北京大学
PEKING UNIVERSITY

Chain of Thoughts

Kun Yuan

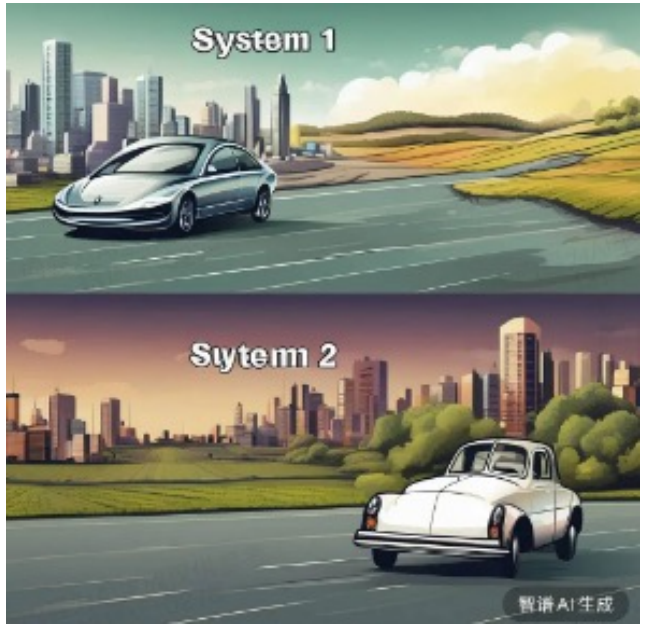
Center for Machine Learning Research @ Peking University

Two systems in our mind

北京大学 PEKING UNIVERSITY

- **System-I (slow system)** tasks can be done quickly and intuitively by humans
  - E.g., sentiment analysis; topic classification

1. **快系统（系统1）**：这是一个快速、自动、经常是无意识的思考过程。它负责直觉反应和快速的判断，如识别物体、人脸，或者对简单问题的迅速回答。快系统依赖习惯、经验以及情感，但同时也可能产生认知偏差和错误。



- **System-II (fast system)** tasks require slow and deliberate thinking
  - E.g., logical, mathematical, and commonsense reasoning

2. **慢系统（系统2）**：这是一个慢速、逻辑性、需要意志努力的思考过程。它涉及到复杂的计算和决策，如解决数学问题或做出重要的决策。慢系统更加注重事实和逻辑，能够纠正快系统可能产生的错误，但它的过程需要更多的精力和时间。

LLM performs well for fast-system tasks

- When model size gets larger, performances on slow systems get significantly better

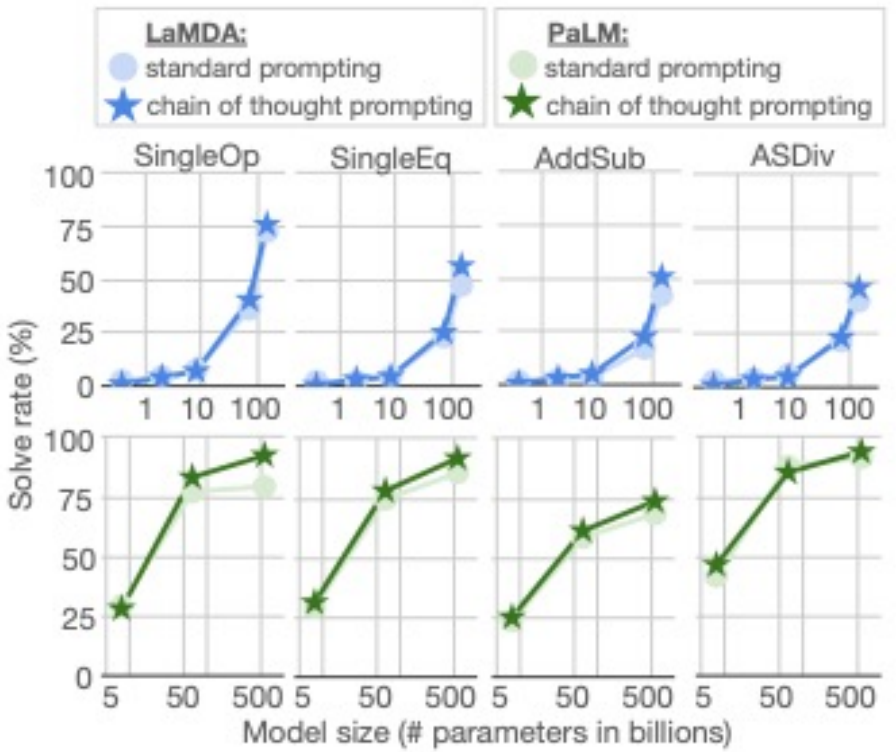


Figure 2. When scaling up the model already facilitates good performance, chain of thought prompting does as well or better.

LLM performs badly for slow-system tasks

- When model size gets larger, performances on slow systems improve slowly

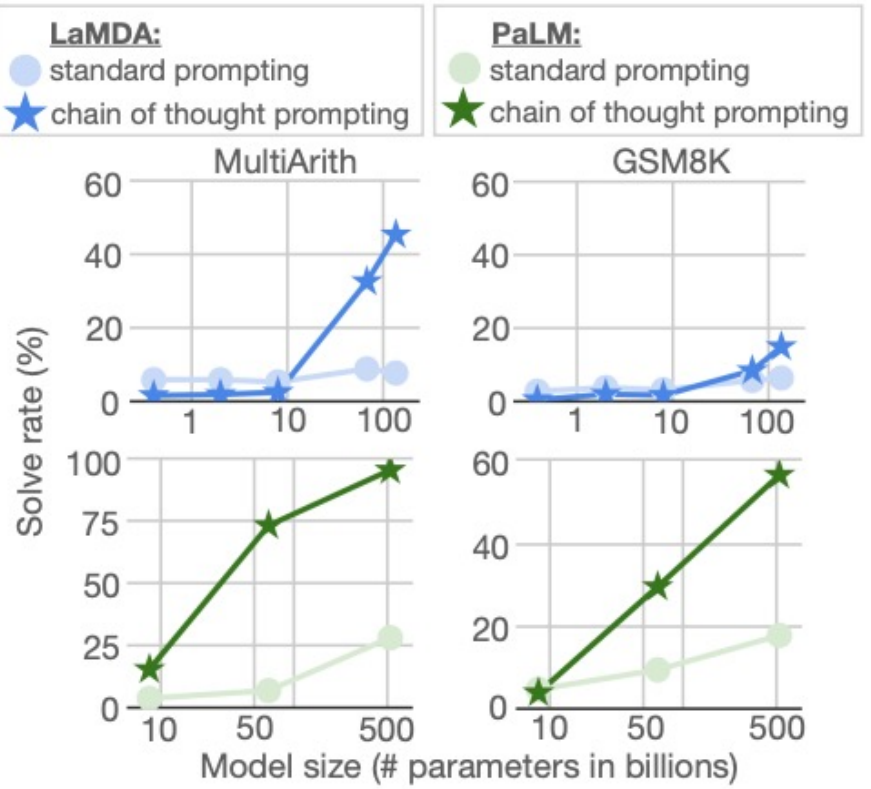


Figure 3. Employing chain of thought enables language models to solve challenging math word problems for which standard prompting has a mostly flat scaling curve.

Chain-of-Thought Prompting Elicits Reasoning in Large Language Models

Jason Wei **Xuezhi Wang** **Dale Schuurmans** **Maarten Bosma**
Brian Ichter **Fei Xia** **Ed H. Chi** **Quoc V. Le** **Denny Zhou**

Google Research, Brain Team
{jasonwei,dennyzhou}@google.com

Chain of thoughts can help slow-system tasks

北京大学 PEKING UNIVERSITY

Standard prompting

Input:
Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?
A: The answer is 11.

...

Q: John takes care of 10 dogs. Each dog takes .5 hours a day to walk and take care of their business. How many hours a week does he spend taking care of dogs?
A:

Model output:
The answer is 50. ✗

Chain of thought prompting

Input:
Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?
A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. 5 + 6 = 11. The answer is 11.

...

Q: John takes care of 10 dogs. Each dog takes .5 hours a day to walk and take care of their business. How many hours a week does he spend taking care of dogs?
A:

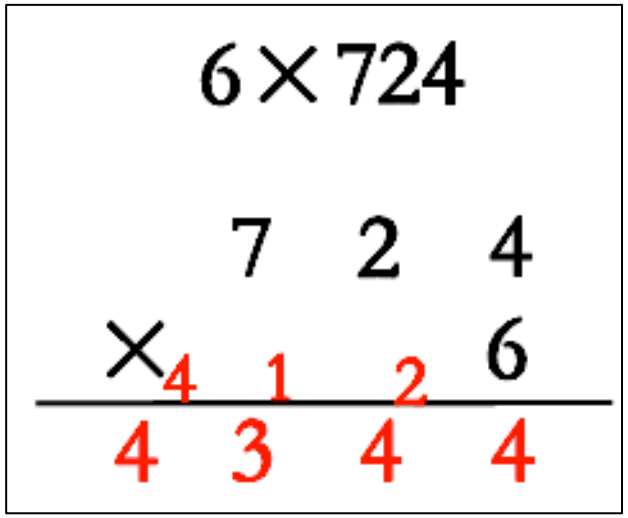
Model output:
John takes care of 10 dogs. Each dog takes .5 hours a day to walk and take care of their business. So that is 10 x .5 = 5 hours a day. 5 hours a day x 7 days a week = 35 hours a week. The answer is 35 hours a week. ✓

What is chain of thought (CoT)?

Definition: a coherent series of short sentences that lead to the answer for a reasoning problem

Prompt with COT will mimic the reasoning process of humans

Humans typically decompose the problem into intermediate steps and solve each before giving the final answer



Why is CoT useful?

Allow models to decompose multi-step problems into intermediate steps, which means that additional computation can be allocated to problems that require more reasoning steps

Instruct LLM on how it might have arrived at a particular answer

Allow LLM to solve any slow-system problem which can be expressed by human language

How to write CoT?

北京大学 PEKING UNIVERSITY

Math reasoning

Table 20. Few-shot exemplars for full chain of thought prompt for math word problems. This set of exemplars was used for all math word problem datasets.

PROMPT FOR MATH WORD PROBLEMS

Q: There are 15 trees in the grove. Grove workers will plant trees in the grove today. After they are done, there will be 21 trees. How many trees did the grove workers plant today?

A: There are 15 trees originally. Then there were 21 trees after some more were planted. So there must have been 21 - 15 = 6. The answer is 6.

Q: If there are 3 cars in the parking lot and 2 more cars arrive, how many cars are in the parking lot?

A: There are originally 3 cars. 2 more cars arrive. 3 + 2 = 5. The answer is 5.

Q: Leah had 32 chocolates and her sister had 42. If they ate 35, how many pieces do they have left in total?

A: Originally, Leah had 32 chocolates. Her sister had 42. So in total they had 32 + 42 = 74. After eating 35, they had 74 - 35 = 39. The answer is 39.

Q: Jason had 20 lollipops. He gave Denny some lollipops. Now Jason has 12 lollipops. How many lollipops did Jason give to Denny?

A: Jason started with 20 lollipops. Then he had 12 after giving some to Denny. So he gave Denny 20 - 12 = 8. The answer is 8.

How to write CoT?

北京大学 PEKING UNIVERSITY

Symbolic reasoning

Table 4. Few-shot exemplars for chain of thought prompting for symbolic reasoning datasets. Chains of thought are highlighted. Exact prompts with all exemplars are shown in Table 21–Table 23 in the Appendix.

LAST LETTER CONCATENATION **Q:** Take the last letters of the words in "Elon Musk" and concatenate them. **A:** The last letter of "Elon" is "n". The last letter of "Musk" is "k". Concatenating them is "nk". So the answer is nk.
REVERSE LIST **Q:** Reverse the sequence "glasses, pen, alarm, license". **A:** First is glasses. Second is pen. Third is alarm. Fourth is license. Now to reverse, change the order to: Fourth is license. Third is alarm. Second is pen. First is glasses. So the answer is "license, alarm, pen, glasses".
COIN FLIP **Q:** A coin is heads up. Maybelle flips the coin. Shalonda does not flip the coin. Is the coin still heads up? **A:** The coin was flipped by Maybelle. So the coin was flipped 1 time, which is an odd number. The coin started heads up, so after an odd number of flips, it will be tails up. So the answer is no.

How to write CoT?

北京大学 PEKING UNIVERSITY

Common sense reasoning

Table 5. Few-shot exemplars for chain of thought prompting for commonsense reasoning datasets. Chains of thought are highlighted. Exact prompts with all exemplars are shown in Table 24–Table 27 in the Appendix.

COMMONSENSEQA

Q: Sammy wanted to go to where the people were. Where might he go?
Options: (a) race track (b) populated areas (c) desert (d) apartment (e) roadblock

A: The answer must be a place with a lot of people. Race tracks, desert, apartments, and roadblocks don't have a lot of people, but populated areas do. So the answer is (b).

STRATEGYQA

Q: Yes or no: Could a llama birth twice during War in Vietnam (1945-46)?

A: The War in Vietnam was 6 months. The gestation period for a llama is 11 months. So a llama could not give birth twice during the War in Vietnam. So the answer is no.

DATE UNDERSTANDING

Q: The concert was scheduled to be on 06/01/1943, but was delayed by one day to today. What is the date 10 days ago in MM/DD/YYYY?

A: One day after 06/01/1943 is 06/02/1943, so today is 06/02/1943. 10 days before today is 05/23/1943. So the answer is 05/23/1943.

CoT performance

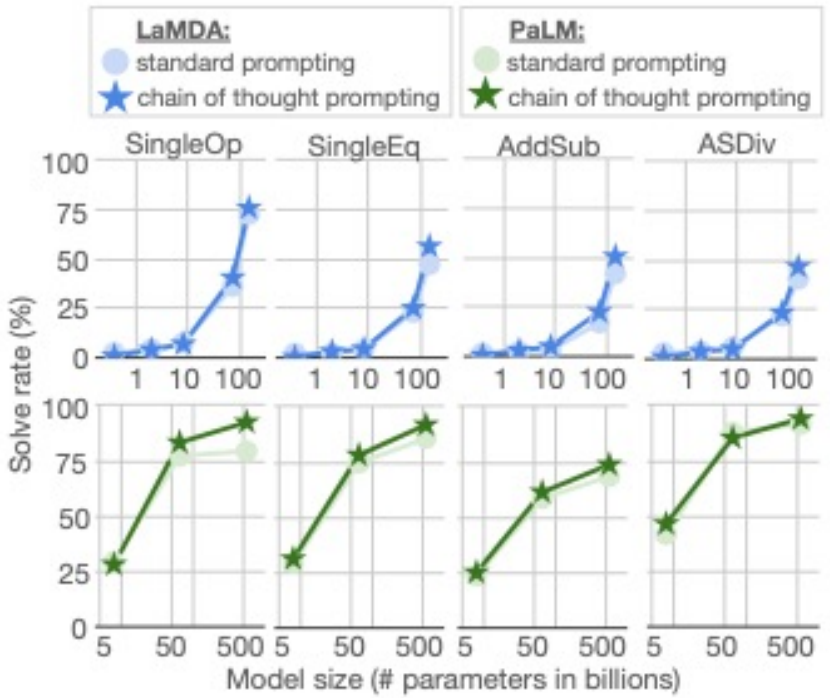


Figure 2. When scaling up the model already facilitates good performance, chain of thought prompting does as well or better.

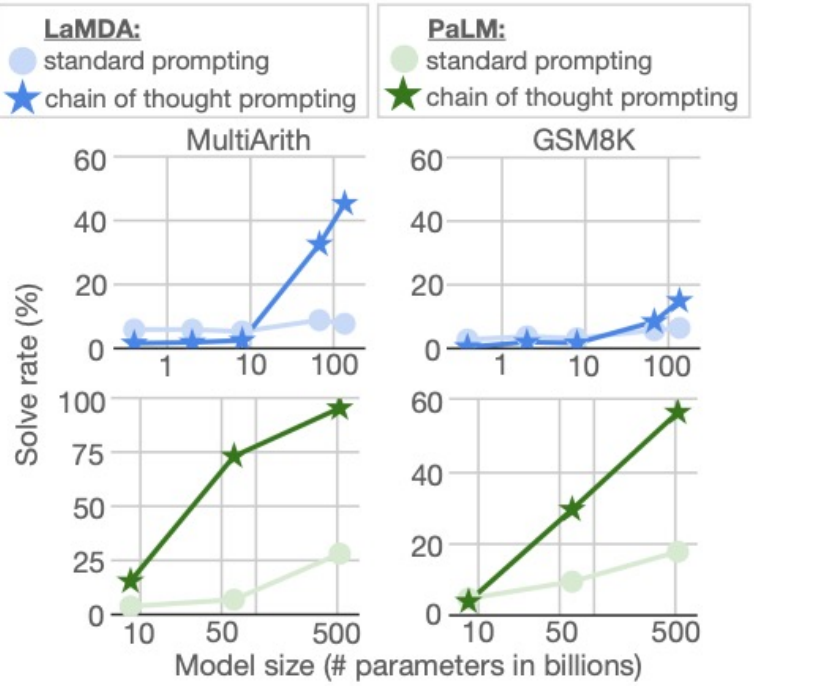


Figure 3. Employing chain of thought enables language models to solve challenging math word problems for which standard prompting has a mostly flat scaling curve.

Fast-system

Slow-system

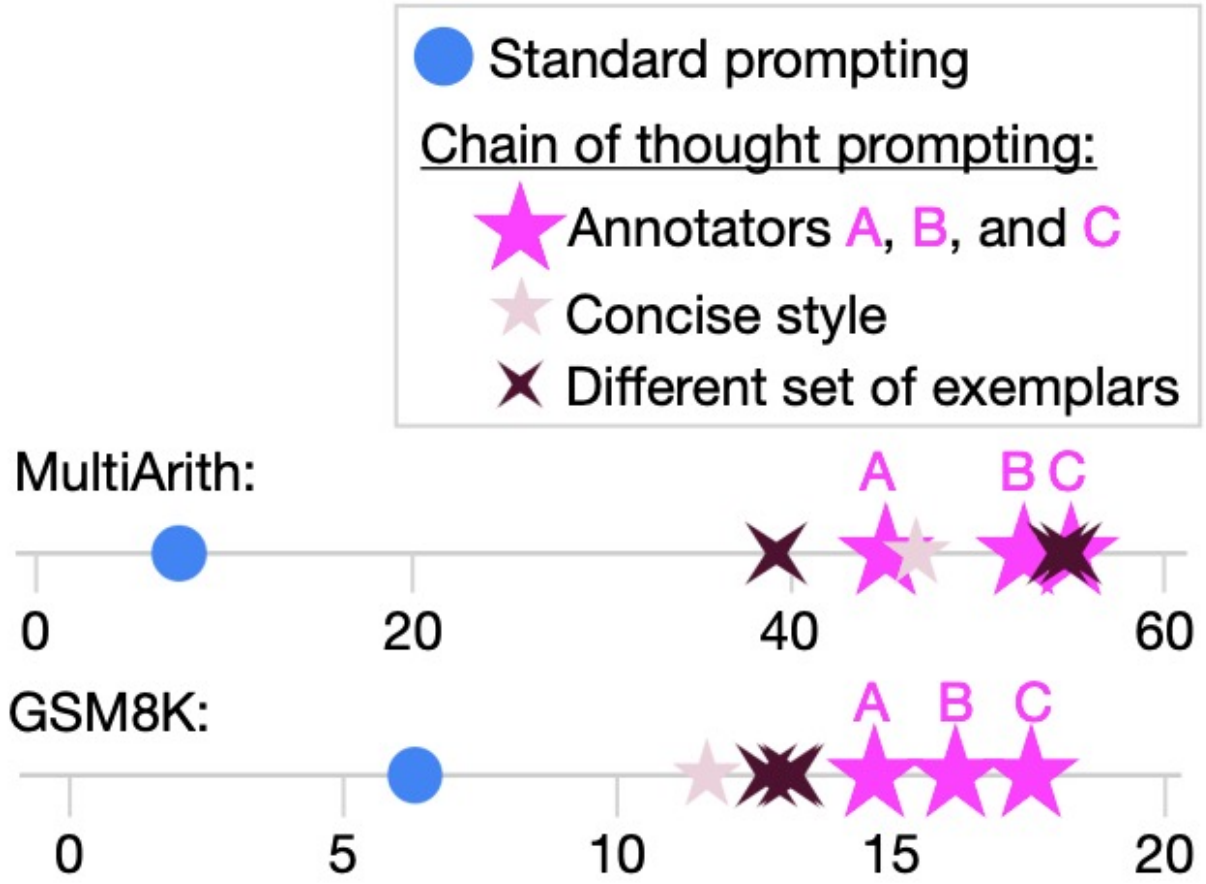
Ablation study

Table 3. Ablation study for different variants of chain of thought, using LaMDA 137B and PaLM 540B.

	GSM8K acc. (%)	
	LaMDA	PaLM
Standard prompting	6.3	17.9
Chain of thought prompting	14.8	56.5
Ablations:		
Equation only	5.7	21.7
Variable compute only	6.0	17.7
Thought after answer	5.9	18.0

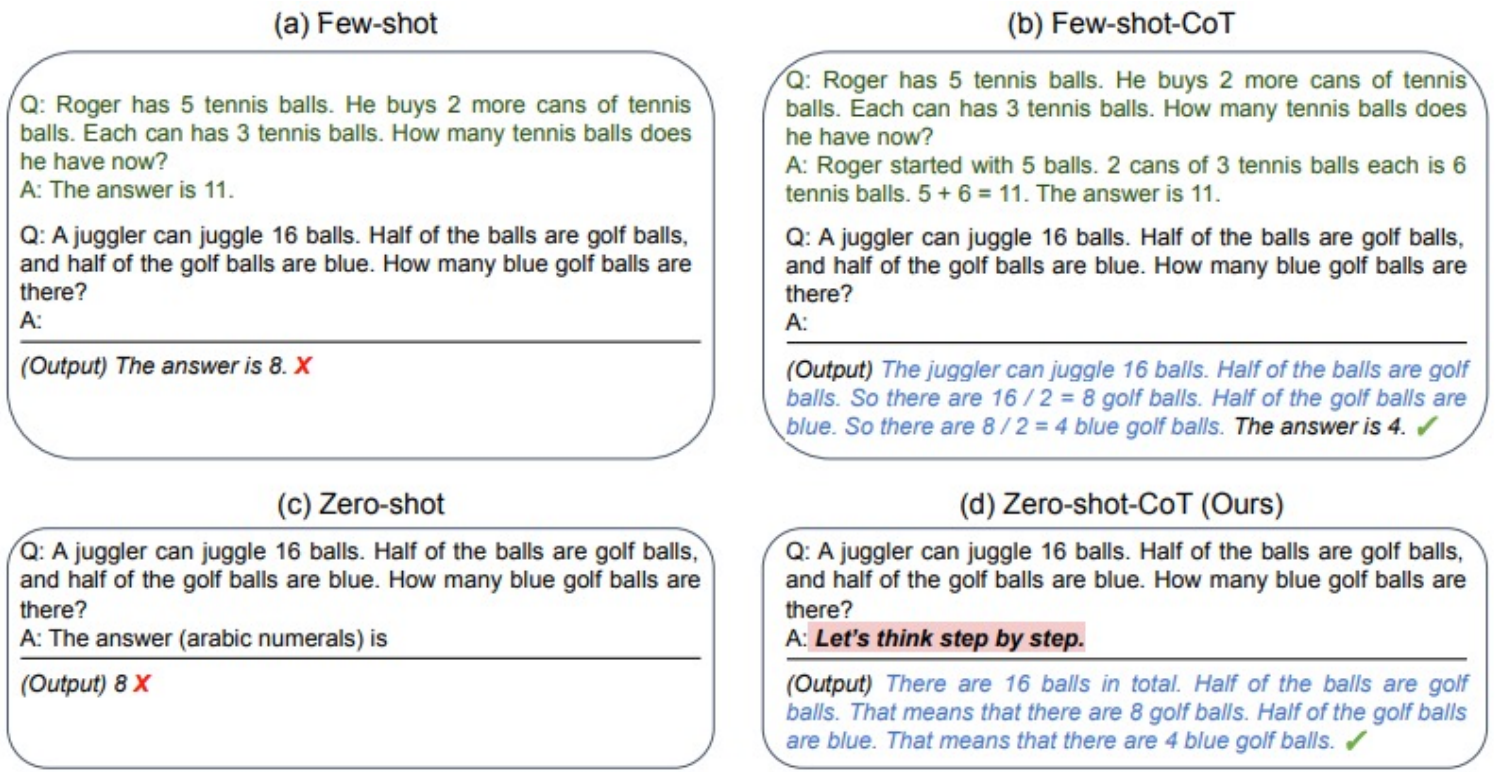
Equations, **sentence description**, and **order** are all important to CoT

Robustness of CoT



Zero-shot CoT: Let's think step by step

Constructing CoT is painful. Can we have Zero-shot CoT? Yes!



[Large Language Models are Zero-Shot Reasoners, NeurIPS 2022]

Zero-shot CoT: Let's think step by step

北京大学 PEKING UNIVERSITY

	Arithmetic					
	SingleEq	AddSub	MultiArith	GSM8K	AQUA	SVAMP
zero-shot	74.6/**78.7**	**72.2/77.0**	17.7/22.7	10.4/12.5	22.4/22.4	58.8/58.7
zero-shot-cot	**78.0/78.7**	69.6/74.7	**78.7/79.3**	**40.7/40.5**	**33.5/31.9**	**62.1/63.7**

	Common Sense		Other Reasoning Tasks		Symbolic Reasoning	
	Common SenseQA	Strategy QA	Date Understand	Shuffled Objects	Last Letter (4 words)	Coin Flip (4 times)
zero-shot	**68.8/72.6**	12.7/**54.3**	49.3/33.6	31.3/29.7	0.2/-	12.8/53.8
zero-shot-cot	64.6/64.0	**54.8**/52.3	**67.5/61.8**	**52.4/52.9**	**57.6**/-	**91.4/87.8**

Zero-shot CoT: Let's think step by step

北京大学 PEKING UNIVERSITY

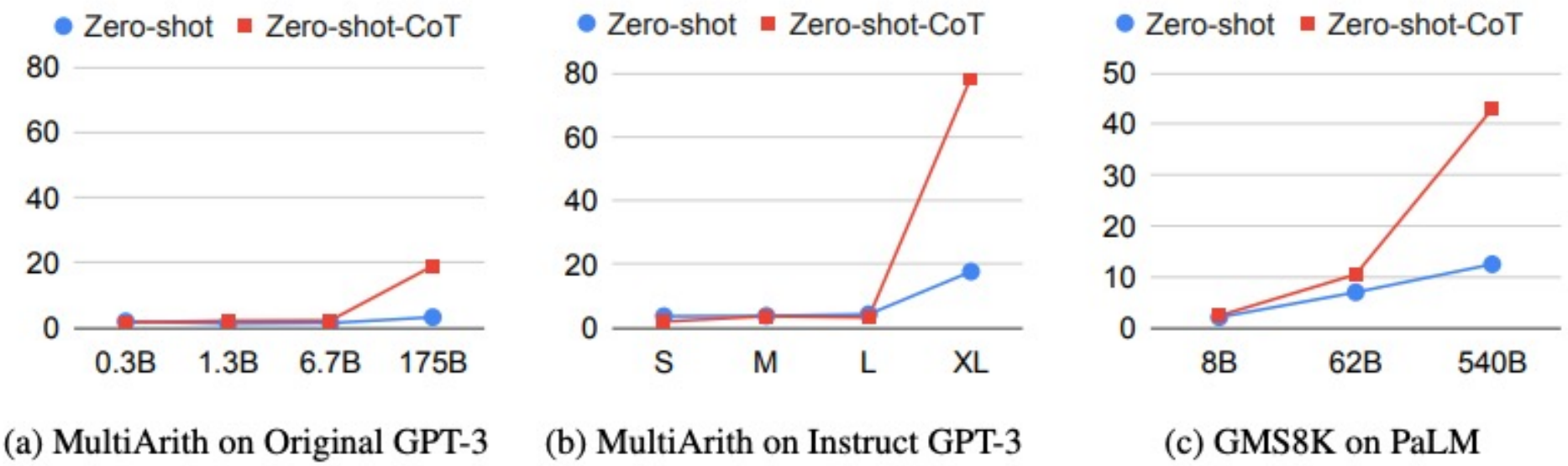
	MultiArith	GSM8K
Zero-Shot	**17.7**	**10.4**
Few-Shot (2 samples)	33.7	15.6
Few-Shot (8 samples)	33.8	15.6
Zero-Shot-CoT	**78.7**	**40.7**
Few-Shot-CoT (2 samples)	84.8	41.3
Few-Shot-CoT (4 samples : First) (*1)	89.2	-
Few-Shot-CoT (4 samples : Second) (*1)	90.5	-
Few-Shot-CoT (8 samples)	93.0	48.7
Zero-Plus-Few-Shot-CoT (8 samples) (*2)	**92.8**	**51.5**

Zero-shot CoT: Let's think step by step

北京大学 PEKING UNIVERSITY

From 17.7% to 78.7%

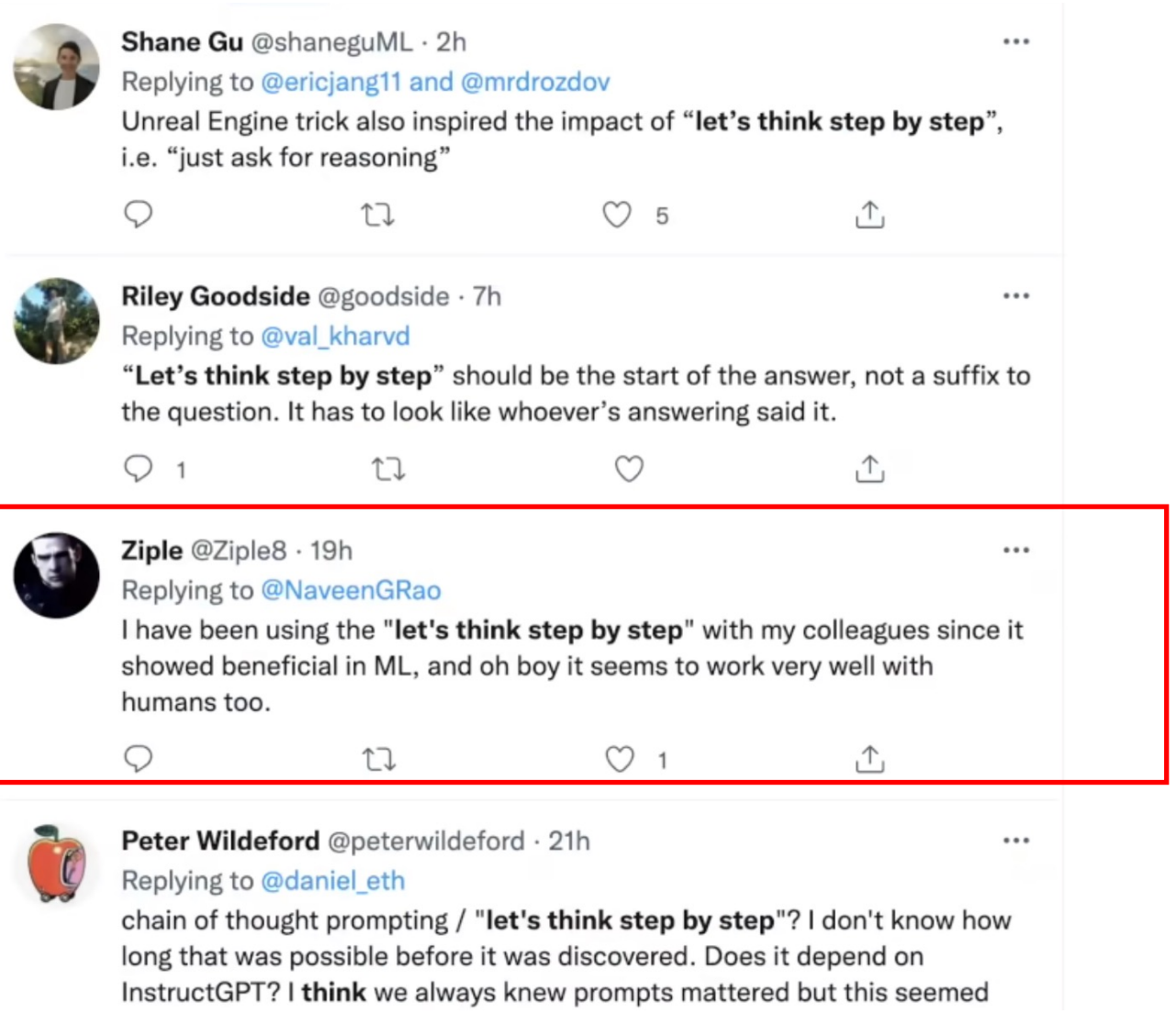
From 10.4% to 40.7%



(a) MultiArith on Original GPT-3 (b) MultiArith on Instruct GPT-3 (c) GMS8K on PaLM

Figure 3: Model scale study with various types of models. S: text-ada-001, M: text-babbage-001, L: text-curie-001, XL: text-davinci-002. See Appendix A.3 and E for the detail.

Zero-shot CoT: Let's think step by step



Shane Gu @shaneguML · 2h
Replying to @ericjang11 and @mrdrozdov
Unreal Engine trick also inspired the impact of "**let's think step by step**", i.e. "just ask for reasoning"
5

Riley Goodside @goodside · 7h
Replying to @val_kharvd
"**Let's think step by step**" should be the start of the answer, not a suffix to the question. It has to look like whoever's answering said it.
1

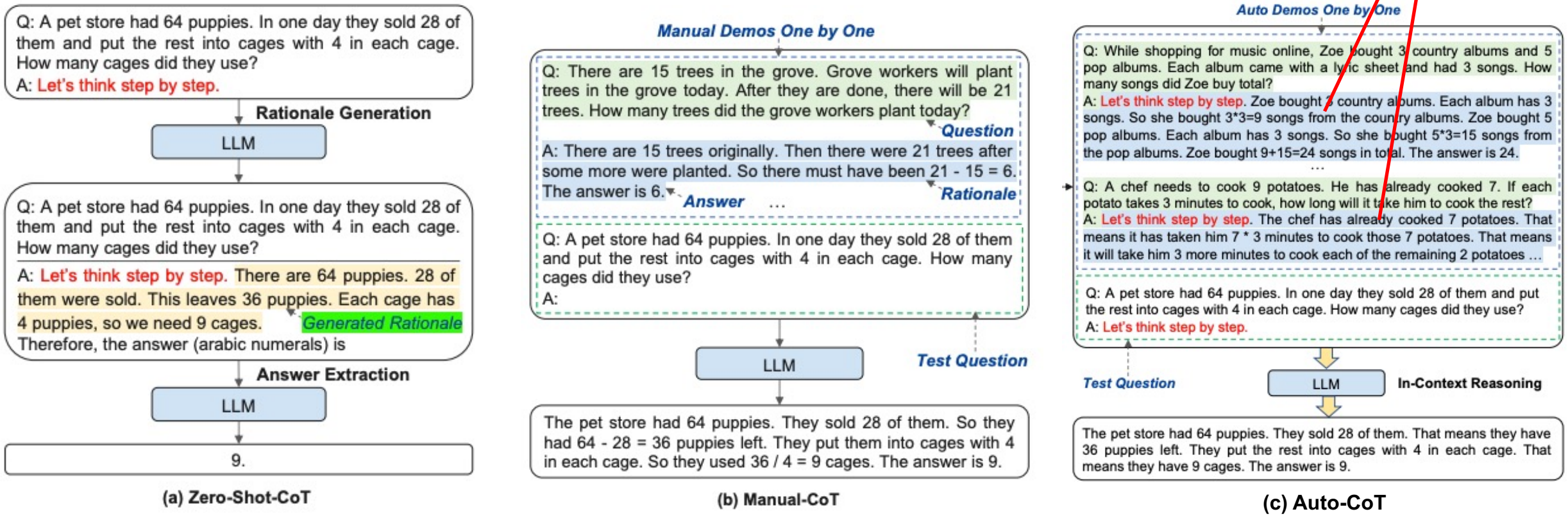
Ziple @Ziple8 · 19h
Replying to @NaveenGRao
I have been using the "**let's think step by step**" with my colleagues since it showed beneficial in ML, and oh boy it seems to work very well with humans too.
1

Peter Wildeford @peterwildeford · 21h
Replying to @daniel_eth
chain of thought prompting / "**let's think step by step**"? I don't know how long that was possible before it was discovered. Does it depend on InstructGPT? I **think** we always knew prompts mattered but this seemed

Auto CoT

北京大学 PEKING UNIVERSITY

Generated automatically



[AUTOMATIC CHAIN OF THOUGHT PROMPTING IN LARGE LANGUAGE MODELS]

Auto CoT

北京大学 PEKING UNIVERSITY

Table 3: Accuracy on ten datasets from three categories of reasoning tasks.

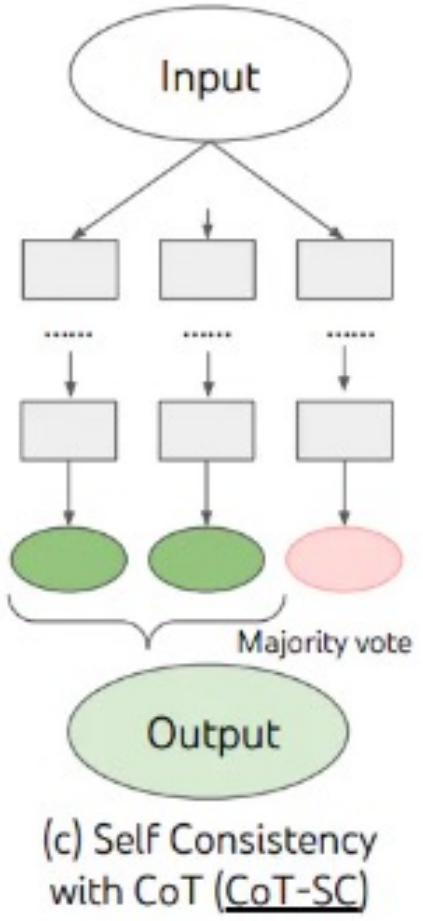
Model	*Arithmetic*						*Commonsense*		*Symbolic*	
	MultiArith	GSM8K	AddSub	AQuA	SingleEq	SVAMP	CSQA	Strategy	Letter	Coin
Zero-Shot	22.7	12.5	77.0	22.4	78.7	58.8	72.6	54.3	0.2	53.8
Zero-Shot-CoT	78.7	40.7	74.7	33.5	78.7	63.7	64.6	54.8	57.6	91.4
Few-Shot	33.8	15.6	83.3	24.8	82.7	65.7	**79.5**	**65.9**	0.2	57.2
Manual-CoT	91.7	46.9	81.3	35.8	86.6	68.9	73.5	65.4	59.0	97.2
Auto-CoT	**92.0**	**47.9**	**84.8**	**36.5**	**87.0**	**69.5**	74.4	65.4	**59.7**	**99.9**

CoT with self-consistency

A complex reasoning problem typically admits **multiple** different ways of thinking leading to its **unique** correct answer

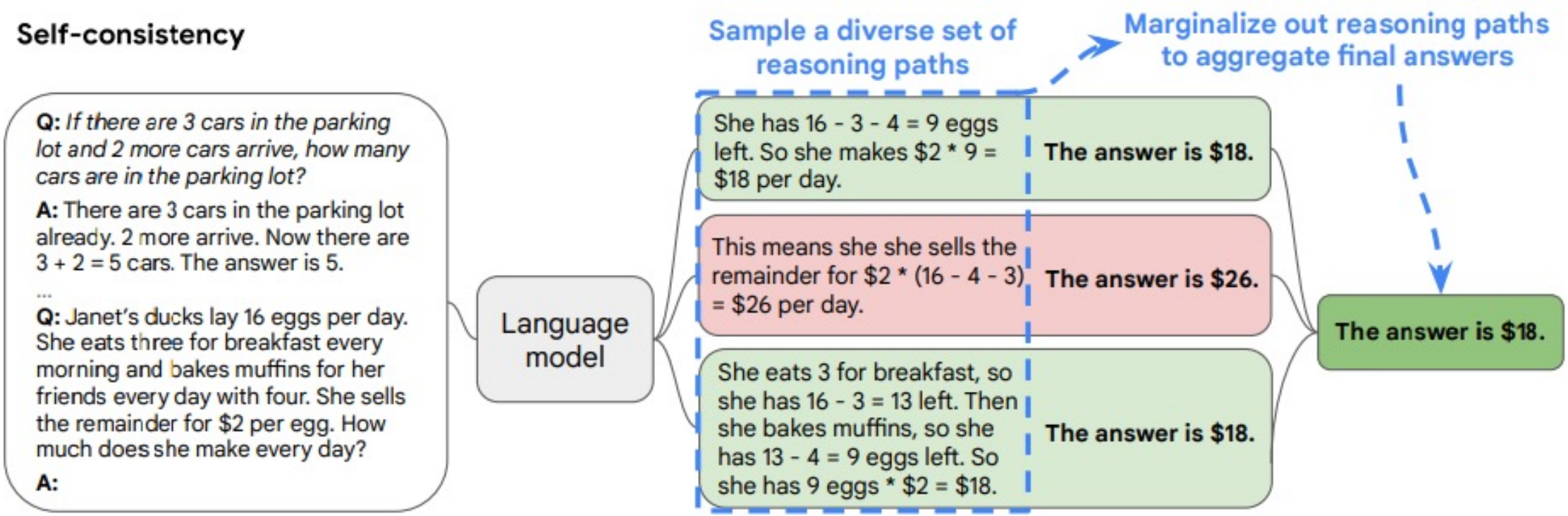
CoT with SC first samples a **diverse** set of reasoning paths instead of only taking the greedy one, and then selects **the most consistent** answer

[SELF-CONSISTENCY IMPROVES CHAIN OF THOUGHT REASONING IN LANGUAGE MODELS, ICLR 2023]



(c) Self Consistency with CoT (CoT-SC)

CoT with self-consistency



CoT with self-consistency

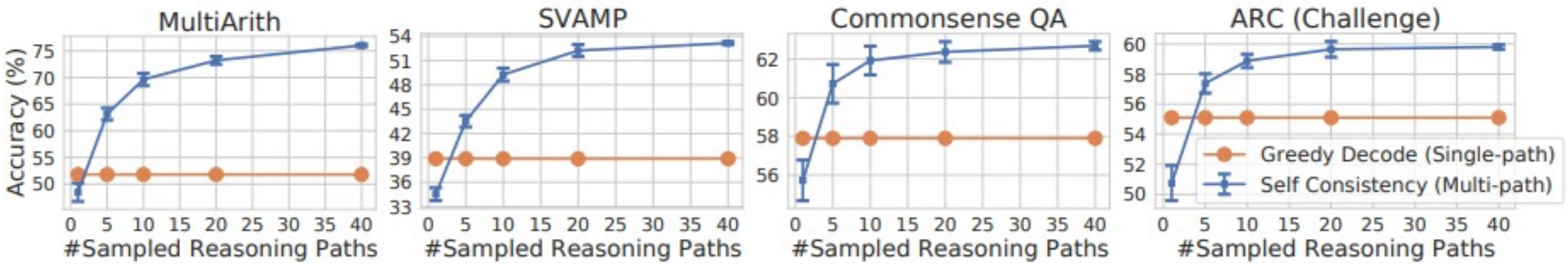
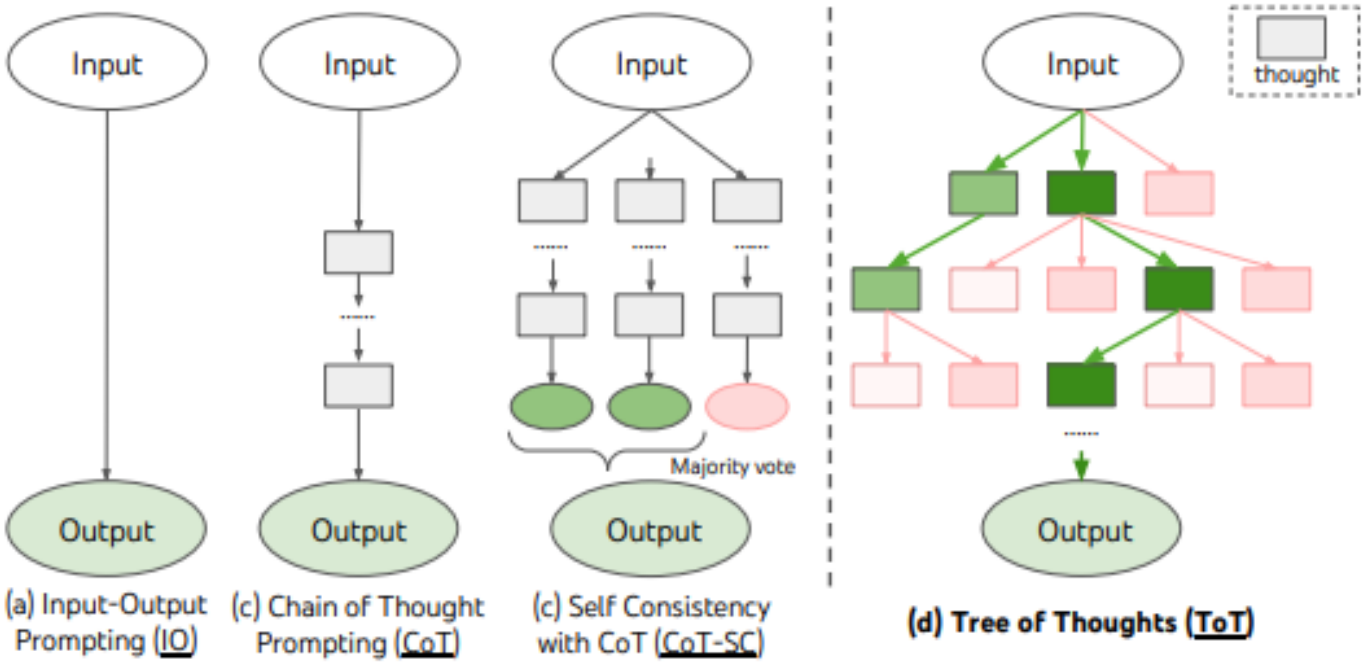


Figure 2: Self-consistency (blue) significantly improves accuracy over CoT-prompting with greedy decoding (orange) across arithmetic and commonsense reasoning tasks, over LaMDA-137B. Sampling a higher number of diverse reasoning paths consistently improves reasoning accuracy.

Tree of Thoughts (ToT)

A genuine problem-solving process involves the repeated use of available information to initiate exploration, which discloses, in turn, more information until a way to attain the solution is finally discovered.—— Newell et al. [21]



[Tree of Thoughts: Deliberate Problem Solving with Large Language Models, NeurIPS 2023]

Tree of Thoughts (ToT)

chain of thoughts 和 tree of thoughts的区别在哪里?

"Chain of Thoughts"（思考链）和"Tree of Thoughts"（思考树）是两种不同的思维模式或方法，它们在解决问题、做决策或进行推理时有所区别。

1. Chain of Thoughts（思考链）：
   - 思考链通常是指一种线性的、连续的思维过程，其中每个想法或结论都是基于前一个想法或结论。
   - 这种模式强调逻辑和连续性，每个步骤都依赖于前一个步骤，形成一条链条。
   - 思考链可能更适用于解决那些需要逐步推理和逻辑推断的问题。
2. Tree of Thoughts（思考树）：
   - 思考树则是一种分支式的思维模式，它从一个中心点开始，然后分支出多个可能的想法或方向。
   - 这种模式强调探索和多样性，允许从多个角度或路径来探索问题或解决方案。
   - 思考树适用于那些需要创造性思维、探索多种可能性或需要考虑多个因素的问题。

总的来说，思考链更侧重于逻辑和连续性，而思考树则更侧重于探索和多样性。在实际应用中，根据问题的性质和需求，可以选择合适的思维模式来解决问题。

Fail and Trial!

Search, Search, Research!

Tree of Thoughts (ToT)

给我一个实际的创作案例，应用思考链和思考树的具体过程。

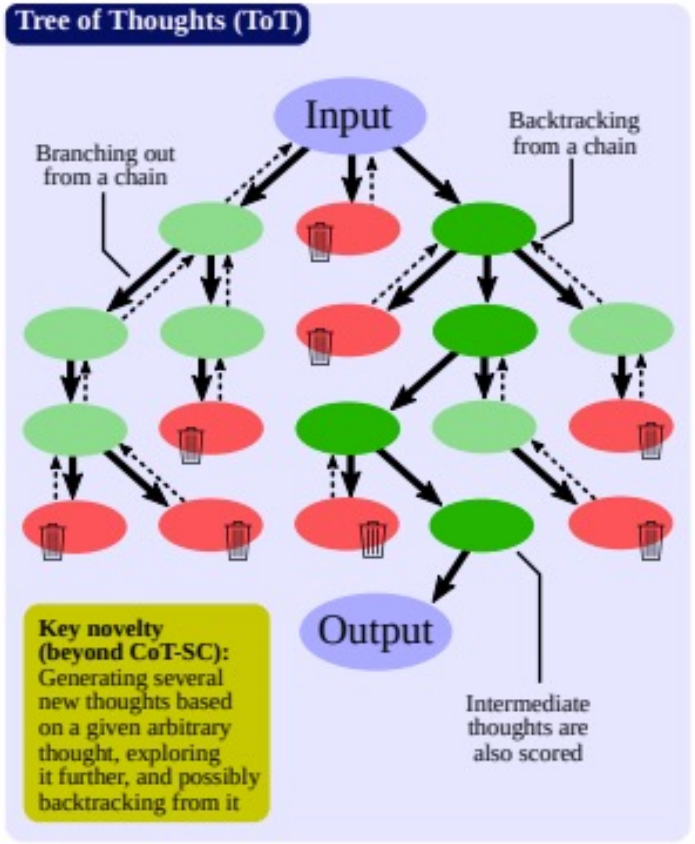
好的，让我们以一个广告创意过程为例，来展示思考链和思考树的应用。

案例：为一个新款运动鞋创建广告创意

1. 思考链（Chain of Thoughts）过程：a. 了解产品：首先，团队需要了解新款运动鞋的特点、目标市场和竞争优势。b. 确定广告目标：根据产品特点和市场定位，确定广告的主要目标，例如提高品牌知名度、强调产品舒适性或突出其高科技特性。c. 创意构思：基于广告目标，团队开始构思创意，这可能包括故事情节、视觉元素和口号。d. 制作和执行：一旦创意确定，团队将进入制作阶段，包括拍摄、编辑和后期制作。e. 发布和评估：最后，广告将被发布到不同的媒体平台上，并对其效果进行评估，以便为未来的广告活动提供依据。
2. 思考树（Tree of Thoughts）过程：a. 了解产品：与思考链相同，首先需要了解产品特点和目标市场。b. 创意发散：团队开始进行头脑风暴，从不同角度和方向生成多个创意。例如，可能探索不同的故事情节、角色、场景和视觉风格。c. 筛选和评估：根据预算、时间和市场定位等因素，对创意进行筛选和评估，选择最符合要求的创意。d. 制作和执行：与思考链相同，进入制作阶段，将选定的创意转化为实际的广告内容。e. 发布和评估：最后，将广告发布到媒体平台上，并对其效果进行评估。

总之，在这个广告创意过程中，思考链侧重于逻辑和连续性，每个步骤都是基于前一个步骤的结果。而思考树则强调探索和多样性，通过头脑风暴和创意发散，从多个方向探索可能的创意，然后根据实际情况进行筛选和评估。

Tree of Thoughts (ToT)



Algorithm 1 ToT-BFS$(x, p_\theta, G, k, V, T, b)$

Require: Input x, LM p_θ, thought generator $G()$ & size limit k, states evaluator $V()$, step limit T, breadth limit b.
$S_0 \leftarrow \{x\}$
for $t = 1, \cdots, T$ **do**
  $S'_t \leftarrow \{[s, z] \mid s \in S_{t-1}, z_t \in \mathrm{G}(p_\theta, s, k)\}$
  $V_t \leftarrow V(p_\theta, S'_t)$
  $S_t \leftarrow \arg\max_{S \subset S'_t, |S|=b} \sum_{s \in S} V_t(s)$
end for
return $G(p_\theta, \arg\max_{s \in S_T} V_T(s), 1)$

Breadth First Search

Algorithm 2 ToT-DFS$(s, t, p_\theta, G, k, V, T, v_{th})$

Require: Current state s, step t, LM p_θ, thought generator $G()$ and size limit k, states evaluator $V()$, step limit T, threshold v_{th}
if $t > T$ **then** record output $G(p_\theta, s, 1)$
end if
for $s' \in G(p_\theta, s, k)$ **do** ▷ sorted candidates
  if $V(p_\theta, \{s'\})(s) > v_{thres}$ **then** ▷ pruning
    DFS$(s', t+1)$
  end if
end for

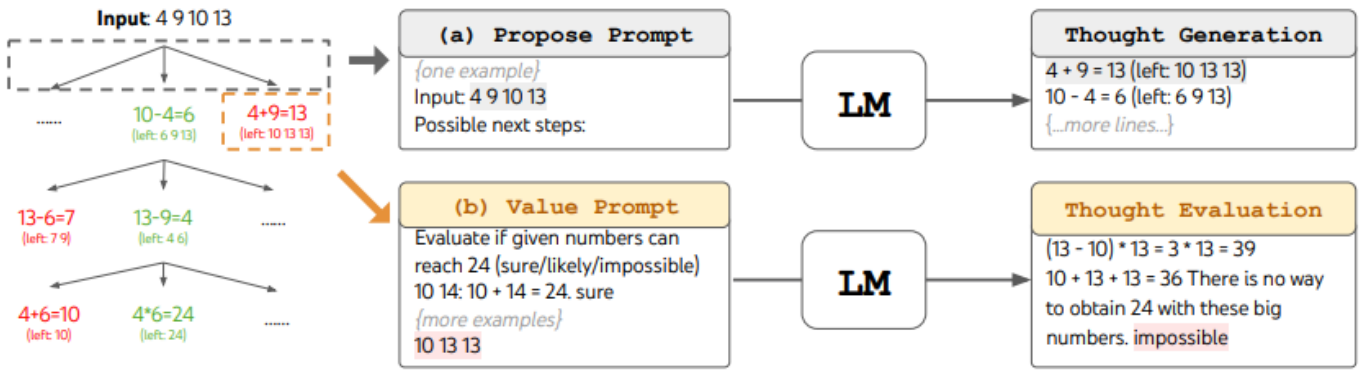
Depth First Search

Tree of Thoughts (ToT)

北京大学 PEKING UNIVERSITY

4.1 Game of 24

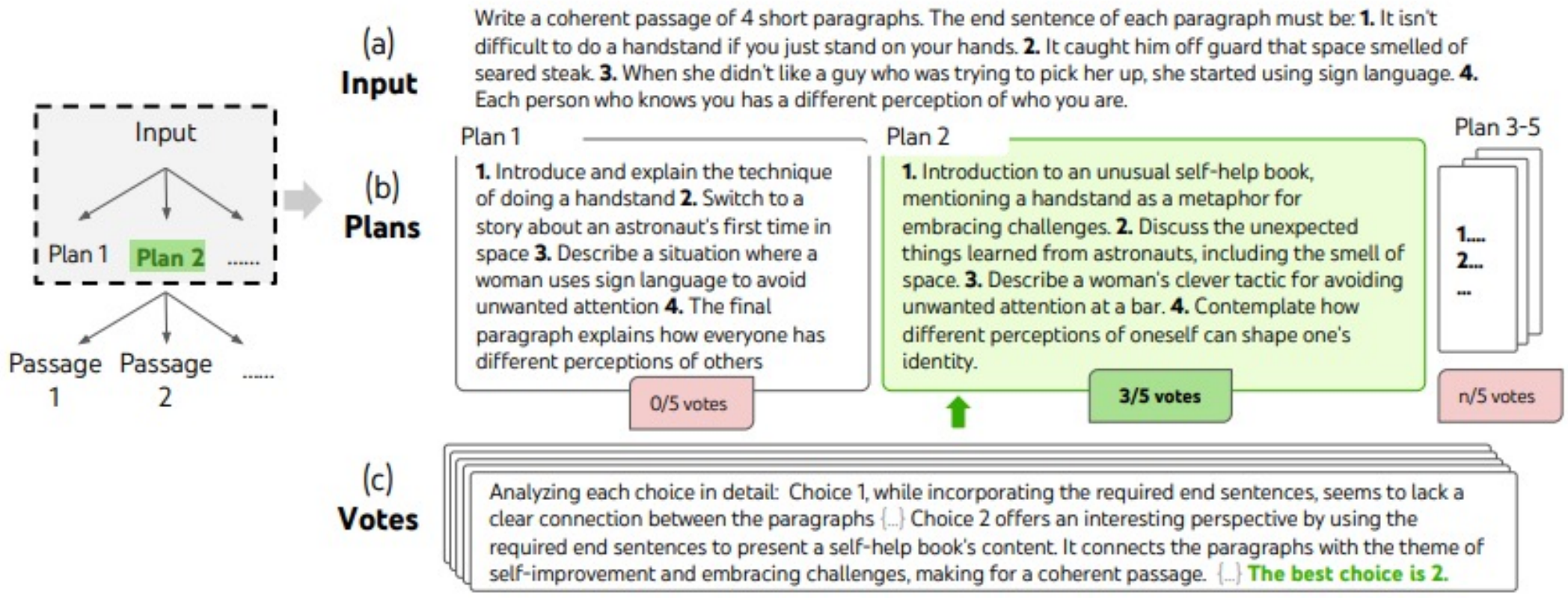
Game of 24 is a mathematical reasoning challenge, where the goal is to use 4 numbers and basic arithmetic operations (+-*/) to obtain 24. For example, given input "4 9 10 13", a solution output could be "(10 - 4) * (13 - 9) = 24".



Method	Success
IO prompt	7.3%
CoT prompt	4.0%
CoT-SC (k=100)	9.0%
ToT (ours) (b=1)	45%
ToT (ours) (b=5)	**74%**
IO + Refine (k=10)	27%
IO (best of 100)	33%
CoT (best of 100)	49%

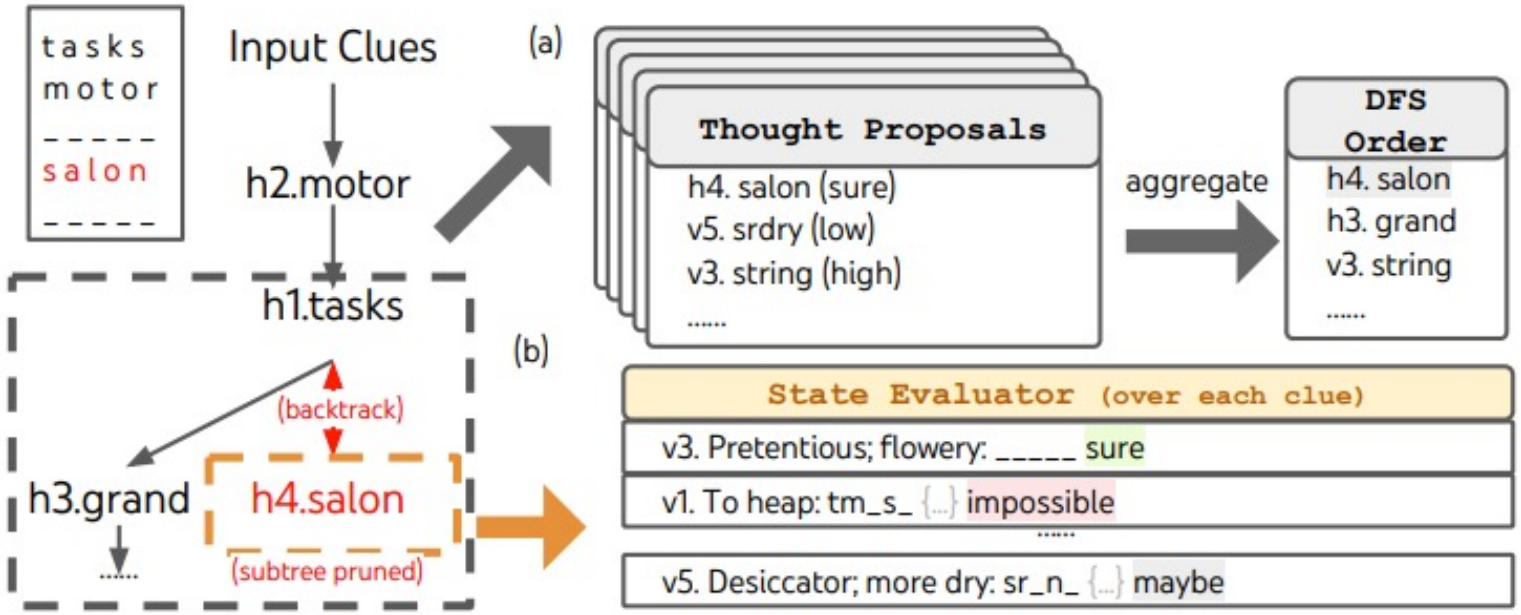
Table 2: Game of 24 Results.

Tree of Thoughts (ToT)



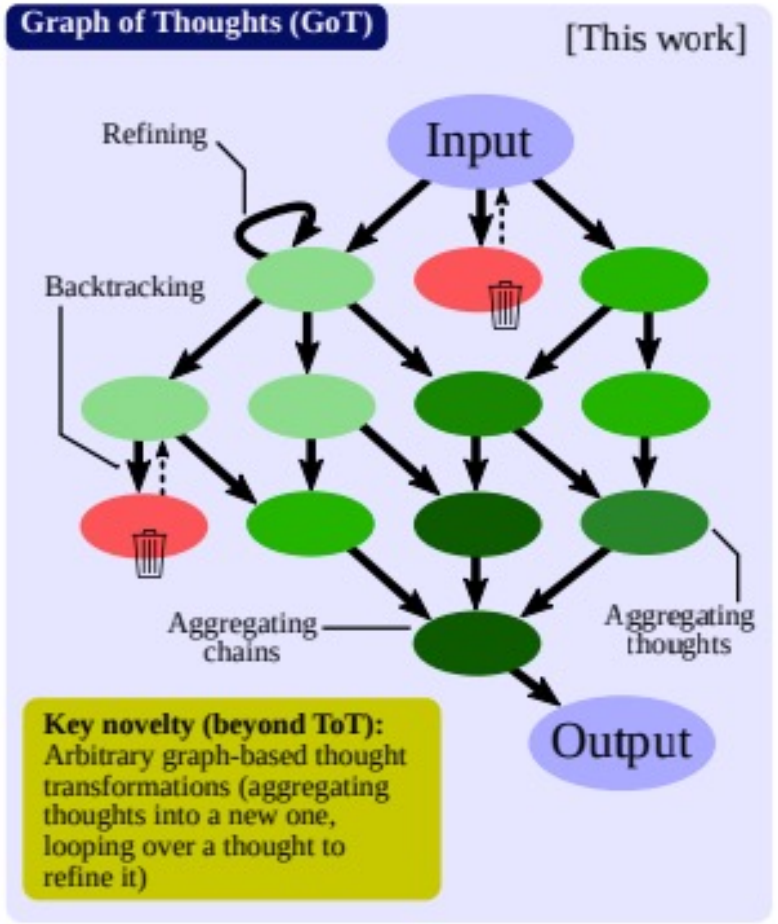
Creative writing

Tree of Thoughts (ToT)

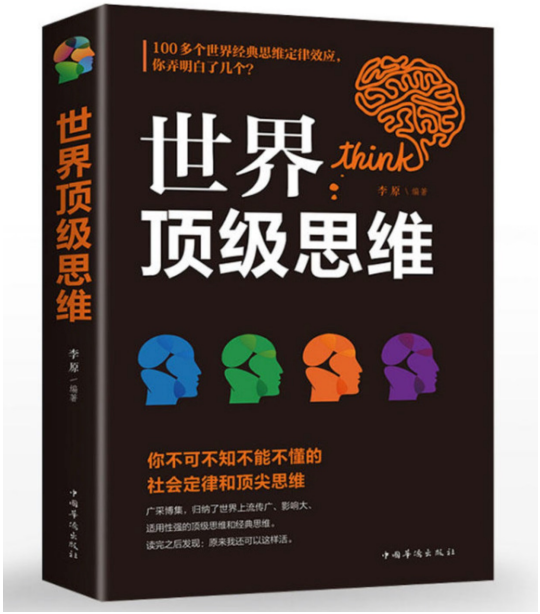


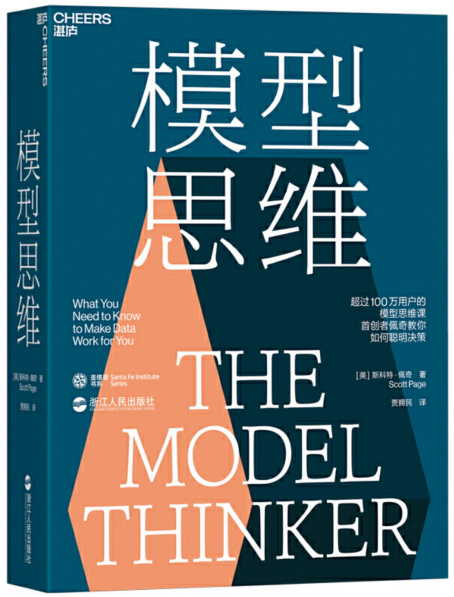
Mini-Crosswords

Graph of Thoughts

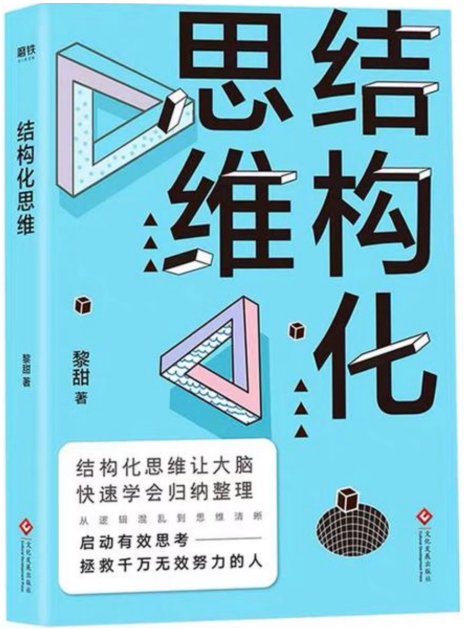


What's next?


100多个世界经典思维定律效应，
你弄明白了几个？
世界顶级思维
think
李原 编著
你不可不知不能不懂的
社会定律和顶尖思维
广采博集，归纳了世界上流传广、影响大、
适用性强的顶级思维和经典思维，
读完之后发现：原来我还可以这样活。
中国华侨出版社


CHEERS
湛庐
模型思维
What You
Need to Know
to Make Data
Work for You
超过100万用户的
模型思维课
首创者佩奇教你
如何聪明决策
THE
MODEL
THINKER
[美] 斯科特·佩奇 著
Scott Page
浙江人民出版社


深度思维
通过复杂直抵本质的
跨越式成长方法论
那些高手从未告诉你的
顶级思维
深研学习策略与思维方法10年精定之作！


结构化思维
思维结构化
黎甜 著
结构化思维让大脑
快速学会归纳整理
从逻辑混乱到思维清晰
启动有效思考
拯救千万无效努力的人