Unlocking the Capabilities of Thought: A Reasoning Boundary Framework to Quantify and Optimize Chain-of-Thought





Qiguang Chen[1] Libo Qin[2] Jiaqi Wang[3] Jingxuan Zhou[1] Wanxiang Che[1]

[1]Research Center for Social Computing and Information Retrieval (SCIR), Harbin Institute of Technology

[2]Central South University [3] The Chinese University of Hong Kong

Chain-of-Thought Reasoning

Chain-of-Thought (CoT): Generate a series of reasoning sub-steps while generating answers to improve the performance of problem solving.

[Input]

Question: Xiaoming has 5 ping-pong balls. He bought 2 packs of ping-pong balls, with 3 balls in each pack. How many ping-pong balls does Xiaoming have now?

Answer:

[Output]

Step 1: Xiaoming initially has 5 ping-pong balls.

Step 2: Xiaoming bought 2 packs of ping-pong balls, with 3 balls in each pack.

Step 3: After buying the 2 packs, he acquired an additional $2 \times 3 = 6$ balls.

Step 4: In total, Xiaoming has $5 + 6 = 11$ ping-pong balls.

Final answer: 11 balls.

Chain-of-Thought Reasoning

There exists a reasoning upper boundary during the Chain-of-Thought process.

- The single-step mathematical computation has a boundary and cannot solve problems with excessively long input sequences.

Computation Boundary

Arithmetic Expression

Input:

$$(7+5) \div (6+4\times 3-2\times 7) =$$

Output:

$$\begin{aligned} & 12 \div (6+4\times 3-2\times 7) \\ =\ & 12 \div (6+12-2\times 7) \\ =\ & 12 \div (18-2\times 7) \\ =\ & 12 \div (18-14) \\ =\ & 12 \div 4 \\ =\ & 3 \end{aligned}$$

Linear Equations

Input:

$$\begin{aligned} 3x+3y+12z&=6; \\ 2x+5y+14z&=7; \\ 2x+4y+15z&=6; \end{aligned}$$

$$\Longrightarrow$$

Output:

$$\begin{aligned} x+1y+4z&=2; \\ 3y+6z&=3; \\ 2y+7z&=2; \\ \Longrightarrow\quad x+2z&=1; \\ y+2z&=1; \\ 3z&=0; \\ \Longrightarrow\quad x&=1; \\ y&=1; \\ z&=0; \end{aligned}$$

Towards Revealing the Mystery behind Chain of Thought: A Theoretical Perspective. Feng et al., NeurIPS 2023

Chain-of-Thought Reasoning

There exists a reasoning upper boundary during the Chain-of-Thought process.

- There is also a boundary to the planning capabilities, making it unable to handle excessively long planning chains.

- Computation Boundary
- Planning Boundary

Domain	Method	Instances correct				
		GPT-4	GPT-3.5	I-GPT3.5	I-GPT3	GPT-3
Blocksworld (BW)	One-shot	206/600 (34.3%)	37/600 (6.1%)	54/600 (9%)	41/600 (6.8%)	6/600 (1%)
	Zero-shot	210/600 (34.6%)	8/600 (1.3%)	-	-	-
	COT	214/600 (35.6%)	-	-	-	-
Logistics Domain	One-shot	28/200 (14%)	1/200 (0.5%)	6/200 (3%)	3/200 (1.5%)	-
	Zero-shot	15/200 (7.5%)	1/200 (0.5%)	-	-	-
Mystery BW (Deceptive)	One-shot	26/600 (4.3%)	0/600 (0%)	4/600 (0.6%)	14/600 (2.3%)	0/600 (0%)
	Zero-shot	1/600 (0.16%)	0/600 (0%)	-	-	-
	COT	54/600 (9%)	-	-	-	-
Mystery BW (Randomized)	One-shot	12/600 (2%)	0/600 (0%)	5/600 (0.8%)	5/600 (0.8%)	1/600 (0.1%)
	Zero-shot	0/600 (0%)	0/600 (0%)	-	-	-

On the Planning Abilities of LaRBe Language Models : A Critical Investigation. Valmeekam et al., NeurIPS 2023

Motivation

Problems with existing work:

- It only conducted qualitative analysis and did not perform quantitative analysis of the reasoning boundary.
- It did not provide guidance on optimizing Chain-of-Thought (CoT).

◄ MATHEMATICAL ► Solve a grade-school level math reasoning problems

Question: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?
Thought: Shawn started with 5 toys. If he got 2 toys each from his mom and dad, then that is 4 more toys. 5 + 4 = 9.
Symbols: Numbers: 5, 4, 9
Patterns: Equations: 5 + 4 = 9. The equations typically appear at the end of the thought, and are almost always involved in generating the final answer.

◄ COMMONSENSE ► (SPORTS) Verify the accuracy of a statement linking an athlete with a sport.

Question: Is the following sentence plausible? "Jamal Murray was perfect from the line."'
Thought: Jamal Murray is a basketball player. Being perfect from the line is part of basketball.
Symbols: Person and activity: **Jamal Murray**, **Being perfect from the line**
Patterns: Consistent sentence structure PERSON *belongs to* SPORT. ACTIVITY *belongs to* SPORT, where *belongs to* is a phrase that connects a sports personality with an activity. The answer is yes if both the person and the activity are associated with the same sport.

◄ COMMONSENSE ► (DATE) Reason about dates

Question: It is 4/19/1969 today. What is the date 24 hours later in MM/DD/YYYY?
Thought: Today is 04/19/1969. 24 hours later is one day after today, which would be 04/20/1969. The answer is 04/20/1969.
Symbols: Dates: **04/19/1969**, **04/20/1969**
Patterns: Reasoning flows in two steps: **initial calculation** (*Today is 04/19/1969...*), followed by **generation of output** (*The answer is...*)

◄ SYMBOLIC ► (SORTING) Sort integers between 1-9

Question: 3, 1, 2, 7, 8, 5, 6, 9, 4
Thought: 1 < 2 < ... < 9
Symbols: Numbers: **2**, **4**, **9**
Patterns: Smaller number < larger number (**1 < 2**)

Table 1: ◄ **Symbols** ► and ◄ **Patterns** ► for different tasks.

Lack of quantitative analysis

Arithmetic Expression

Input:
$(7+5) \div (6+4\times 3-2\times 7) =$

Output:
$$\begin{aligned} & 12 \div (6+4\times 3-2\times 7) \\ = & 12 \div (6+12-2\times 7) \\ = & 12 \div (18-2\times 7) \\ = & 12 \div (18-14) \\ = & 12 \div 4 \\ = & 3 \end{aligned}$$

Linear Equations

Input:
$$\begin{aligned} 3x+3y+12z &= 6; \\ 2x+5y+14z &= 7; \\ 2x+4y+15z &= 6; \end{aligned}$$
$\Longrightarrow$

Output:
$$\begin{aligned} x+1y+4z &= 2; \\ 3y+6z &= 3; \\ 2y+7z &= 2; \\ \Longrightarrow \quad x+2z &= 1; \\ y+2z &= 1; \\ 3z &= 0; \\ \Longrightarrow \quad x &= 1; \\ y &= 1; \\ z &= 0; \end{aligned}$$

Absence of optimization guidance

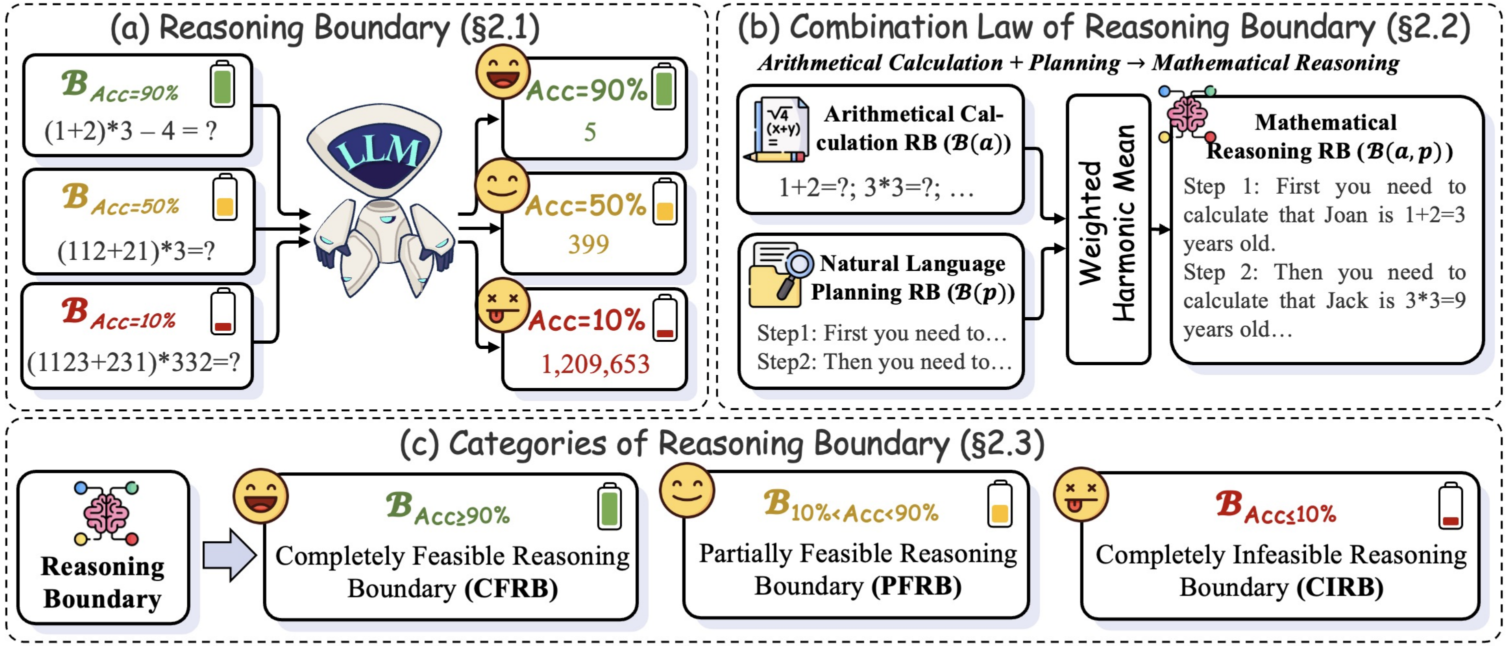
A Reasoning Boundary Framework to Quantify and Optimize CoT

Contributions & Main Framework

- First Systematically define and comprehensively validate reasoning boundary
- Conduct quantitative analysis of the reasoning boundary
- Propose Minimal Acceptable Reasoning Path Prompting to optimize the performance.

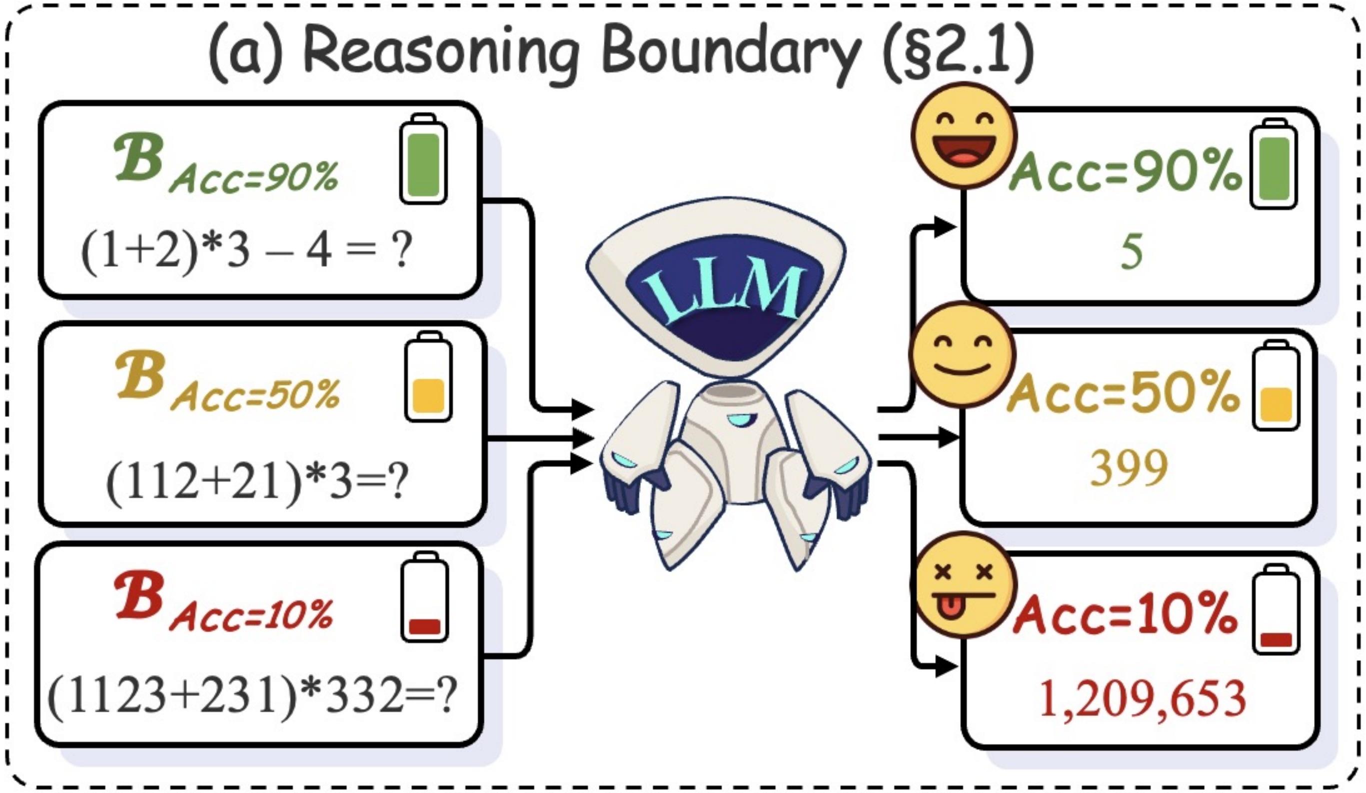
Reasoning Boundary Framework

Reasoning Boundary Hypothesis → Combined Reasoning Boundary Hypothesis → Minimal Acceptable Reasoning Path Chain-of-Thought



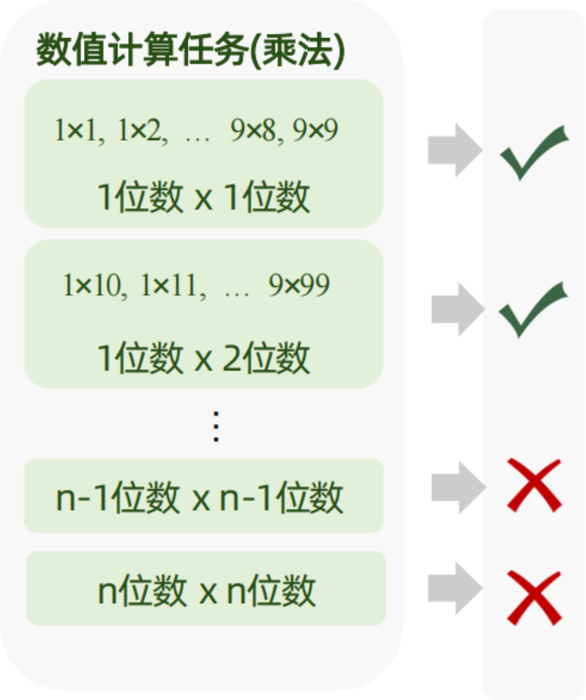
Reasoning Boundary Hypothesis

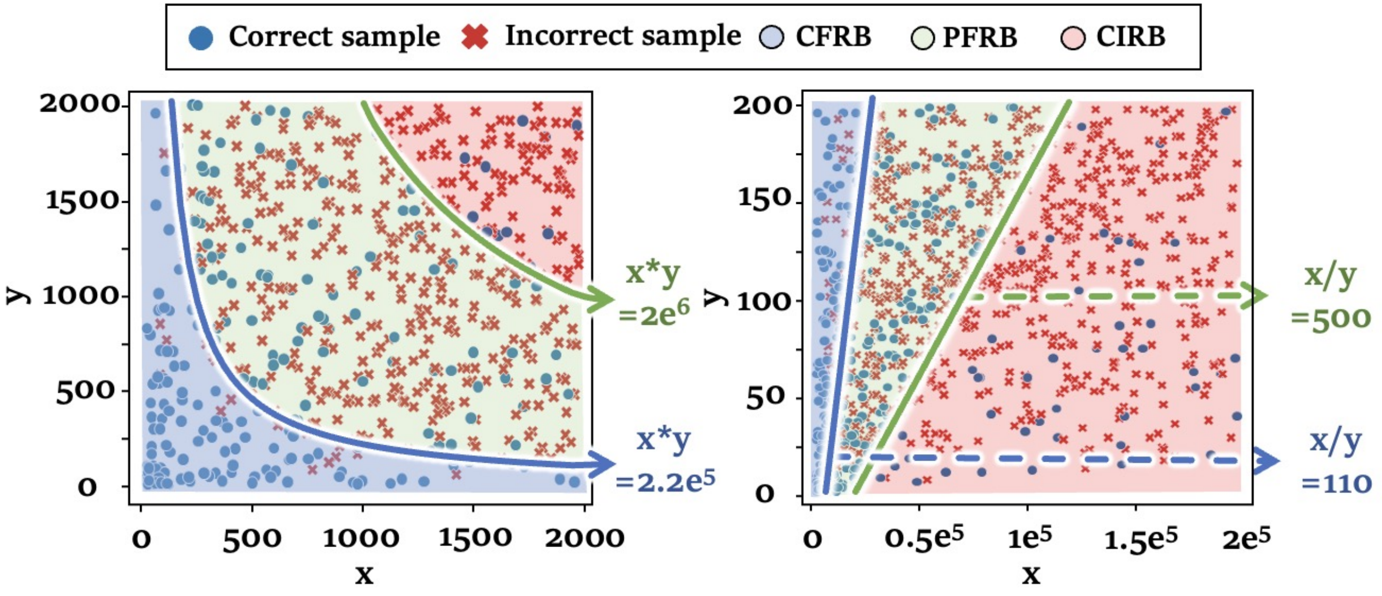
For certain tasks and models, during the CoT reasoning, each reasoning capability has an upper-bound, known as the **reasoning boundary**. Exceeding this boundary prevents reasoning from proceeding as expected.



Reasoning Boundary Hypothesis

Computation capability has the reasoning boundary.





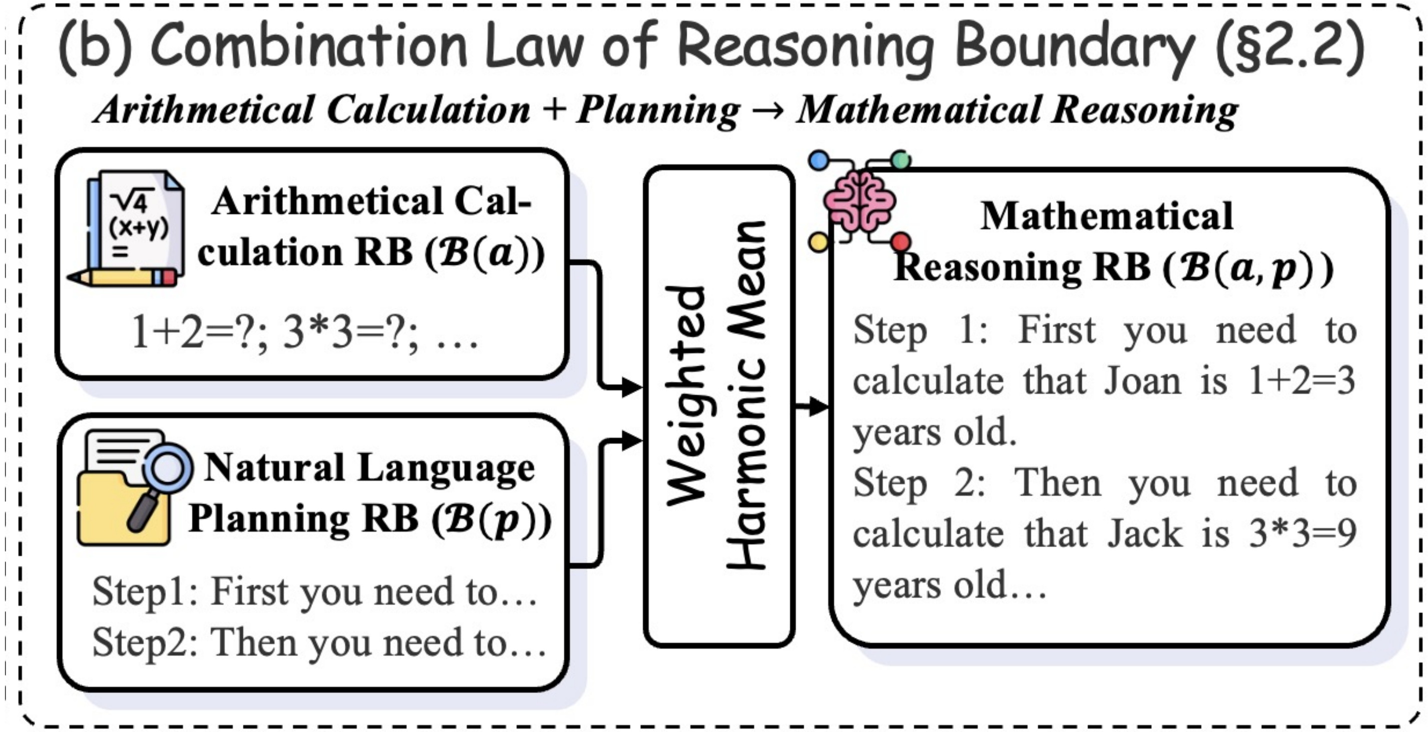
(a) Distribution of correct predictions for x*y samples.

(b) Distribution of correct predictions for x/y samples.

Combined Reasoning Boundary Hypothesis

For real-world tasks, LLMs need to utilize more different fundamental boundaries for combined reasoning to solve problems.

- Practical combined reasoning boundary can be calculated as the weighted harmonic mean of the fundamental reasoning boundaries.

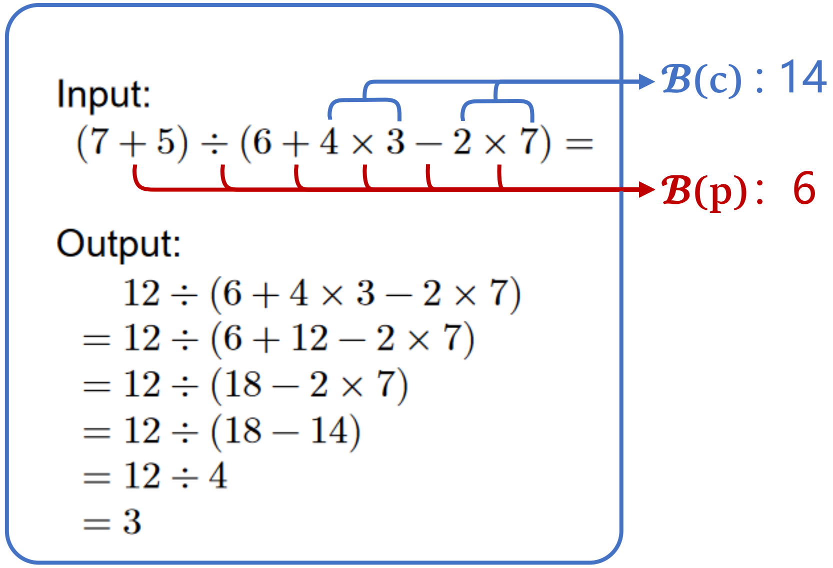


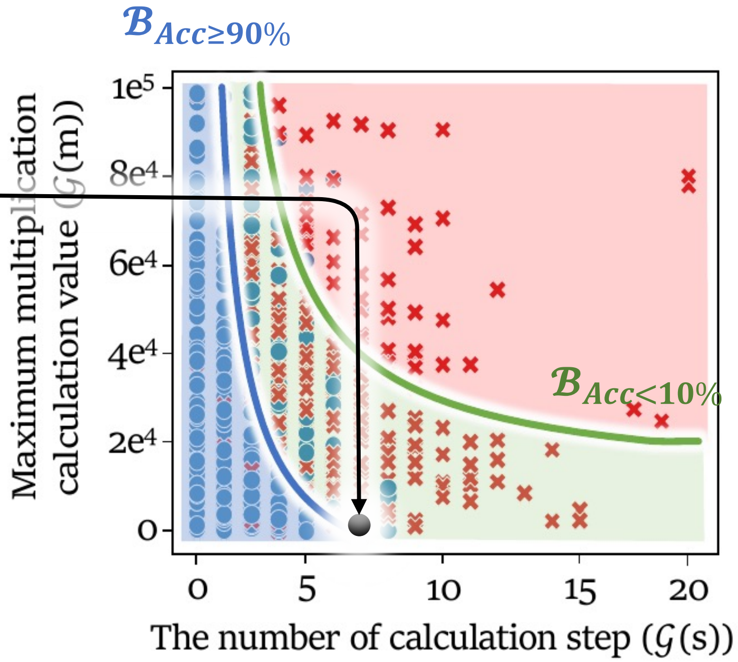
Combined Reasoning Boundary Hypothesis

Task: Multi-step Mathematical Calculations

Observation: The Combined reasoning boundaries are computed as the **weighted harmonic mean.**

Multi-step mathematical calculations

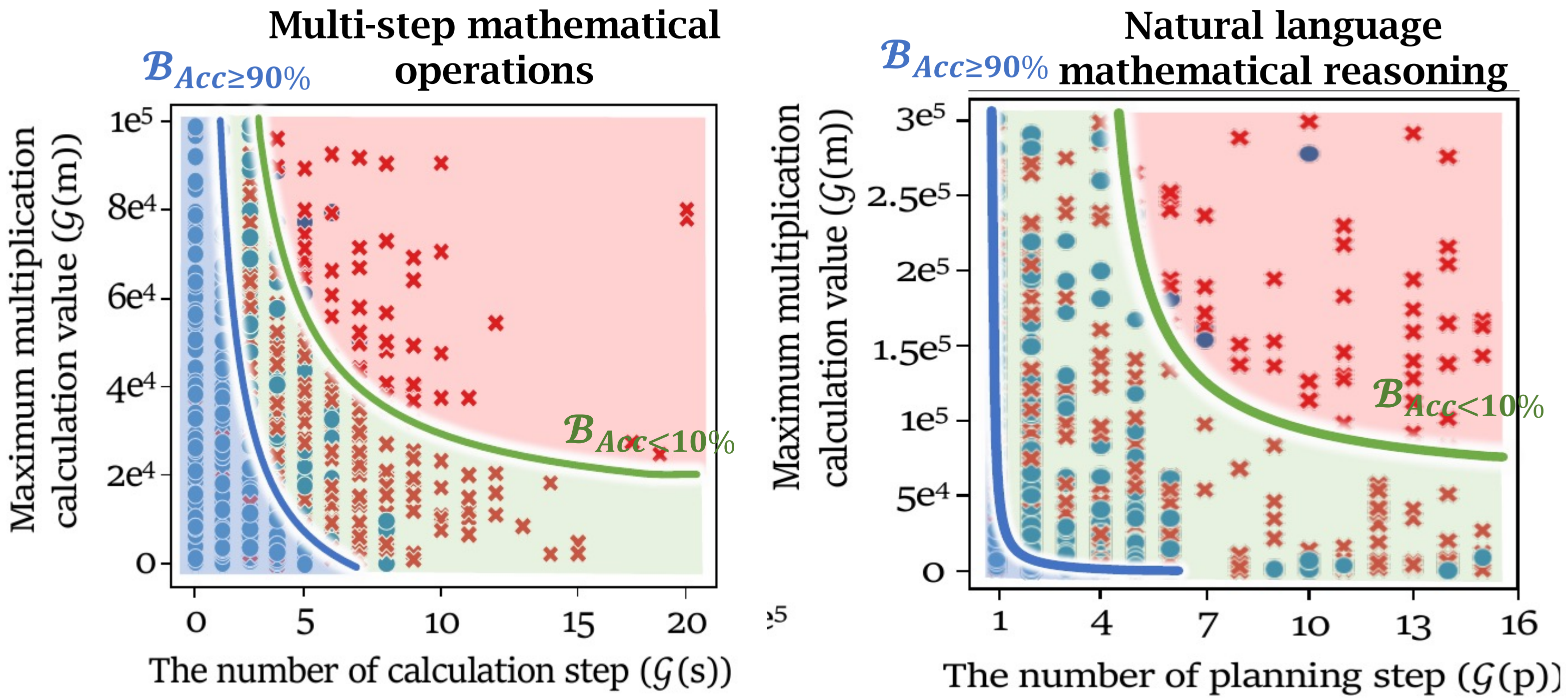




$$\mathcal{B}^{\text{CoT}}(c,p) = \frac{1}{\frac{N_{11}}{(\mathcal{B}(c)-b_1)} + \frac{N_{21}}{(\mathcal{B}(p)-b_2)}}.$$

Combined Reasoning Boundary Hypothesis

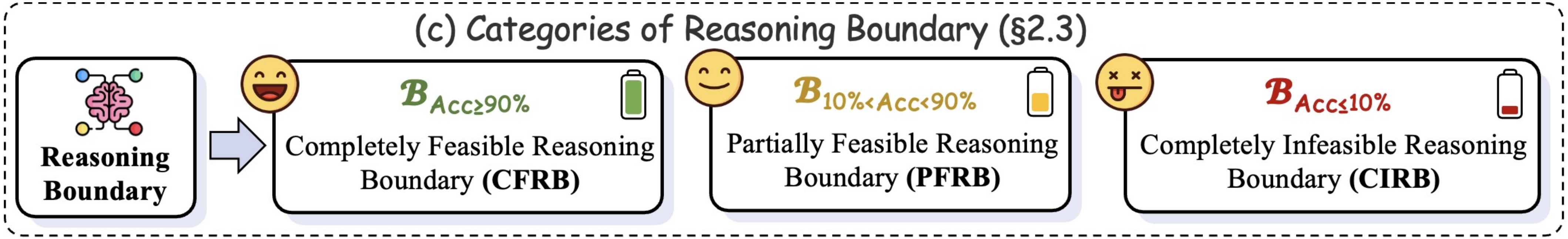
In various **tasks and models**, the Combined reasoning boundaries are computed as the weighted harmonic mean.

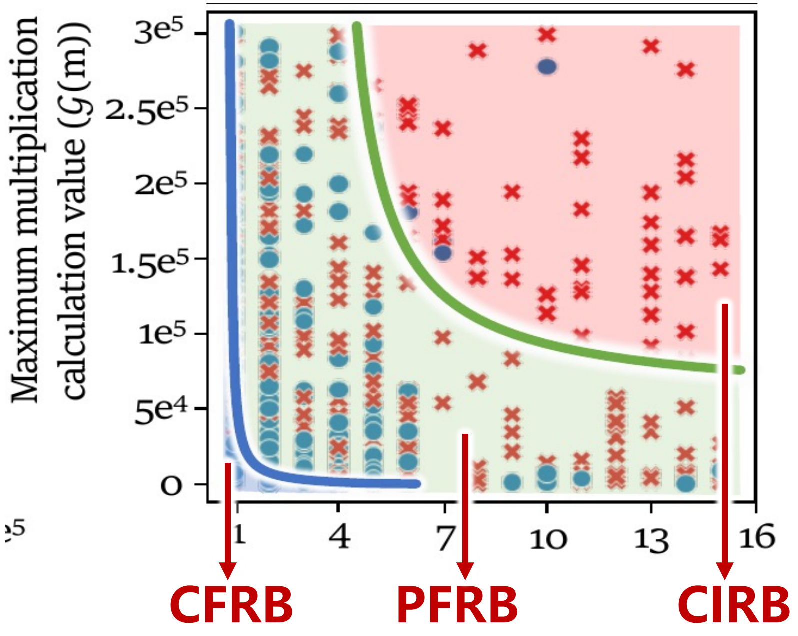


$$\mathcal{B}^{\texttt{CoT}}(c,p) = \frac{1}{\frac{N_{11}}{(\mathcal{B}(c)-b_1)} + \frac{N_{21}}{(\mathcal{B}(p)-b_2)}}.$$

Classification of Reasoning Boundaries

We have divided the reasoning boundary into three regions:

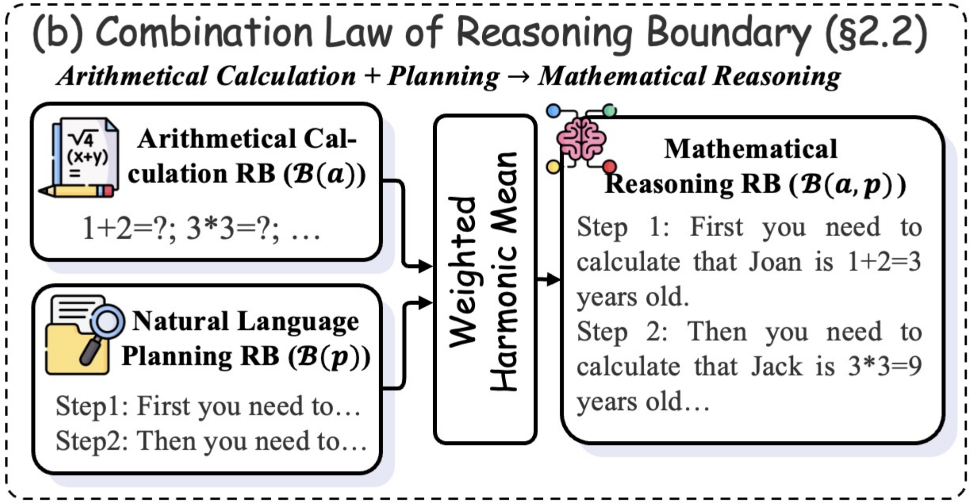




Optimization of Reasoning Boundaries

Optimization of combined reasoning boundaries:

- Effective optimization of combined reasoning boundaries
- Optimizing the reasoning path within a fixed reasoning boundary



Optimizing the Combined RB to enhance capabilities

- Tool Usage
- Program of Thought
- …

Optimizing the reasoning path within a fixed RB to reduce reasoning difficulty

- Least-to-Most
- Complex-CoT
- …

Optimization of Reasoning Boundaries

In practical scenarios, when the model framework cannot optimize RB, we should focus on optimizing the problem itself.

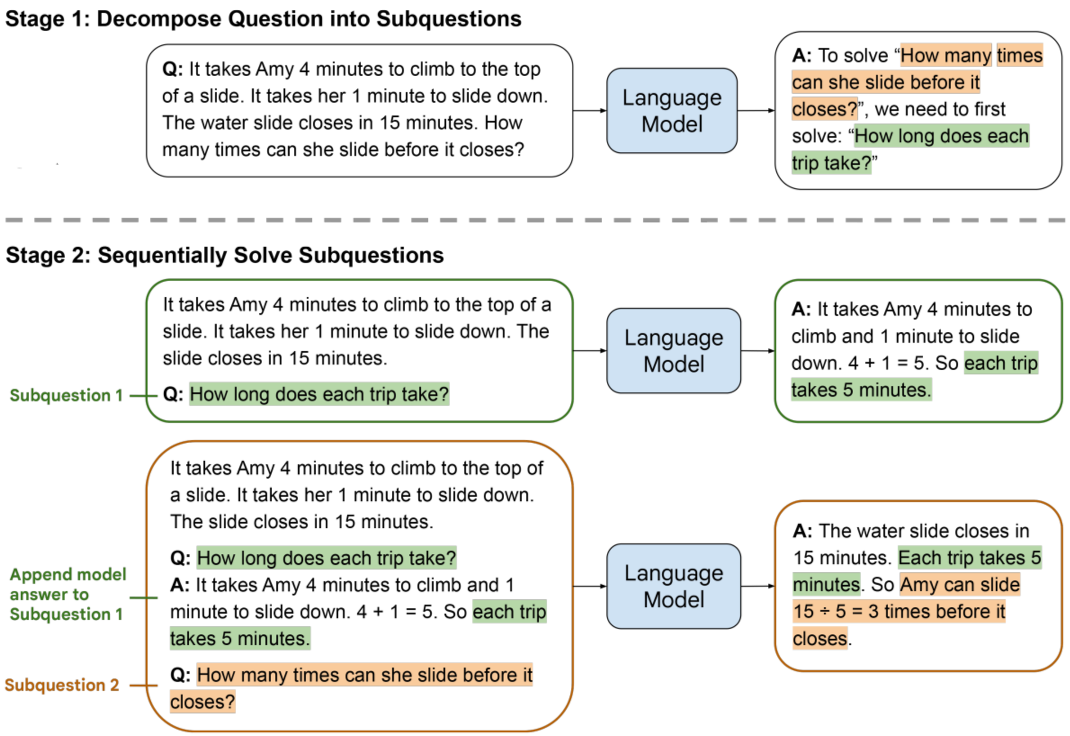
By reducing difficulty is less than the original RB, the model can achieve better results.

- **Optimizing Planning Difficulty:** Least-to-Most
- **Optimizing Calculation Difficulty:** Complex CoT

Optimization of Reasoning Boundaries

Optimizing Planning Difficulty: Least-to-Most

- **Definition:** Divide the problem into subproblems, planning only a few steps at a time.



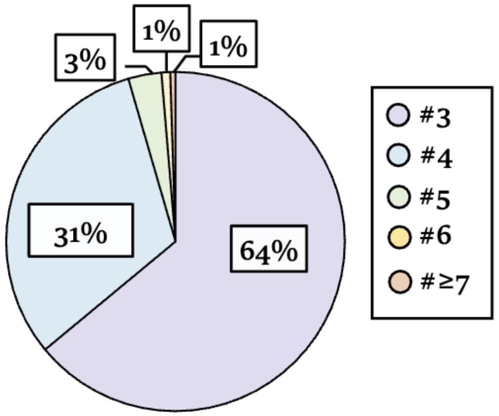
Model	BigGSM		
	Acc. (↑)	Input Token (↓)	Output Token (↓)
CoT	57.00 ±0.93	780.43	96.76 ±3.22
RG-Optimized Methods			
Tool Usage	71.64 ±0.66	688.43	129.53 ±3.82
PoT	78.25 ±1.09	657.43	78.25 ±1.09
Reasoning-Path-Optimized Methods			
Least-to-most	58.25 ±3.28	679.59	176.09 ±15.22

Optimization of Reasoning Boundaries

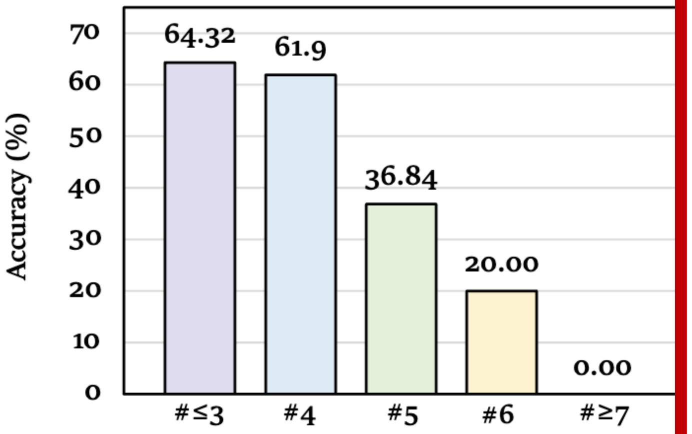
Optimizing Planning Difficulty: Least-to-Most

- **Definition:** Divide the problem into subproblems, planning only a few steps at a time.
- **Drawback:** Introducing additional RB for global planning the overall problem.

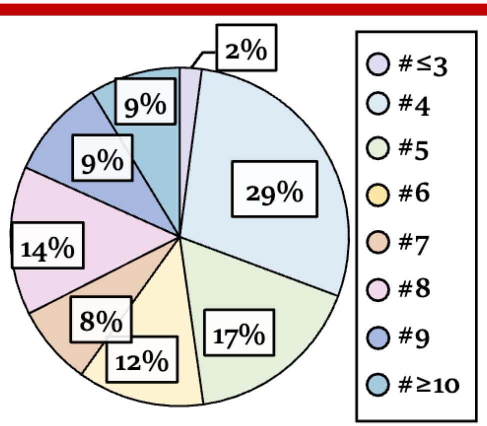
Excessive global planning



(a) Distribution of **the number of CoT steps** performed for each sub-question.



(b) The **Accuracy Distribution** on the CoT steps performed for each sub-question.



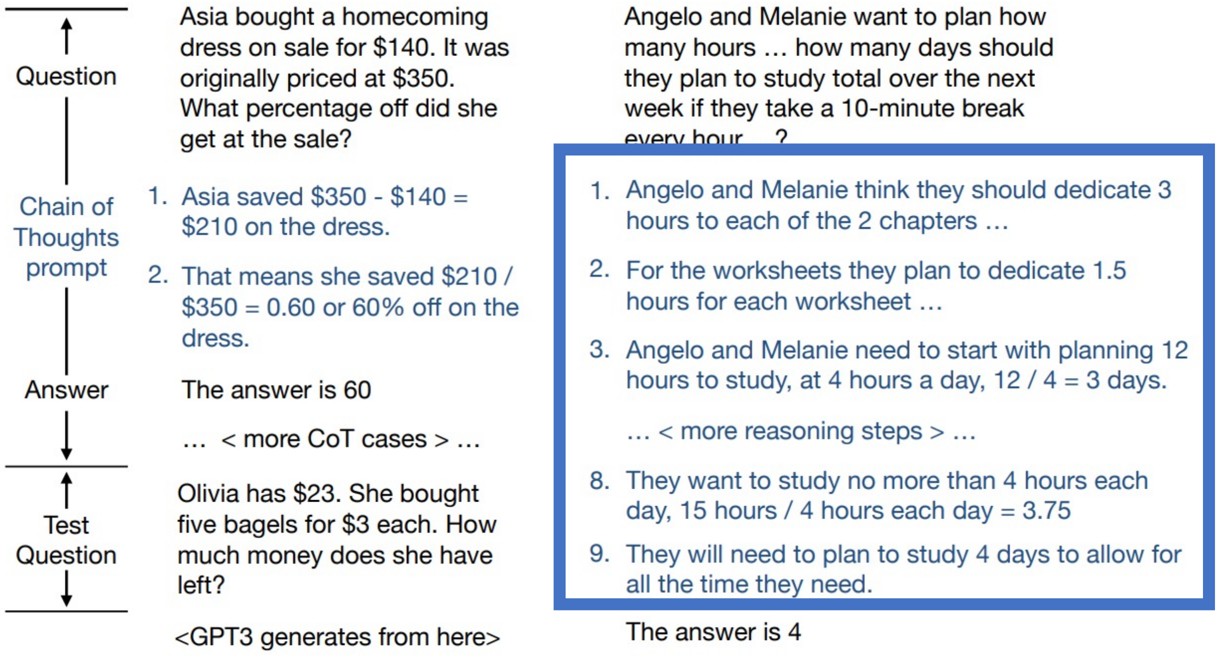
(c) Distribution of the **number of sub-question**.

$$\mathcal{B}^{\texttt{LtM}}(d,p,c) = \frac{1}{\frac{2N_{11}}{(\mathcal{B}'(c)-b_1)} + \frac{2N_{21}}{(\mathcal{B}(p)-b_2)} + \frac{2N_{31}}{(\mathcal{B}(d)-b_3)}}$$

Optimization of Reasoning Boundaries

Optimizing Calculation Difficulty: Complex CoT

- **Definition:** Increase the number of steps to simplify the calculation difficulty.



A. Workflow of chain of thoughts prompting **B.** Example complex chain, 9 reasoning steps

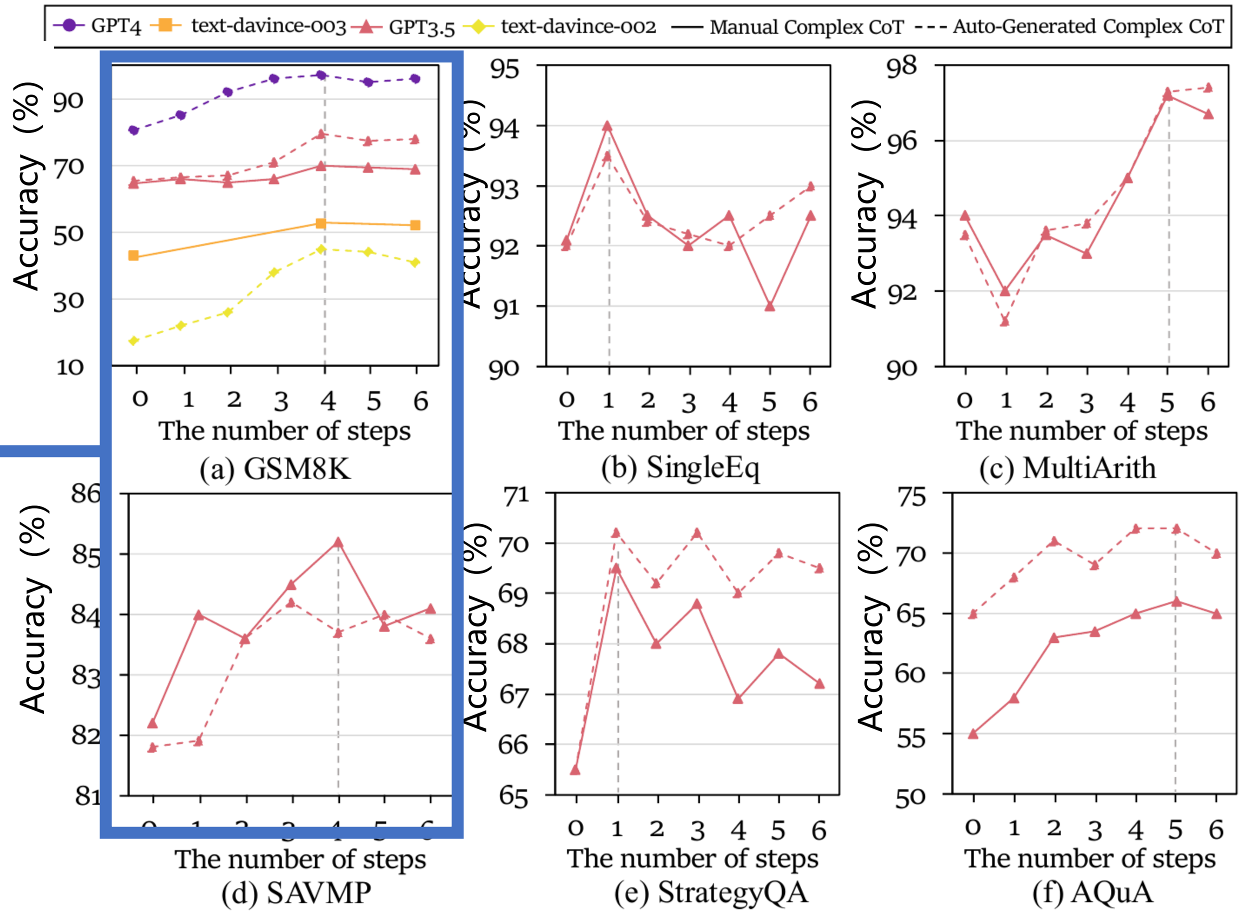
Model	BigGSM		
	Acc. (↑)	Input Token (↓)	Output Token (↓)
CoT	57.00 ±0.93	780.43	96.76 ±3.22
RG-Optimized Methods			
Tool Usage	71.64 ±0.66	688.43	129.53 ±3.82
PoT	78.25 ±1.09	657.43	78.25 ±1.09
Reasoning-Path-Optimized Methods			
Least-to-most	58.25 ±3.28	679.59	176.09 ±15.22
Complex-CoT	59.78 ±0.60	1111.43	131.82 ±1.91

Optimization of Reasoning Boundaries

Optimizing Calculation Difficulty: Complex CoT

- **Definition:** Increase the number of steps to simplify the calculation difficulty.
- **Drawback:** Also increase the overall planning complexity.

Excessive required steps in the prompt render Complex CoT ineffective.



Optimization of Reasoning Boundaries

Minium Acceptable Reasoning Path Prompting (MARP): Based on the maximum RB, complex language processing tasks can be broken down into fewer, model-suitable reasoning steps.

Question: Leo's homework is divided into three parts. He completed the first part in 25 minutes and took twice as long to complete the second part. If he can complete the entire homework in 2 hours, how many minutes does Leo have left to complete the third part?

Original Example Sample:

1. Leo spent 25 x 2 = 50 minutes completing the second part of his homework.
2. Leo spent 25 + 50 = 75 minutes completing the first and second parts.
3. He spent 60×2=120 minutes on the entire homework.
4. Therefore, Leo spent 120−75=45 minutes on the third part of his homework.

45

Optimized Instruction:

Requirements:… Each step should include as many basic operations as possible.

Constraints: … Each step can contain a maximum of 5 basic operations.…

Optimized Example Sample:

1. Leo spent 25 + 25 x 2 = 75 minutes completing the first and second parts of his homework.
2. Therefore, Leo spent 2 x 60 - 75 = 45 minutes on the third part of his homework.

45

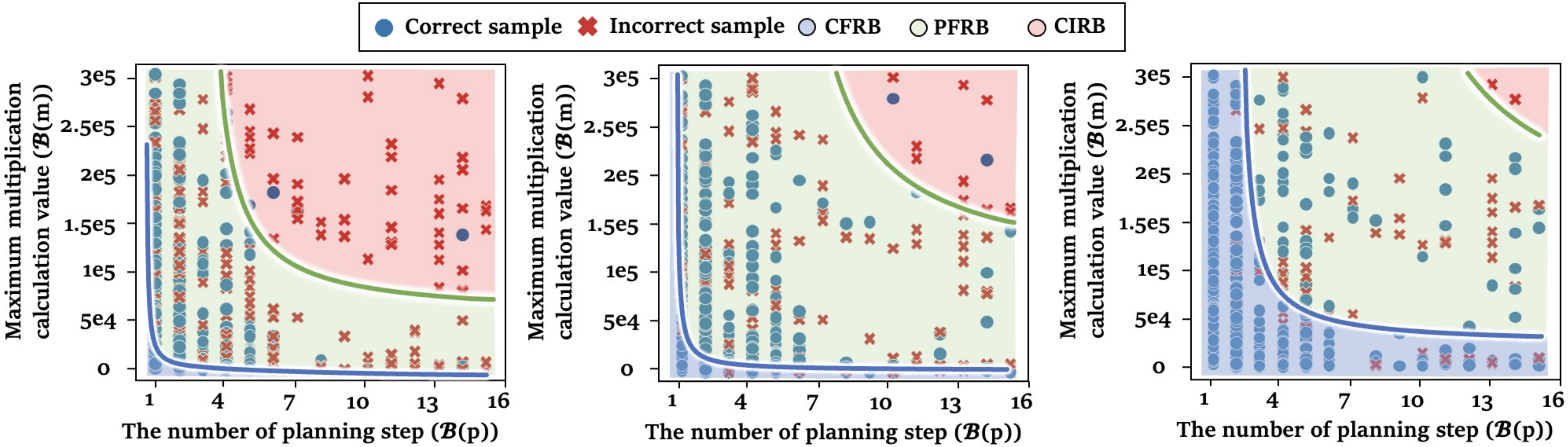
Achieving SOTA Performance!

Exploration:

Discussion on O1

Exploration: Discussion on O1

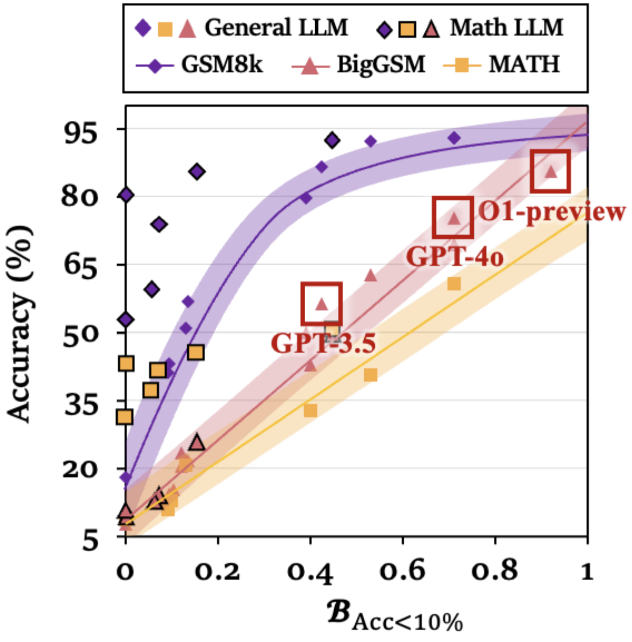
Different reasoning boundaries progressively improve as the model is optimized.



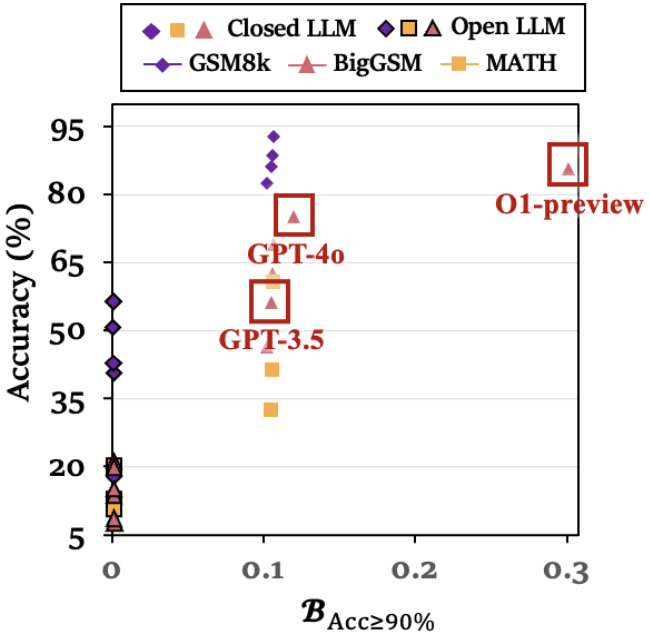
(a) GPT-3.5-turbo (b) GPT-4o (c) O1-preview

Exploration: Discussion on O1

- The CIRB shows a significant improvement, with a **linear** trend.
- The CFRB boundary experiences significant **nonlinear** gains through reinforcement learning combined with Inference Law.



(a) Correlation between the values of CIRB $\mathcal{B}_{Acc<10\%}$ for different **general LLMs** and performance on real benchmarks.



(b) Correlation between the values of CFRB $\mathcal{B}_{Acc\geq90\%}$ for different **closed and open LLMs** and performance on real benchmarks.

Thank you & QA

Paper

Code

Dataset