
Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
0.79
0.83
0.83
0.84
0.85
0.88
0.93
0.93
0.97
Baseline
Retry
Keywords
Advice
Instructions
Explanation
Solution
Composite
Unredacted
Agent

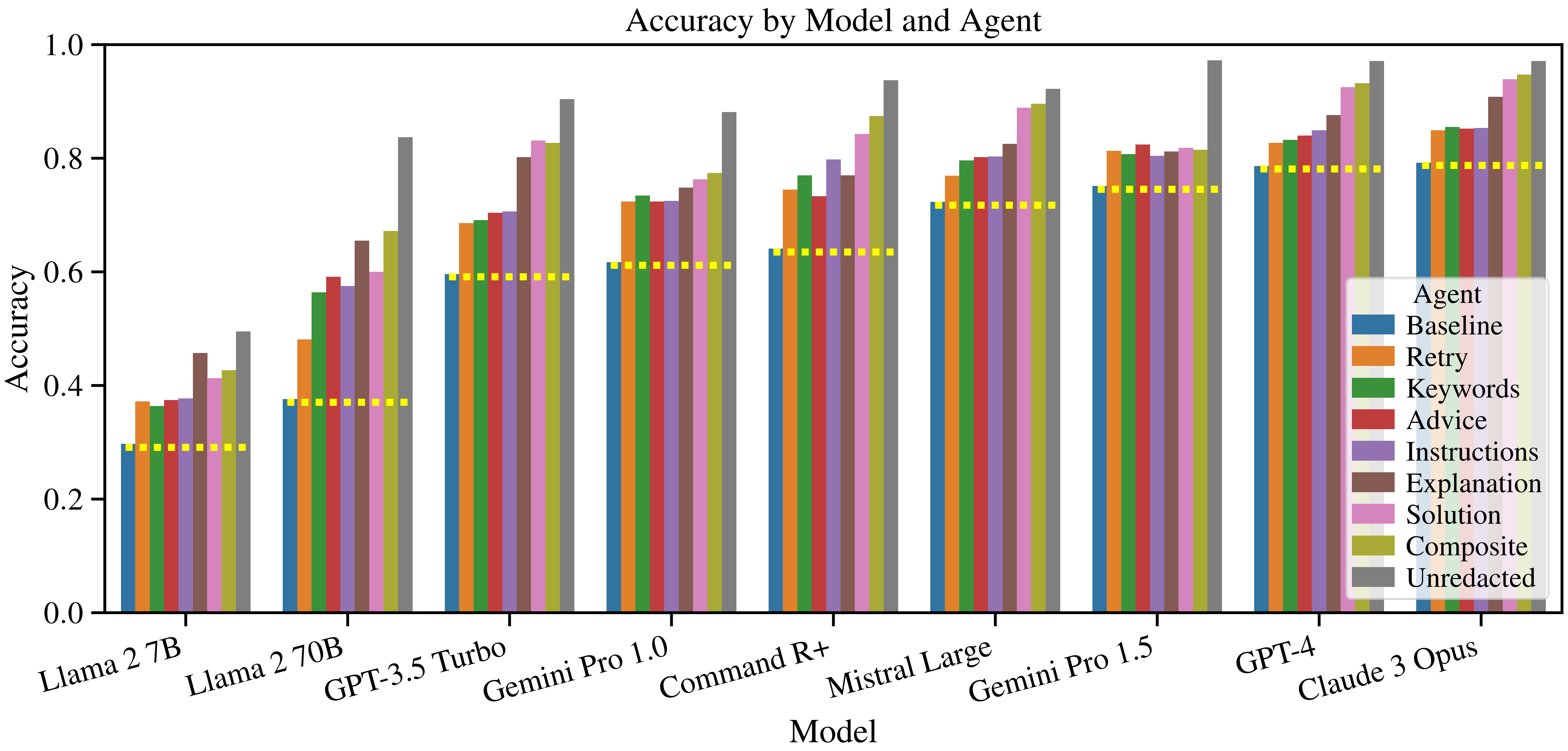

Accuracy by Agent for GPT-4
Accuracy
Agent
1.0
0.8
0.6
0.4
0.2
0.0
0.79
0.83
0.83
0.84
0.85
0.88
0.93
0.93
0.97
Baseline
Retry
Keywords
Advice
Instructions
Explanation
Solution
Composite
Unredacted



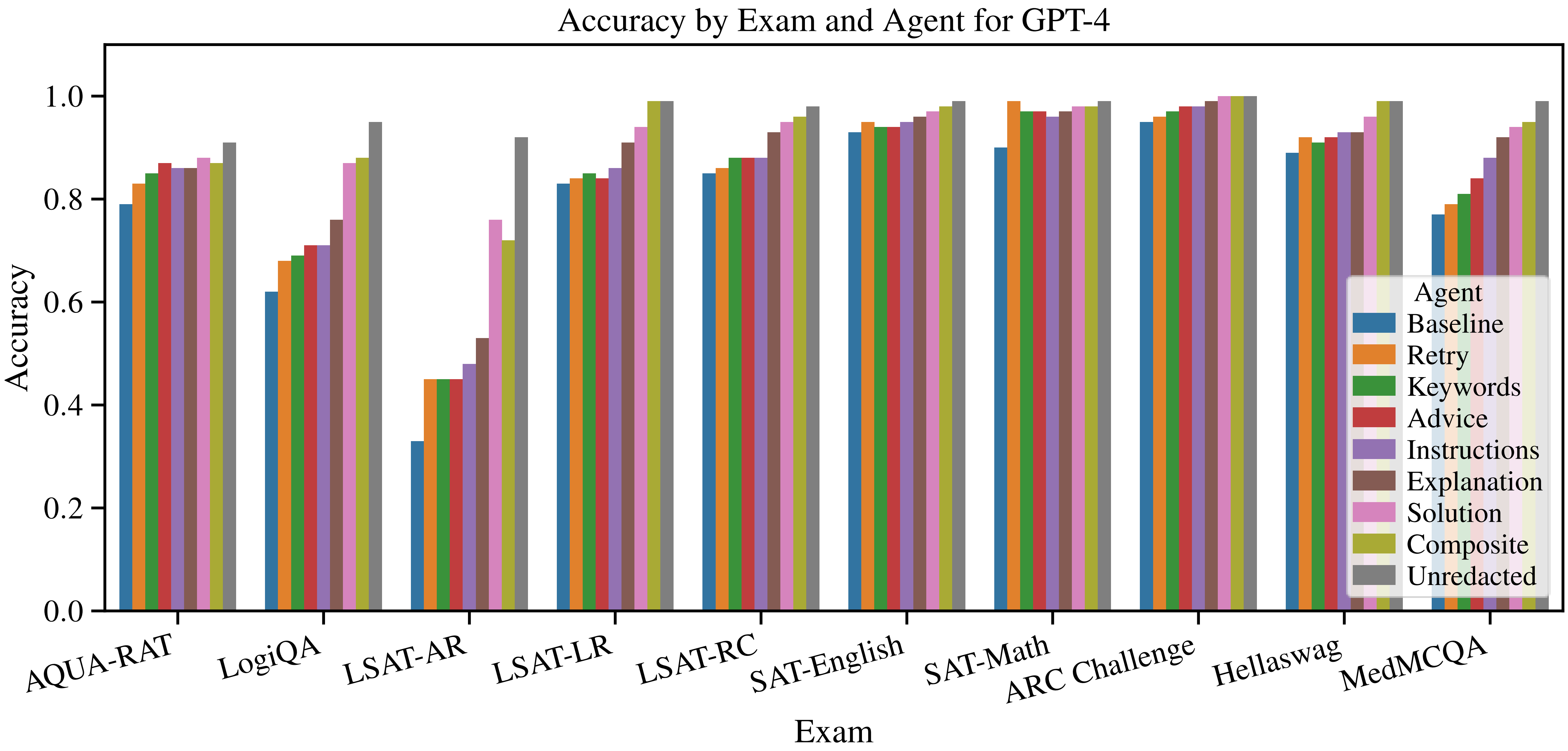
Accuracy by Agent for GPT-4
Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
0.79
0.83
0.83
0.84
0.85
0.88
0.93
0.93
0.97
Baseline
Retry
Keywords
Advice
Instructions
Explanation
Solution
Composite
Unredacted
Agent



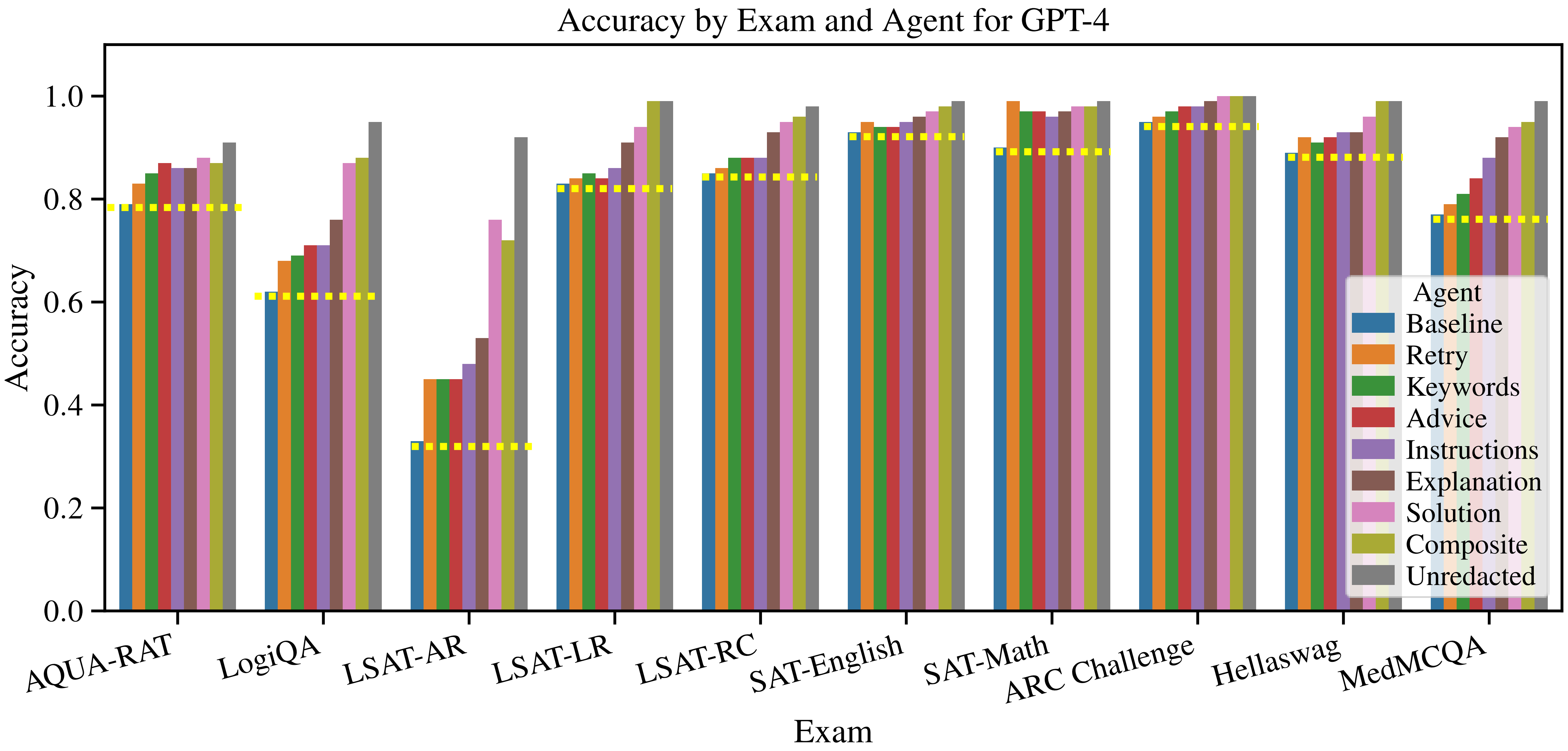
Accuracy by Model and Agent
Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
Llama 2 7B
Llama 2 70B
GPT-3.5 Turbo
Gemini Pro 1.0
Command R+
Mistral Large
Gemini Pro 1.5
GPT-4
Claude 3 Opus
Model
Agent
Baseline
Retry
Keywords
Advice
Instructions
Explanation
Solution
Composite
Unredacted



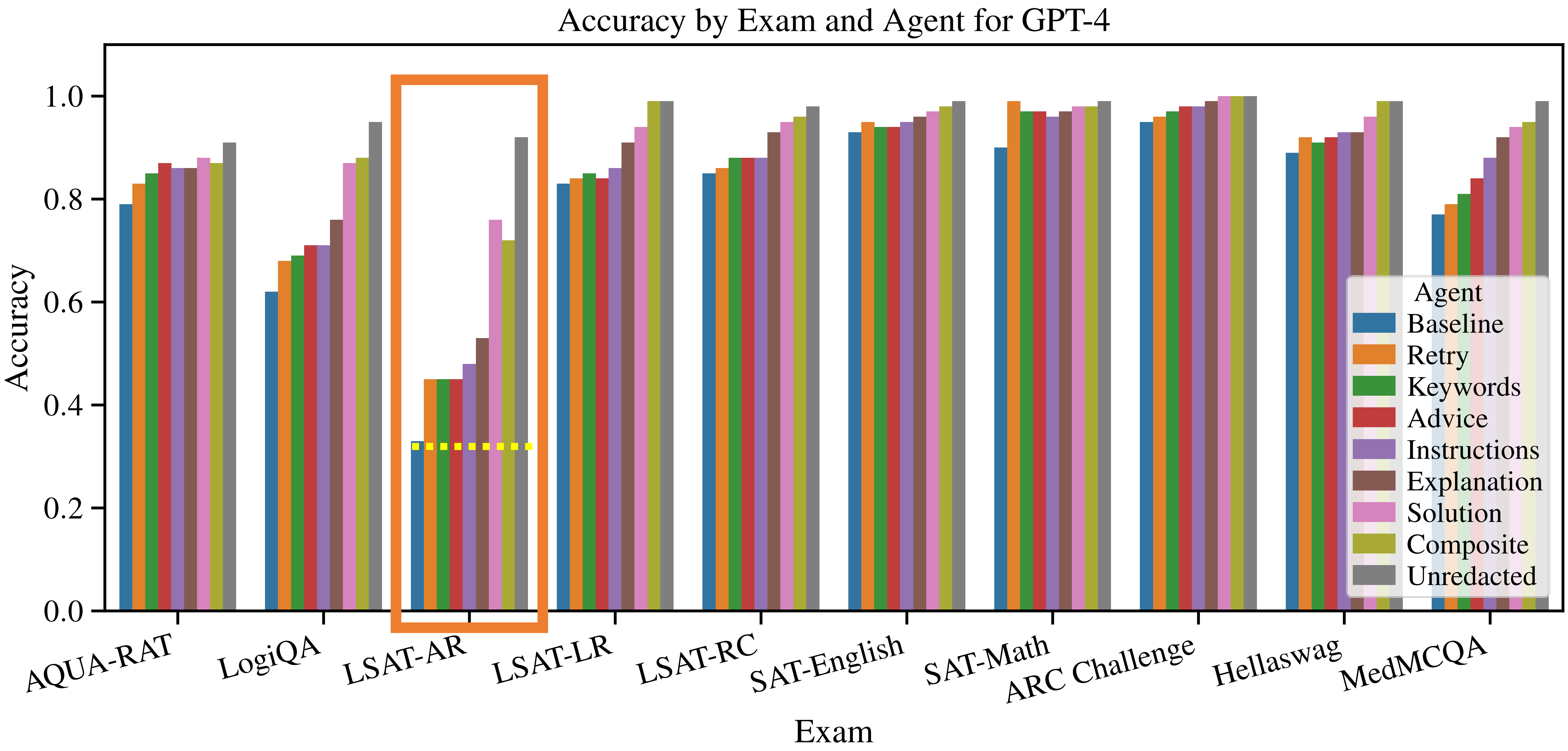
Accuracy by Model and Agent
Accuracy
0.0
0.2
0.4
0.6
0.8
1.0
Llama 2 7B
Llama 2 70B
GPT-3.5 Turbo
Gemini Pro 1.0
Command R+
Mistral Large
Gemini Pro 1.5
GPT-4
Claude 3 Opus
Model
Agent
Baseline
Retry
Keywords
Advice
Instructions
Explanation
Solution
Composite
Unredacted



Accuracy by Exam and Agent for GPT-4
Accuracy
0.0
0.2
0.4
0.6
0.8
1.0
Exam
AQUA-RAT
LogiQA
LSAT-AR
LSAT-LR
LSAT-RC
SAT-English
SAT-Math
ARC Challenge
Hellaswag
MedMCQA
Agent
Baseline
Retry
Keywords
Advice
Instructions
Explanation
Solution
Composite
Unredacted



Accuracy by Exam and Agent for GPT-4
Accuracy
0.0
0.2
0.4
0.6
0.8
1.0
Exam
AQUA-RAT
LogiQA
LSAT-AR
LSAT-LR
LSAT-RC
SAT-English
SAT-Math
ARC Challenge
Hellaswag
MedMCQA
Agent
Baseline
Retry
Keywords
Advice
Instructions
Explanation
Solution
Composite
Unredacted



Accuracy by Exam and Agent for GPT-4
Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
AQUA-RAT
LogiQA
LSAT-AR
LSAT-LR
LSAT-RC
SAT-English
SAT-Math
ARC Challenge
Hellaswag
MedMCQA
Exam
Agent
Baseline
Retry
Keywords
Advice
Instructions
Explanation
Solution
Composite
Unredacted

Discussion

Interpretation

All types of self-reflection improve problem-solving performance

Interpretation

All types of self-reflection improve problem-solving performance

More reflection information leads to better performance

Interpretation

All types of self-reflection improve problem-solving performance

More reflection information leads to better performance

Simply re-answering also improves performance

Interpretation

All types of self-reflection improve problem-solving performance

More reflection information leads to better performance

Simply re-answering also improves performance???

Limitations

Limitations

Only single-step tasks

Limitations

Only single-step tasks

API response errors

Limitations

Only single-step tasks

API response errors

Saturation effect

Limitations

Only single-step tasks

API response errors

Saturation effect

LSAT-AR skew

Future Research

More complex problems

Future Research

More complex problems

Provide agents tools

Future Research

More complex problems

Provide agents tools

Provide memory

Future Research

More complex problems

Provide agents tools

Provide memory

More evidence

Conclusion

Conclusion

Conclusion

All types of self-reflection improve problem-solving performance

Conclusion

All types of self-reflection improve problem-solving performance

More reflection information leads to better performance

Conclusion

All types of self-reflection improve problem-solving performance

More reflection information leads to better performance

Simply re-answering also improves performance

Learn more

https://matthewrenze.com/research/self-reflection-in-llm-agents/