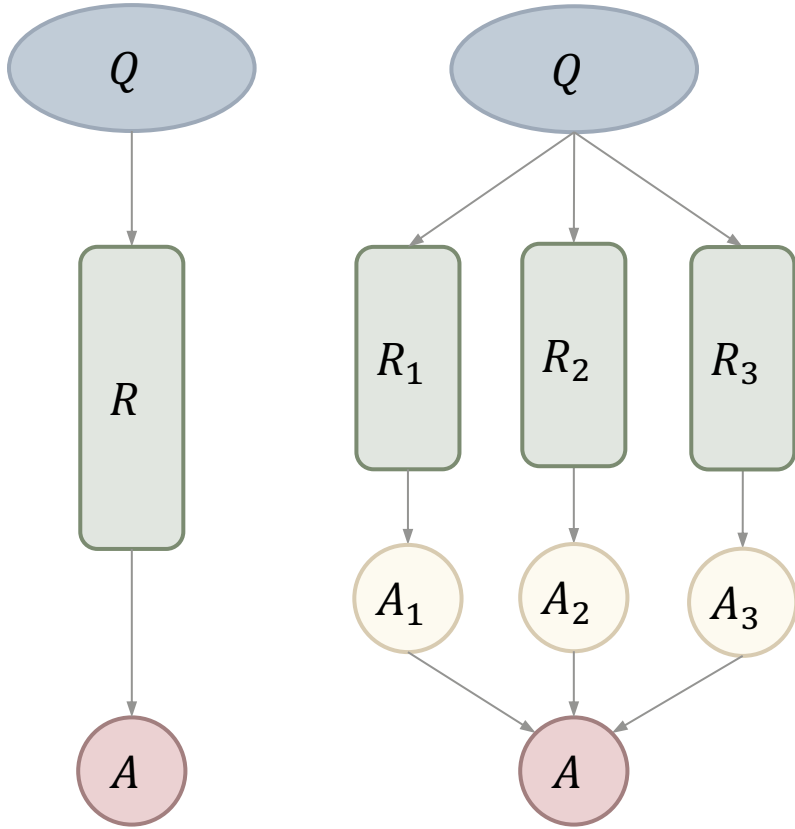


Aggregation of Reasoning: A Hierarchical Framework for Enhancing Answer Selection in Large Language Models

Zhangyue Yin, Qiushi Sun, Qipeng Guo, Zhiyuan Zeng, Xiaonan Li, Tianxiang Sun
Cheng Chang, Qinyuan Cheng, Ding Wang, Xiaofeng Mou, Xipeng Qiu, Xuanjing Huang



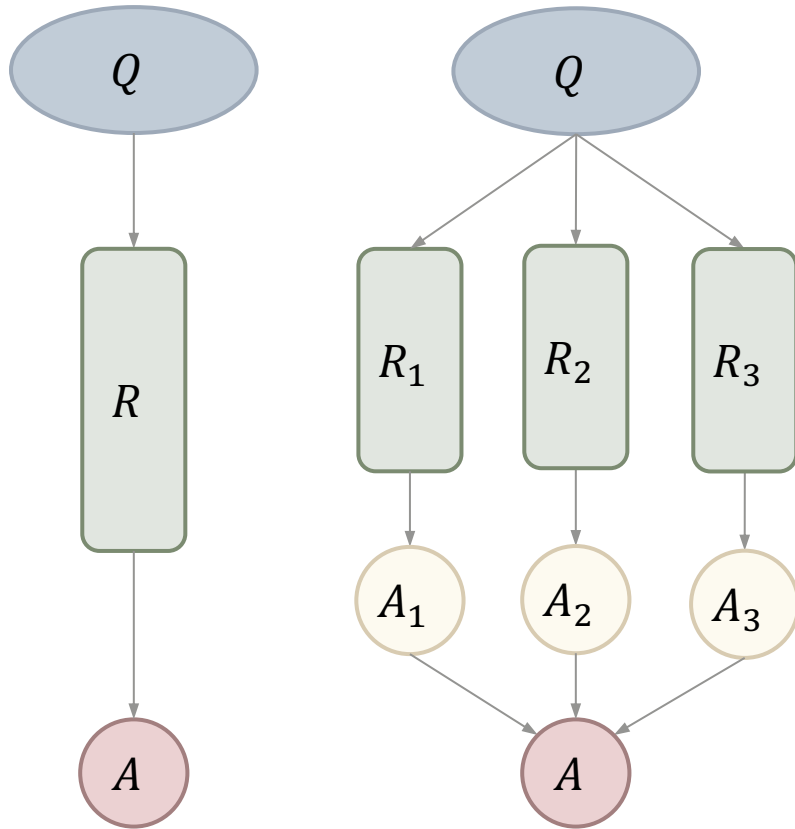
Background: Chain-of-Thought and Self-Consistency



Chain-of-Thought

Self-Consistency

Background: Weakness in Majority Vote



Chain-of-Thought

Self-Consistency

Q : The difference between simple interest and C.I. at the same rate for Rs.5000 for 2 years in Rs.72. The rate of interest is?
(A) 10% (B) 12% (C) 6% (D) 8% (E) 4%

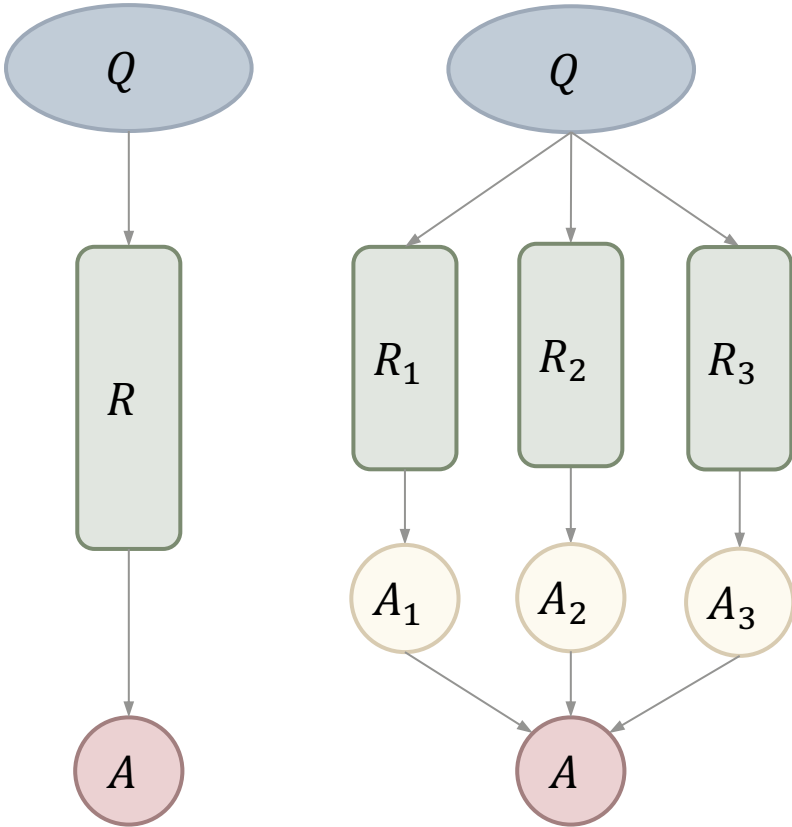
$R1$: Let the rate of interest be r . The difference between simple interest and compound interest at the same rate for 2 years..., the answer is 6%. So the answer is (C). ✗

$R2$: Let the rate of interest be r . $..|(5000 * r / 100) * 2 - 5000 * ((1 + r/100)^2 - 1)| = 72$. Simplifying this equation gives $r = 12$. So the answer is (B). ✓

$R3$: Let the rate of interest be $x\%$. The simple interest for 2 years would be $(5000 * 2 * x) / 100 = 100x...$ So, $5000 * (1 + x/100)^2 - 5000 - 100x = 72$. Solving this equation, we get $x = 6\%$. So the answer is (C). ✗

Erroneous cases of Majority Vote

Background: Weakness in Majority Vote



Chain-of-Thought

Self-Consistency

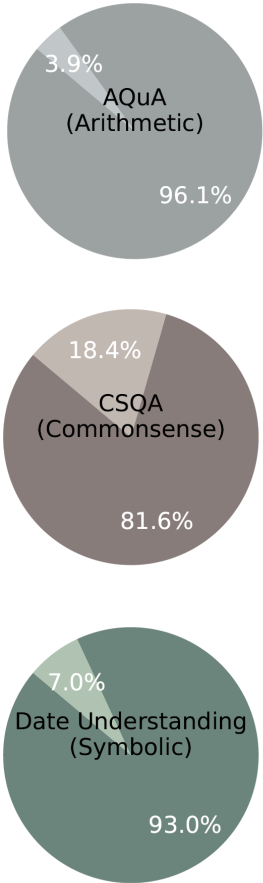
Q: The difference between simple interest and C.I. at the same rate for Rs.5000 for 2 years in Rs.72. The rate of interest is?
(A) 10% (B) 12% (C) 6% (D) 8% (E) 4%

R1: Let the rate of interest be r . The difference between simple interest and compound interest at the same rate for 2 years..., the answer is 6%. So the answer is (C). ✗

R2: Let the rate of interest be r . $..|(5000 * r / 100) * 2 - 5000 * ((1 + r/100)^2 - 1)| = 72$. Simplifying this equation gives $r = 12$. So the answer is (B). ✓

R3: Let the rate of interest be $x\%$. The simple interest for 2 years would be $(5000 * 2 * x) / 100 = 100x...$ So, $5000 * (1 + x/100)^2 - 5000 - 100x = 72$. Solving this equation, we get $x = 6\%$. So the answer is (C). ✗

Erroneous cases of Majority Vote



Proportion

Methodology: Aggregation of Reasoning

Local-Scoring

- Reasoning chains with the same answer.
- Filter out low-quality reasoning chains

Global-Evaluation

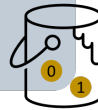
- Reasoning chains with different answers.
- Select the answer corresponding to the reasoning chain with the highest average score.

Question: *The difference between simple interest and C.I. at the same rate for Rs.5000 for 2 years in Rs.72. The rate of interest is?*
 Answer Choices: (A) 10% (B) 12% (C) 6% (D) 8% (E) 4%

R0: Let's first find the simple interest. Simple interest = $(P * R * T) / 100 = (5000 * R * 2) / 100 = 100R$... Since the rate of interest cannot be negative, the rate of interest is 10%. So the answer is (A).

R1: Let the rate of interest be $r\%$. The compound interest would be $5000[(1 + r/100)^2 - 1]$... This quadratic equation gives us $r = 10\%$. So the answer is (A).

Answer: (A)



R2: Let the rate of interest be r . Then, according to the given information, $|(5000 * r / 100) * 2 - 5000 * ((1 + r/100)^2 - 1)| = 72$... Simplifying this equation gives $r = 12$. So the answer is (B).

R6: Let the rate of interest be r . Then, according to the given information, we have C.I. - S.I. = $(P * r/100)^2 / 2$... So the rate of interest is 12%. So the answer is (B).

Answer: (B)

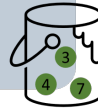


R3: The difference between simple interest and compound interest at the same rate for 2 years..., answer is 6%. So the answer is (C).

R4: The simple interest for 2 years would be $(5000 * 2 * x) / 100 = 100x$..., we get $x = 6\%$. So the answer is (C).

R7: Using the formula, $CI = P(1 + r/100)^n - P$... So the answer is (C).

Answer: (C)



R5: According to the given information, C.I. - Simple Interest = 72... $r = 8\%$. So the answer is (D).

R8: The interest earned after 2 years would be $5000 * r / 100 = 50r$, ... $5r/2 - 5000 = 72$... $r = 8\%$. So the answer is (D).

R9: C.I. - Simple Interest = 72... we take the solution $r = 8\%$. So the answer is (D).

Answer: (D)



Local-Scoring

Prompt: Evaluate the solution process for the problem using the criteria below, with a maximum score of 10 points:

- Logical Consistency (3 points):** Evaluate the coherence and soundness of the reasoning.
- Appropriateness of Method (3 points):** Ensure the method used is suitable, and that the approach is not overcomplicated.
- Completeness and Clarity (2 points):** Ensure all necessary steps are shown clearly, without any omission.
- Application of Knowledge (2 points):** Assess the correct and relevant use of formulas, theorems, or fact.

R2: Let the rate of interest be r . Then, according to the given information, $|(5000 * r / 100) * 2 - 5000 * ((1 + r/100)^2 - 1)| = 72$... Simplifying this equation gives $r = 12$. So the answer is (B).

R6: Let the rate of interest be r . Then, according to the given information, we have C.I. - S.I. = $(P * r/100)^2 / 2$... So the rate of interest is 12%. Therefore, the answer is (B).

Response: Logical Consistency: Solution 1: The formula used is simpler and more direct.... Adding all the points, **R2 receives 10/10 and R6 receives 7/10**. Both responses have done well, but R2 provided a more clear and complete solution. So the score is (R2, 10), (R6, 7).

Global-Evaluation

Prompt: Multiple solution processes are presented below, each leading to a different answer. Only one of these answers is correct. Evaluate each solution process based on:

- Validity of Approach (3 points):** Does the method employed effectively address the problem?
- Consistency of Steps and Answer (3 points):** Are all steps not only correct but also consistent with the final answer?
- Completeness and Clarity (2 points):** Have all essential steps been delineated, and are they presented unambiguously?
- Application of Knowledge (2 points):** Is the use of formulas, theorems, or facts precise and appropriate?

R1: Let the rate of interest be $r\%$. The compound interest would be $5000[(1 + r/100)^2 - 1]$... So the answer is (A).

R2: Let the rate of interest be r . Then, according to the given information, $|(5000 * r / 100) * 2 - 5000 * ((1 + r/100)^2 - 1)| = 72$... Simplifying this equation gives $r = 12$. So the answer is (B).

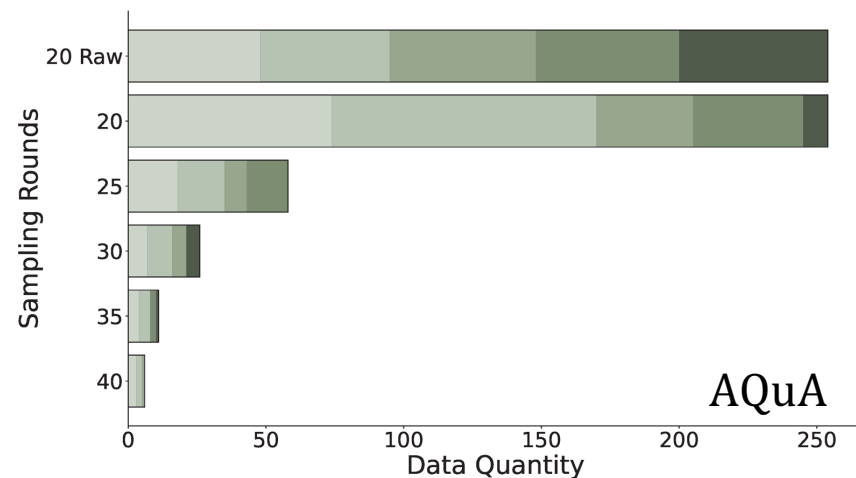
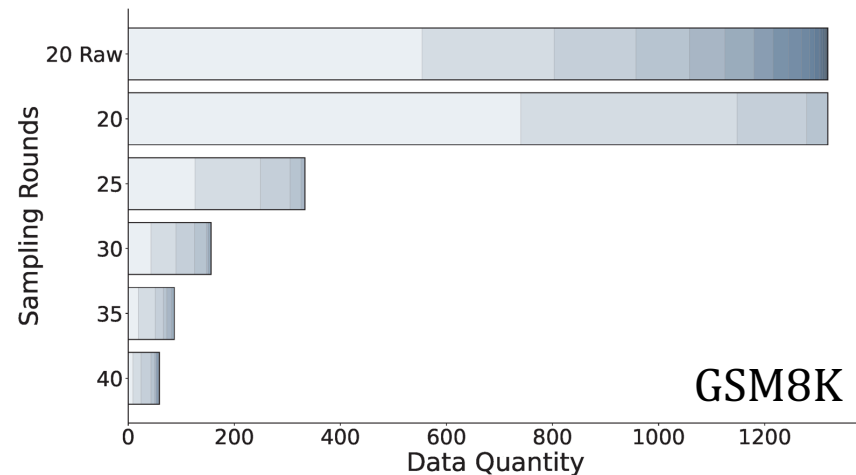
R3: The difference between simple interest and compound interest at the same rate for 2 years... So the answer is (C).

R8: The interest earned after 2 years would be $5000 * r / 100 = 50r$, ... $5r/2 - 5000 = 72$... $r = 8\%$. So the answer is (D).

Response: Let's evaluate each solution: R1: Validity of Approach (3 points): ... Adding all the points, **R1 receives 3/10, R2 gets 10/10, R3 gets 6/10, and R8 gets 5/10**. R2 appearing the most comprehensive and accurate, with R3 following closely behind. R1 and R8 both made crucial errors in their computations. So the score is (R1, 3), (R2, 10), (R3, 6), (R8, 5).

Methodology: Dynamic Sampling

- ❑ Dynamic sampling: Allocate different numbers of reasoning chain sampling attempts for questions of varying difficulty.
 - ❑ Easy questions: few reasoning chains
 - ❑ Hard questions: more reasoning chains
 - ❑ Stopping criterion: The model's confidence in the optimal reasoning chain.
- ❑ Analysis on GSM8K and AQuA
 - ❑ GSM8K: Open-ended questions with a broader distribution of answer quantities.
 - ❑ AQuA: Multiple-choice questions with the maximum number of answers limited to the number of options.
 - ❑ For most questions, only a small number of reasoning chains need to be sampled.

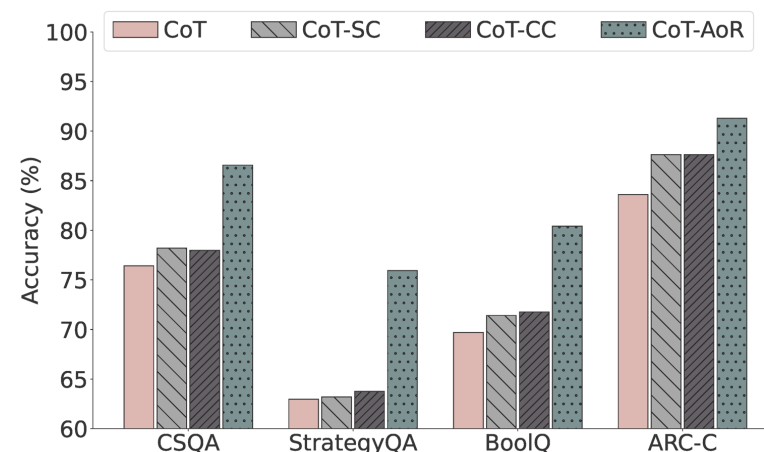


Experiment: Commonsense Reasoning & Symbolic Reasoning

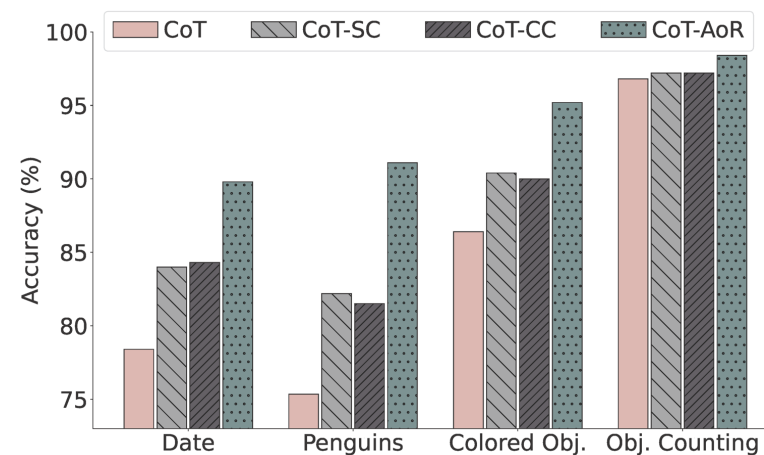
Feature	<i>AoR</i> (our work)	Self-Consistency (Wang et al., 2023b)	ComplexSC (Fu et al., 2023b)	PHP (Zheng et al., 2023)	DiVeRSE (Li et al., 2023b)
Task Agnostic?	✓	✓	✓	✗	✓
Training-Free?	✓	✓	✓	✓	✗
Plug-and-Play?	✓	✓	✓	✗	✗
Dynamic Sampling?	✓	✗	✗	✓	✗

Comparison of different reasoning chains ensemble methods

- ❑ A significant advantage over the previous strong baseline in both commonsense and symbolic reasoning tasks.
- ❑ Compared to previous reasoning chain integration methods, the AoR method has features such as task agnosticism, training-free, plug-and-play, and dynamic sampling.



Commonsense Reasoning

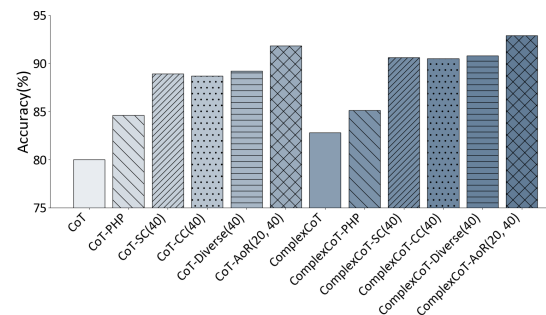


Symbolic Reasoning

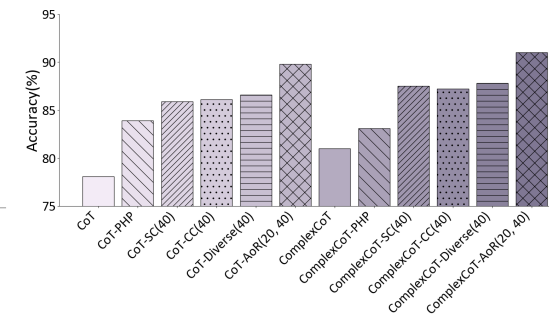
Experiment: Mathematical Reasoning

- ❑ The AoR method outperformed previous reasoning chain ensemble methods.
- ❑ Self-Consistency: Majority voting for generated answers.
- ❑ Complex-based Self-Consistency: Majority voting based on answers from reasoning chains of higher complexity.
- ❑ Diverse: Training a verifier to individually evaluate each reasoning chain.
- ❑ PHP: Using previously generated answers to guide the current reasoning process.

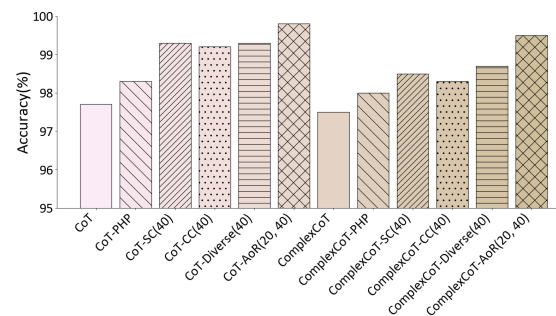
GSM8K



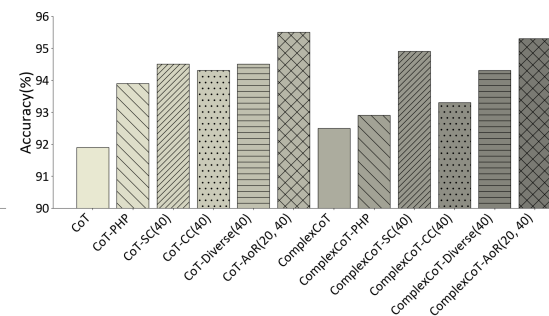
SVAMP



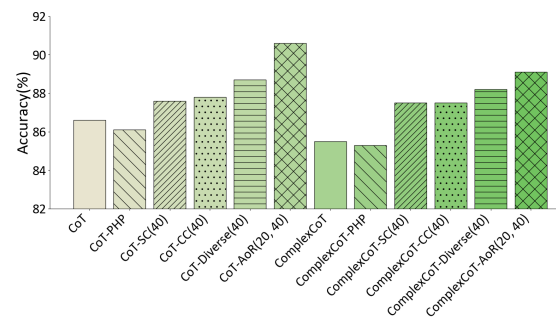
MultiArith



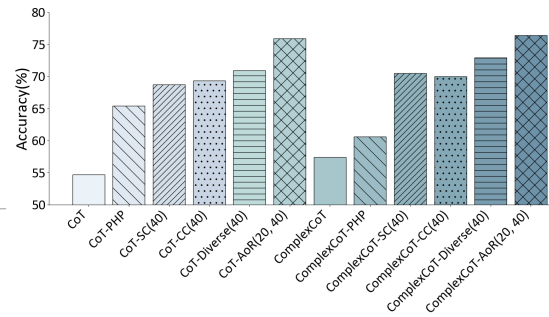
SingleEQ



AddSub

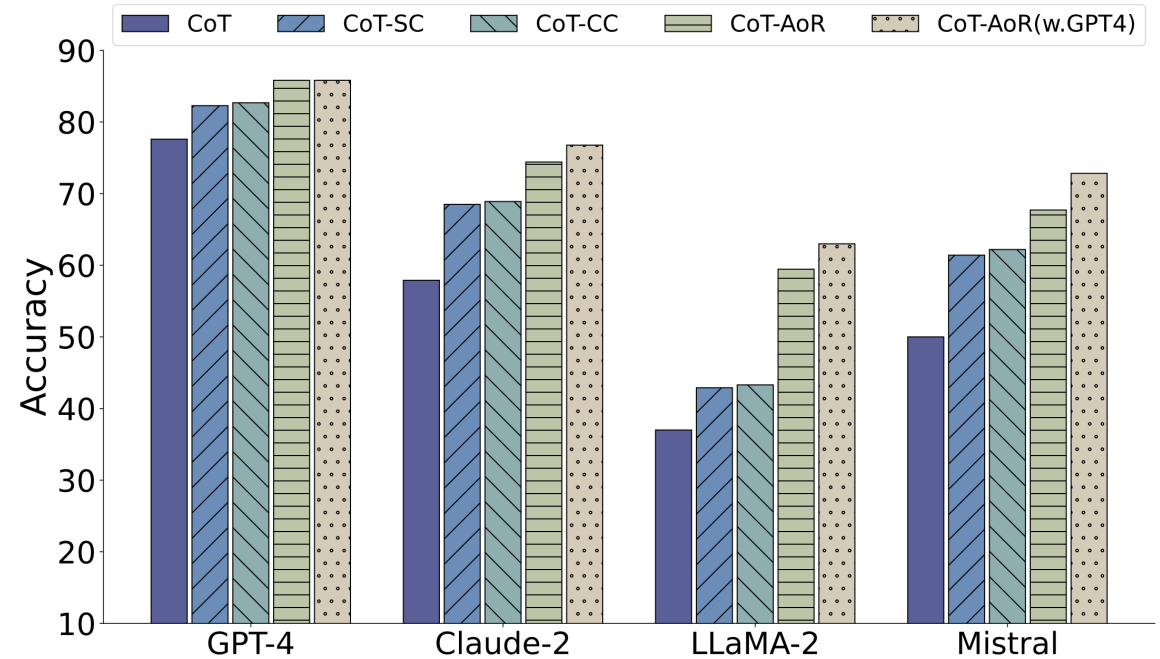


AQuA



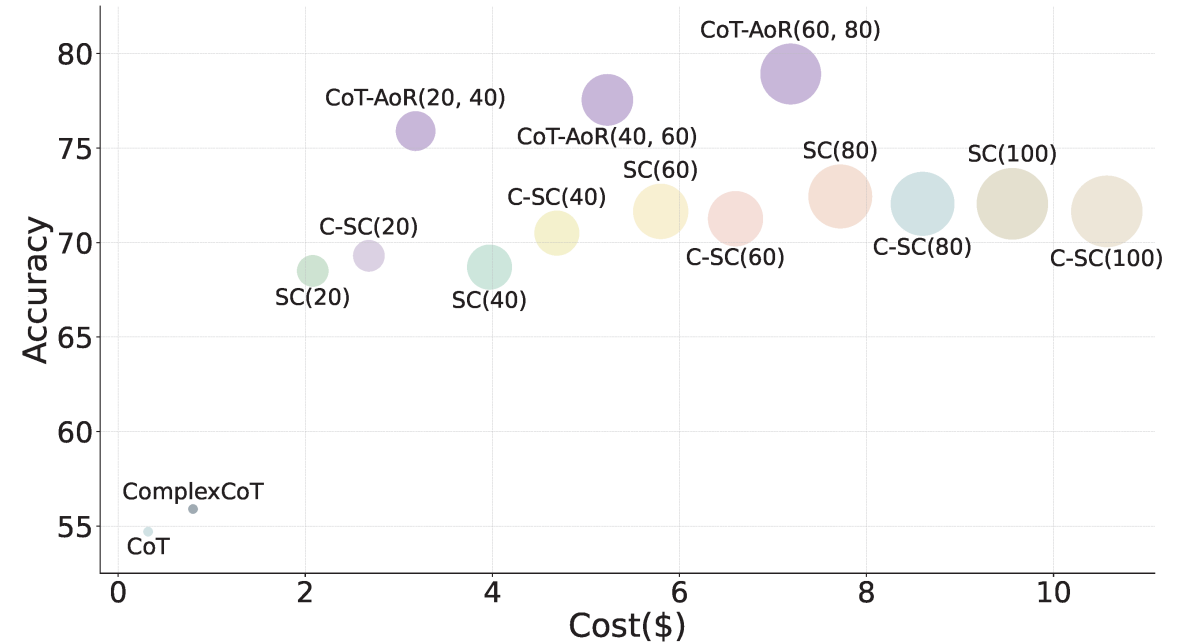
Analysis: AoR on Various LLMs

- ❑ AoR can bring performance improvements to a variety of LLMs.
- ❑ GPT-4: The most powerful and state-of-the-art model from OpenAI.
- ❑ Claude-2: The latest model from Anthropic.
- ❑ LLaMA-2 (70B): An open-source model from Meta.
- ❑ Mistral (8x7B): Sparse Mixture of Experts model from Mistral AI.
- ❑ When using a stronger model (GPT-4) as the evaluator, the potential of AoR can be further unleashed.



Analysis: Cost-Efficiency

- ❑ Compared to Self-Consistency, AoR improved performance by 7.2% while reducing costs by 20%.
- ❑ AoR can overcome the saturation flaw inherent in Self-Consistency, making more effective use of reasoning chains.
- ❑ The AoR method with a sampling limit of 40 surpassed the Self-Consistency method that sampled 100 reasoning chains.



Conclusion

❑ Majority of Incorrect Reasoning Chains in Hard Problems:

- ❑ In challenging scenarios, the model tends to generate a majority of incorrect reasoning chains, rendering traditional methods relying on majority voting ineffective in selecting the correct answers.
- ❑ The prevalence of erroneous reasoning chains underscores the necessity for a more sophisticated approach in answer selection.

❑ Utilization of Model Evaluation:

- ❑ Leveraging the model's evaluation capacity is crucial, as evaluation often proves to be more straightforward and accurate than providing solutions.
- ❑ Effective utilization of evaluation mechanisms enhances the model's ability to discern accurate reasoning chains amidst a sea of possibilities.

❑ Increasing Need for Reasoning Chains with Problem Complexity:

- ❑ As problems become more intricate, the demand for multiple reasoning chains grows, akin to how humans often require more time to tackle difficult problems.

A wide-angle photograph of the Milan skyline at sunset. The Duomo di Milano is prominent on the right, with the snow-capped Alps in the background under a warm, orange-hued sky.

THANK YOU!