思维链

Chain-of-Thought (CoT)

汇报人：Chengcheng Han

2023年03月09日

目录

CONTENTS

PART 01

简介

PART 02

相关工作

PART 03

实施路线

大型预训练语言模型

Large pre-trained language model

LLM的推理能力并不能随着模型的扩大而快速增加。

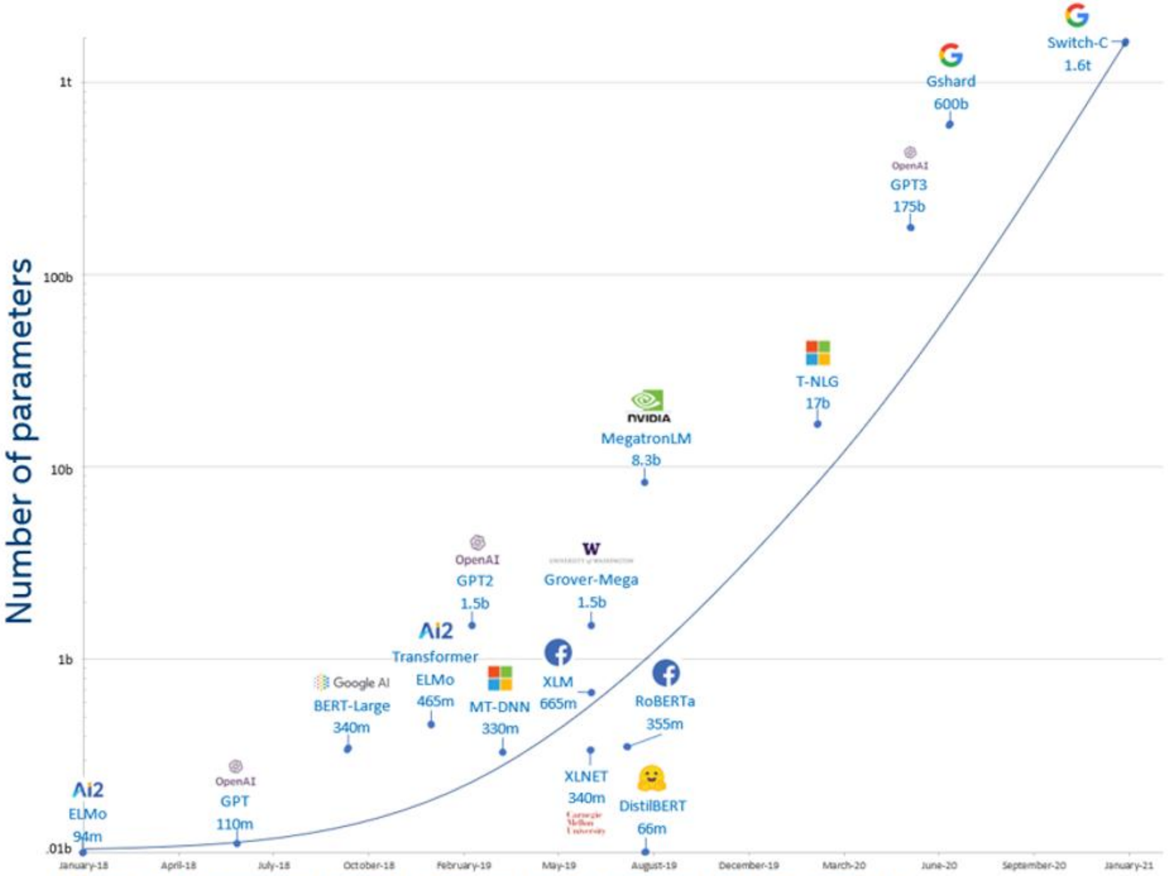
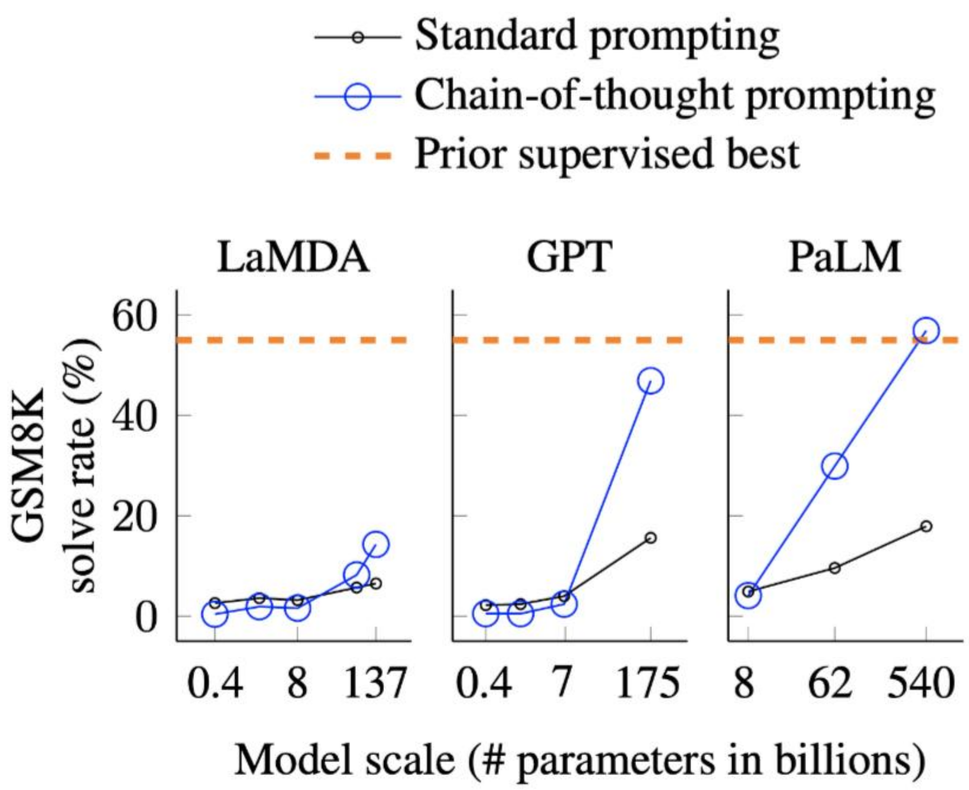


Figure 1: Exponential growth of number of parameters in DL models

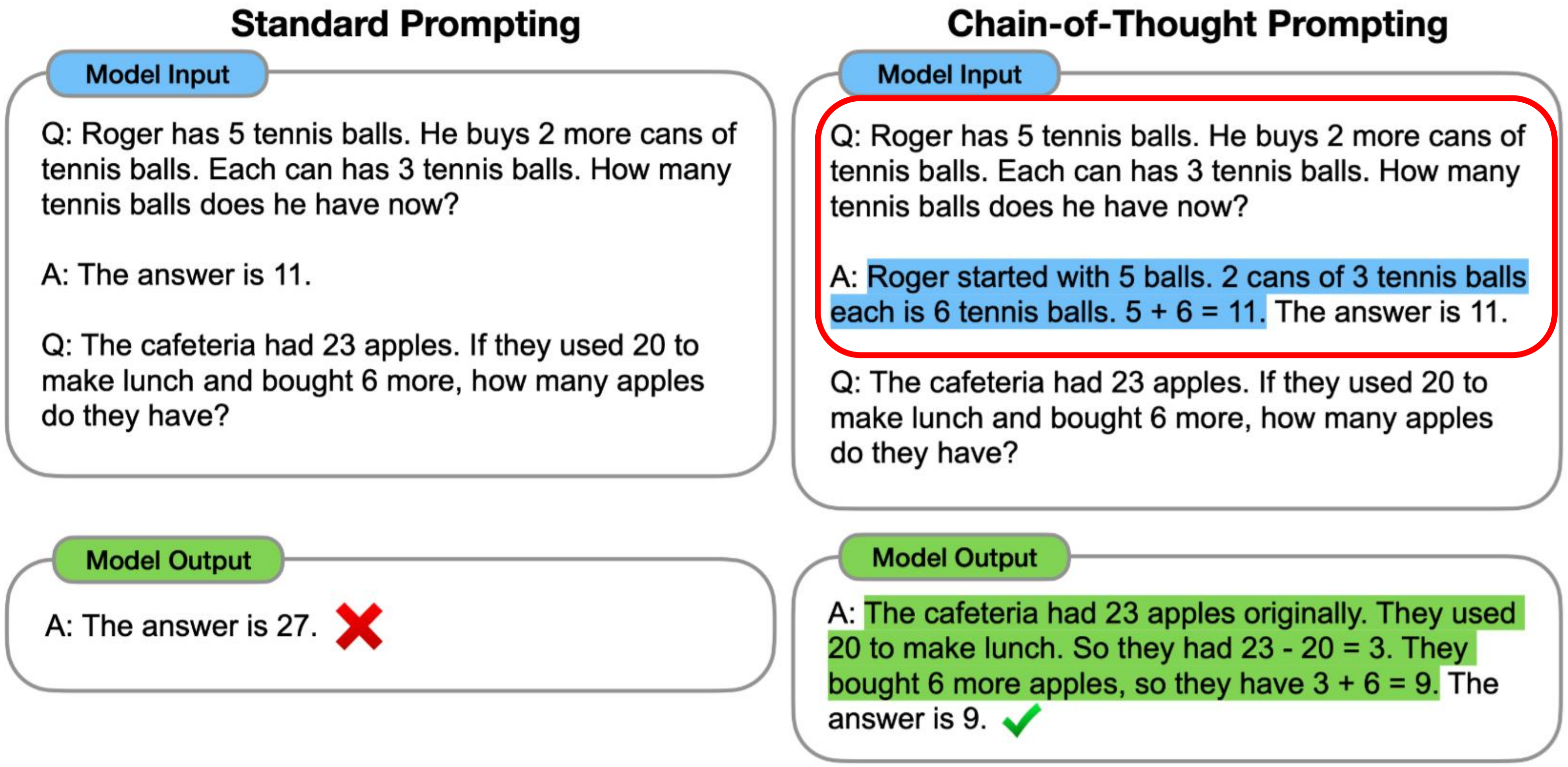


23_ICLR_Chain-of-Thought Prompting Elicits Reasoning in Large Language Models

思维链-定义

CoT-Definition

一系列中间的推理步骤 (a series of intermediate reasoning steps)



Standard Prompting

Model 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: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The answer is 27. ✗

Chain-of-Thought Prompting

Model 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: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

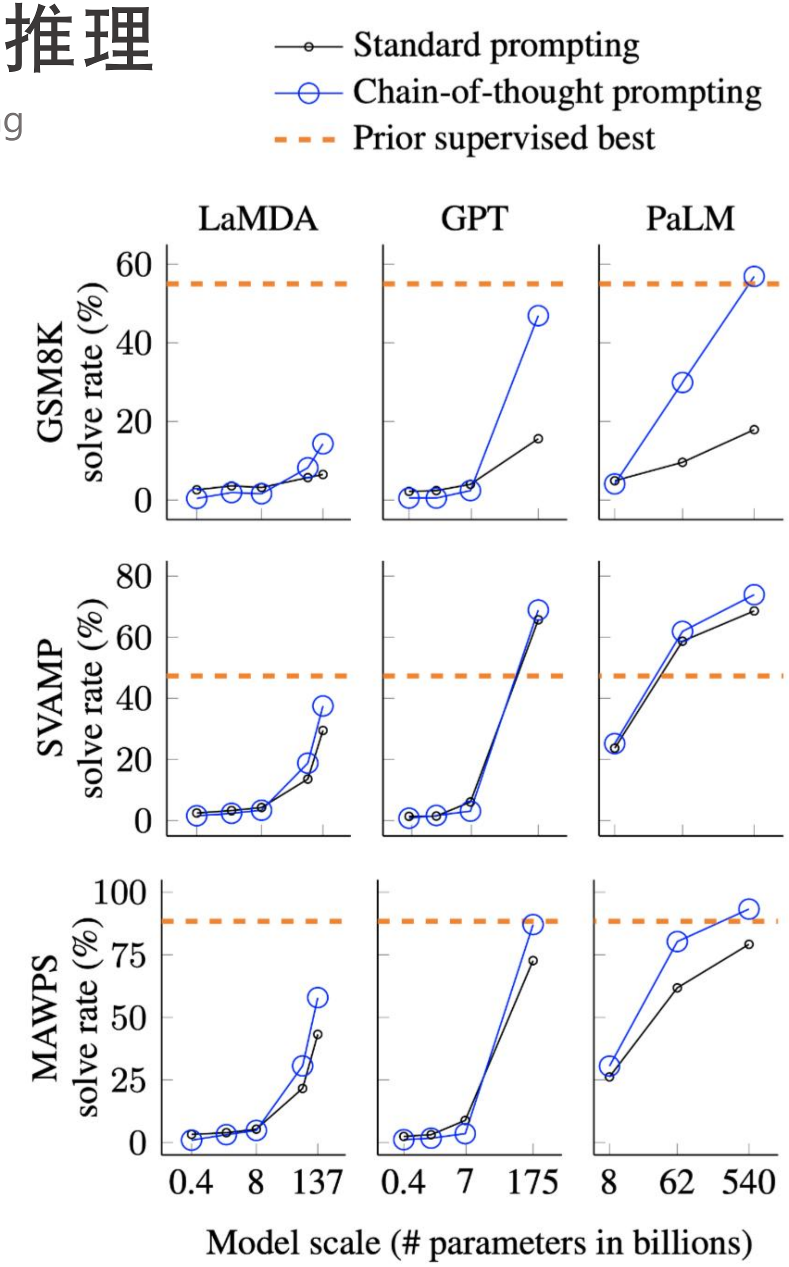
A: The cafeteria had 23 apples originally. They used 20 to make lunch. So they had 23 - 20 = 3. They bought 6 more apples, so they have 3 + 6 = 9. The answer is 9. ✓

Exemplar

23_ICLR_Chain-of-Thought Prompting Elicits Reasoning in Large Language Models

思维链-算术推理

CoT-Arithmetic Reasoning



三个重要结论：

1. CoT对小模型作用不大，模型参数至少达到10B才有效果，达到100B效果才明显。

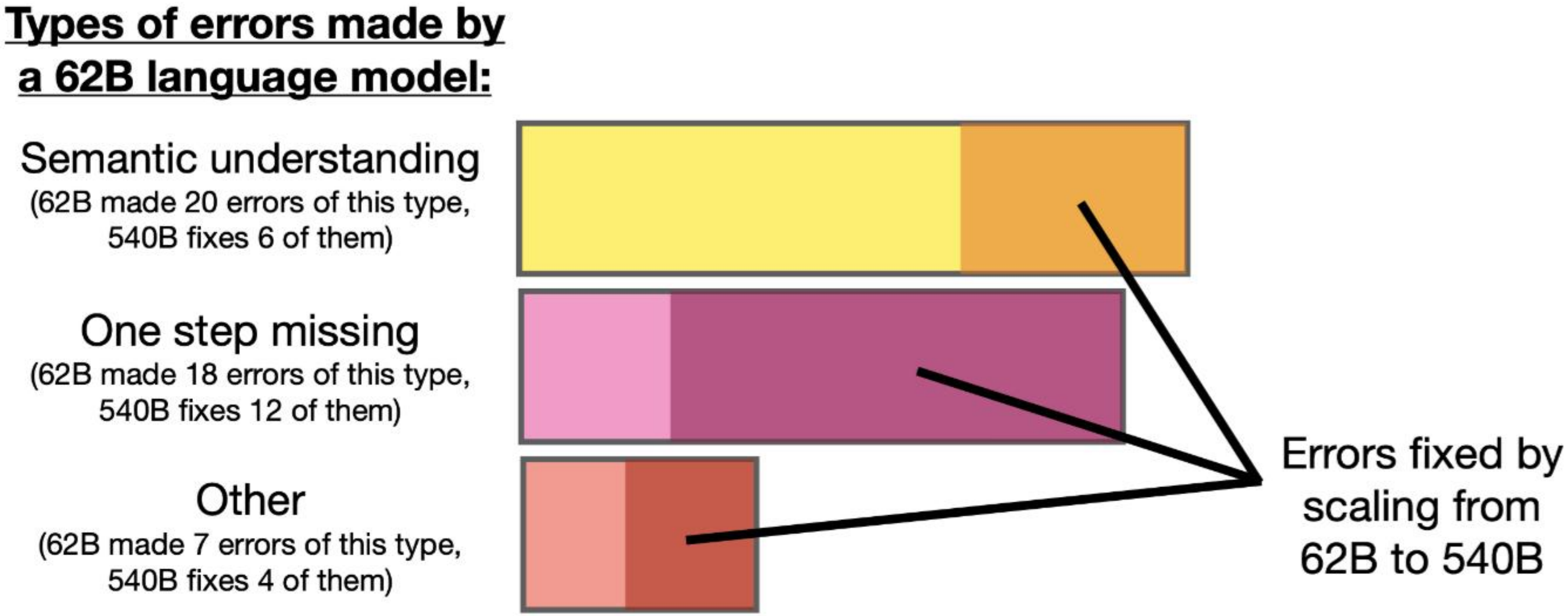
2. CoT对复杂的推理问题的性能增益更大。（为什么呢？猜测是CoT的生成还是不够准确，会存在噪声，导致最终结果变差？）

3. 加上CoT的LLM可以超过专门为特定任务训练的模型的最优结果。

23_ICLR_Chain-of-Thought Prompting Elicits Reasoning in Large Language Models

思维链-分析

CoT-Analysis



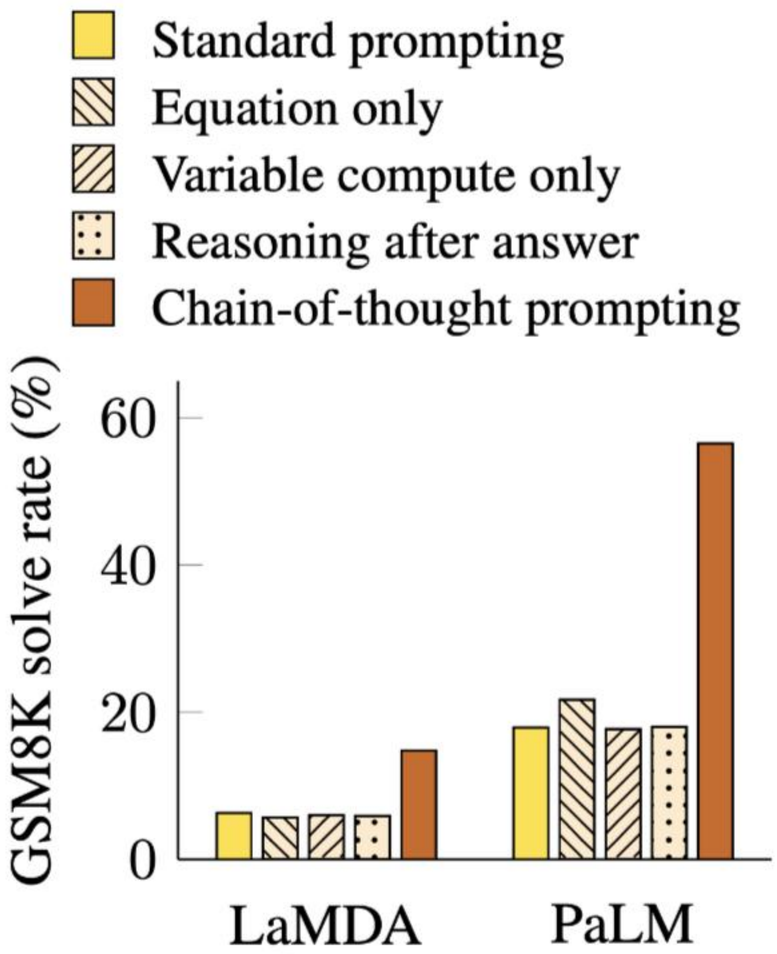
通过增加PaLM的参数量（62B->540B），模型可以修复很大一部分错误。

为什么仅仅增大参数量，就会**涌现**出更强的推理能力？（一个开放的问题）

未来的研究可以集中在那些增大参数量也无法自动修复的问题。

思维链-消融实验

CoT-Ablation Study



三个CoT变种：

1.把CoT中的文字去掉，只保留公式部分。

2. 把CoT中的token全换成点（…）。

3. 把思维链放到生成结果之后。

结论：

三个CoT变种和原始CoT的效果相差甚远。

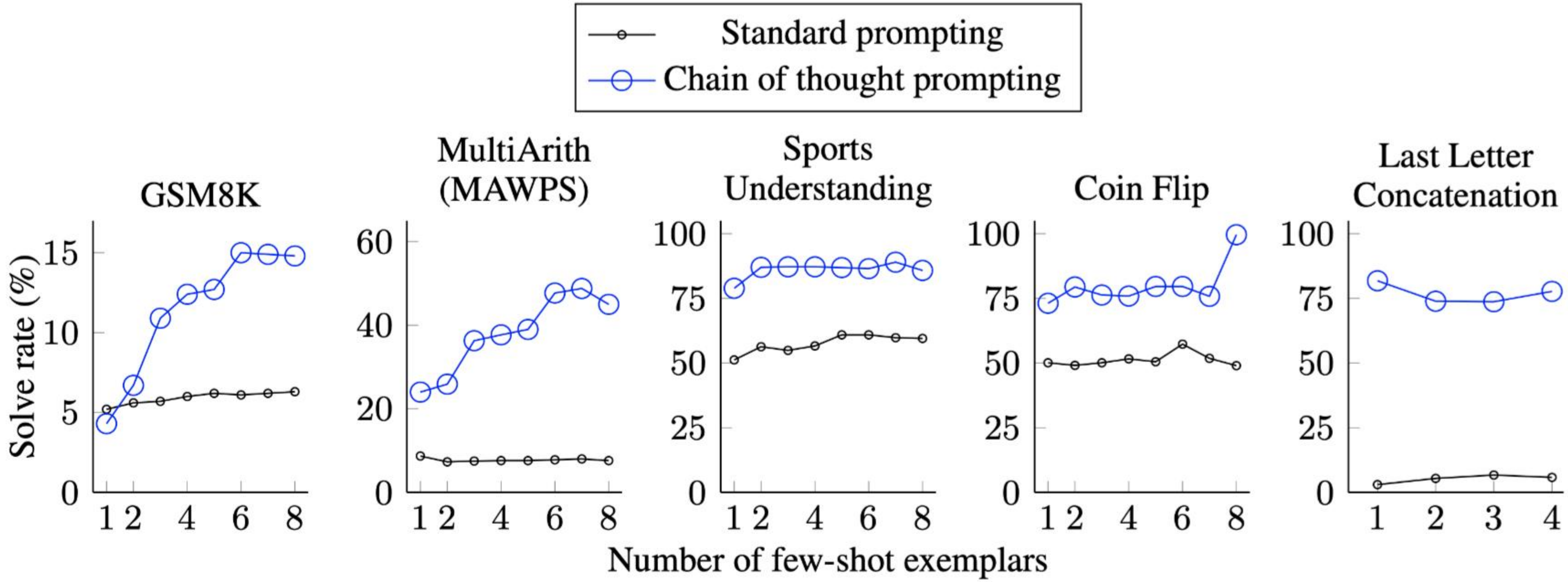
1和2说明CoT中的自然语言部分很重要。

3说明模型确实是依赖于生成的思维链一步一步得到的最终结果。

23_ICLR_Chain-of-Thought Prompting Elicits Reasoning in Large Language Models

思维链-稳健性分析

CoT-Robustness



Prompt Engineering仍然很重要！

不同的prompt（CoT）的设计/数量/顺序都会对模型产生不同的影响，且方差还是很大的。

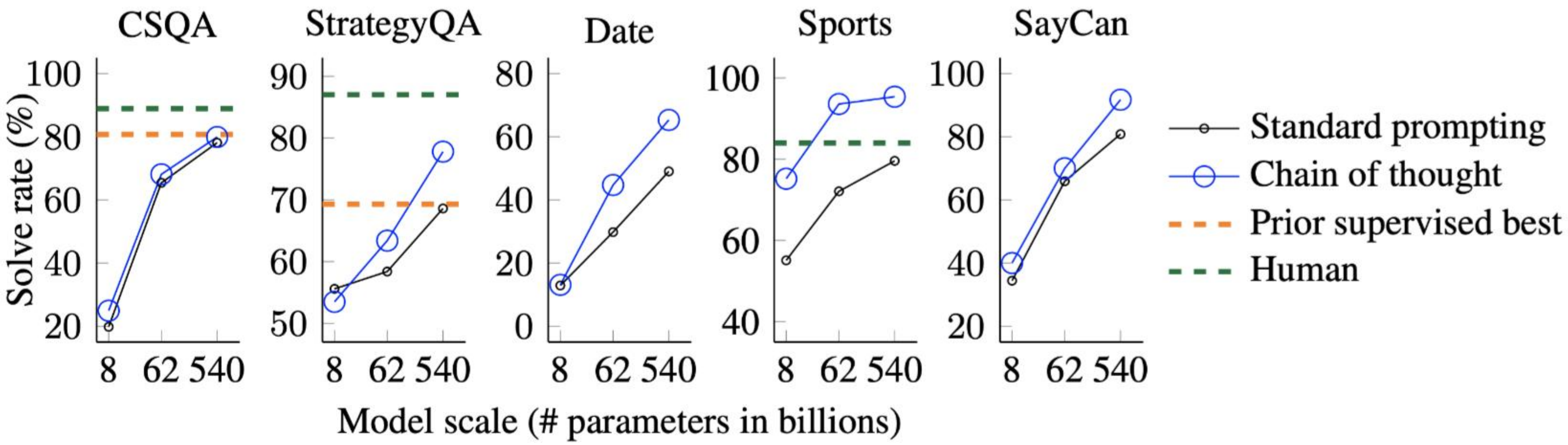
因此未来的一个方向可能是探索一种annotation的模型来得到稳健的CoT（Prompts）。

或许可以用一个LLM自动生成CoT用于Prompting。 —> Auto CoT。

23_ICLR_Chain-of-Thought Prompting Elicits Reasoning in Large Language Models

思维链-常识推理

CoT-Commonsense Reasoning



结论：说明CoT在常识推理任务上也有用。

但是为什么在CSQA上没用？

针对哪种类型的常识性推理任务有用？为什么会失效或奏效都是可以继续研究的问题。

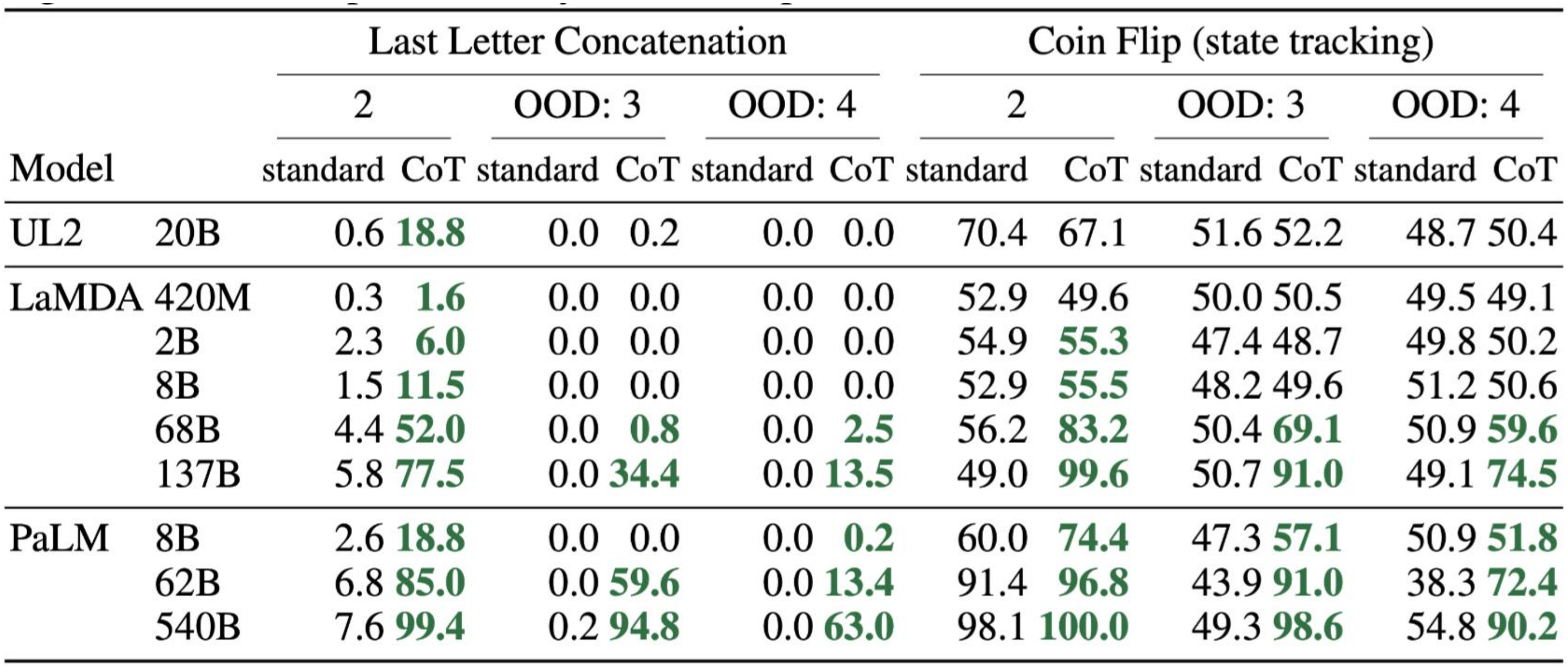
23_ICLR_Chain-of-Thought Prompting Elicits Reasoning in Large Language Models

思维链-符号推理

CoT-Symbolic Reasoning

Last letter concatenation (e.g., “Amy Brown” → “yn”)

Coin flip (e.g., “A coin is heads up. Bob flips the coin. Lily does not flip the coin. Is the coin still heads up?” → “no”)



		Last Letter Concatenation						Coin Flip (state tracking)					
		2		OOD: 3		OOD: 4		2		OOD: 3		OOD: 4	
Model		standard	CoT	standard	CoT	standard	CoT	standard	CoT	standard	CoT	standard	CoT
UL2	20B	0.6	**18.8**	0.0	0.2	0.0	0.0	70.4	67.1	51.6	52.2	48.7	50.4
LaMDA	420M	0.3	**1.6**	0.0	0.0	0.0	0.0	52.9	49.6	50.0	50.5	49.5	49.1
	2B	2.3	**6.0**	0.0	0.0	0.0	0.0	54.9	**55.3**	47.4	48.7	49.8	50.2
	8B	1.5	**11.5**	0.0	0.0	0.0	0.0	52.9	**55.5**	48.2	49.6	51.2	50.6
	68B	4.4	**52.0**	0.0	**0.8**	0.0	**2.5**	56.2	**83.2**	50.4	**69.1**	50.9	**59.6**
	137B	5.8	**77.5**	0.0	**34.4**	0.0	**13.5**	49.0	**99.6**	50.7	**91.0**	49.1	**74.5**
PaLM	8B	2.6	**18.8**	0.0	0.0	0.0	**0.2**	60.0	**74.4**	47.3	**57.1**	50.9	**51.8**
	62B	6.8	**85.0**	0.0	**59.6**	0.0	**13.4**	91.4	**96.8**	43.9	**91.0**	38.3	**72.4**
	540B	7.6	**99.4**	0.2	**94.8**	0.0	**63.0**	98.1	**100.0**	49.3	**98.6**	54.8	**90.2**

结论：

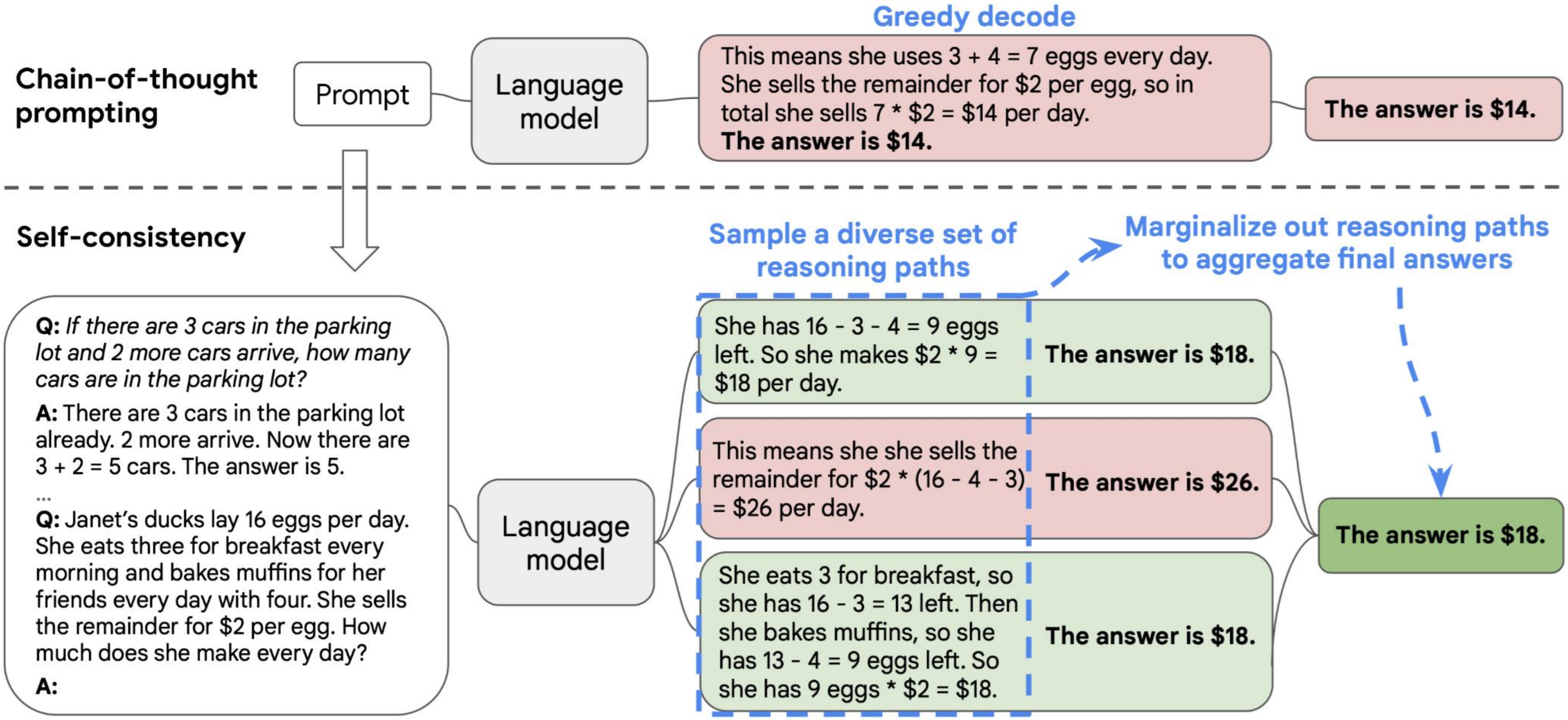
1. 对于LLC任务来说，10B以下的模型，无论in-domain还是OOD，模型都完全不会做（0.0%）。

2. CoT在符号推理任务上效果卓著。

23_ICLR_Chain-of-Thought Prompting Elicits Reasoning in Large Language Models

思维链-一致性方法

CoT-Self-Consistency



Chain-of-thought prompting

Prompt

Language model

Greedy decode

This means she uses 3 + 4 = 7 eggs every day. She sells the remainder for $2 per egg, so in total she sells 7 * $2 = $14 per day. **The answer is $14.**

The answer is $14.

Self-consistency

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 3 cars in the parking lot already. 2 more arrive. Now there are 3 + 2 = 5 cars. The answer is 5.

...

Q: Janet's ducks lay 16 eggs per day. She eats three for breakfast every morning and bakes muffins for her friends every day with four. She sells the remainder for $2 per egg. How much does she make every day?

A:

Language model

Sample a diverse set of reasoning paths

She has 16 - 3 - 4 = 9 eggs left. So she makes $2 * 9 = $18 per day. **The answer is $18.**

This means she she sells the remainder for $2 * (16 - 4 - 3) = $26 per day. **The answer is $26.**

She eats 3 for breakfast, so she has 16 - 3 = 13 left. Then she bakes muffins, so she has 13 - 4 = 9 eggs left. So she has 9 eggs * $2 = $18. **The answer is $18.**

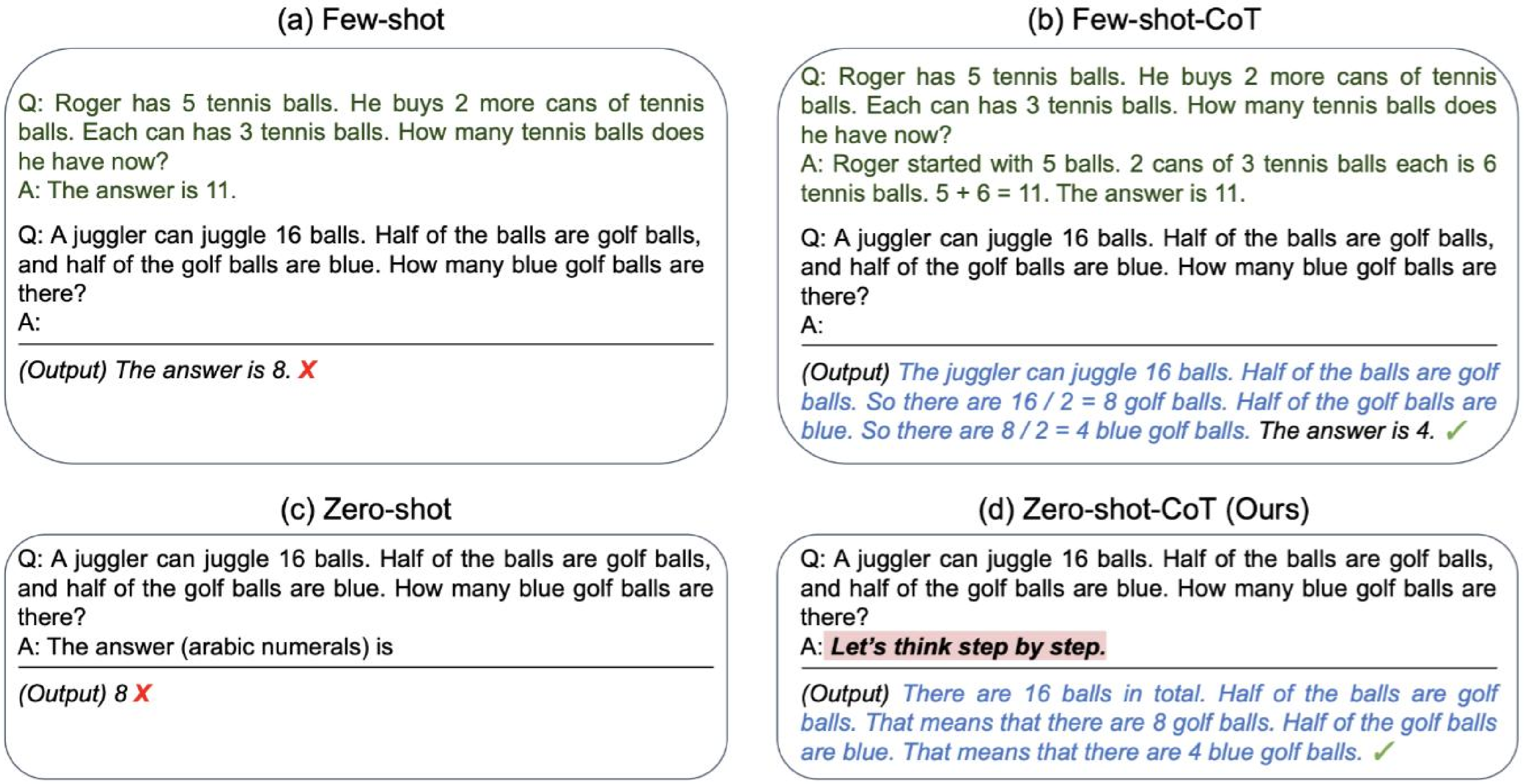
Marginalize out reasoning paths to aggregate final answers

The answer is $18.

23_ICLR_Self-Consistency Improves Chain of Thought Reasoning in Language Models

思维链-零样本设定

CoT Zero-shot



(a) Few-shot

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: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?
A:

(Output) The answer is 8. X

(b) Few-shot-CoT

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: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?
A:

(Output) The juggler can juggle 16 balls. Half of the balls are golf balls. So there are 16 / 2 = 8 golf balls. Half of the golf balls are blue. So there are 8 / 2 = 4 blue golf balls. The answer is 4. ✓

(c) Zero-shot

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?
A: The answer (arabic numerals) is

(Output) 8 X

(d) Zero-shot-CoT (Ours)

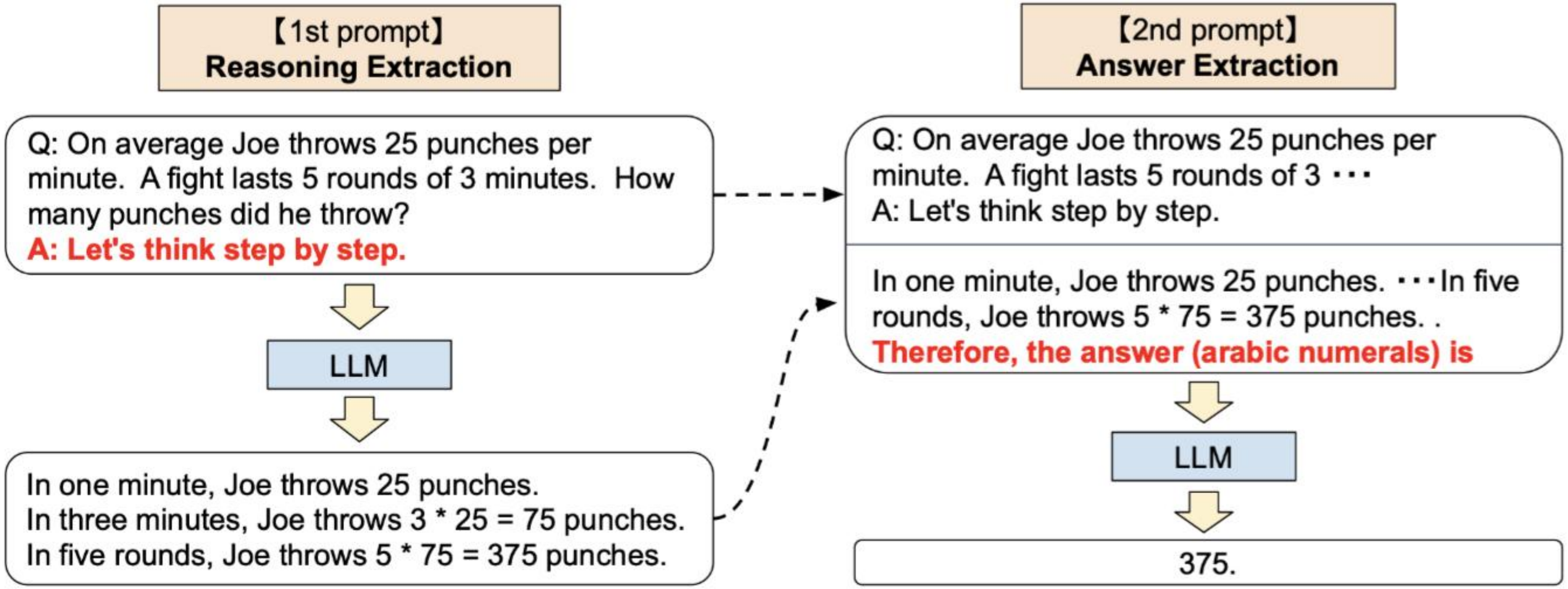
Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?
A: ***Let's think step by step.***

(Output) There are 16 balls in total. Half of the balls are golf balls. That means that there are 8 golf balls. Half of the golf balls are blue. That means that there are 4 blue golf balls. ✓

22_NIPS_Large Language Models are Zero-Shot Reasoners

思维链-零样本设定

CoT Zero-shot



【1st prompt】
Reasoning Extraction

Q: On average Joe throws 25 punches per minute. A fight lasts 5 rounds of 3 minutes. How many punches did he throw?
A: Let's think step by step.

LLM

In one minute, Joe throws 25 punches.
In three minutes, Joe throws 3 * 25 = 75 punches.
In five rounds, Joe throws 5 * 75 = 375 punches.

【2nd prompt】
Answer Extraction

Q: On average Joe throws 25 punches per minute. A fight lasts 5 rounds of 3 ···
A: Let's think step by step.

In one minute, Joe throws 25 punches. ···In five rounds, Joe throws 5 * 75 = 375 punches. .
Therefore, the answer (arabic numerals) is

LLM

375.

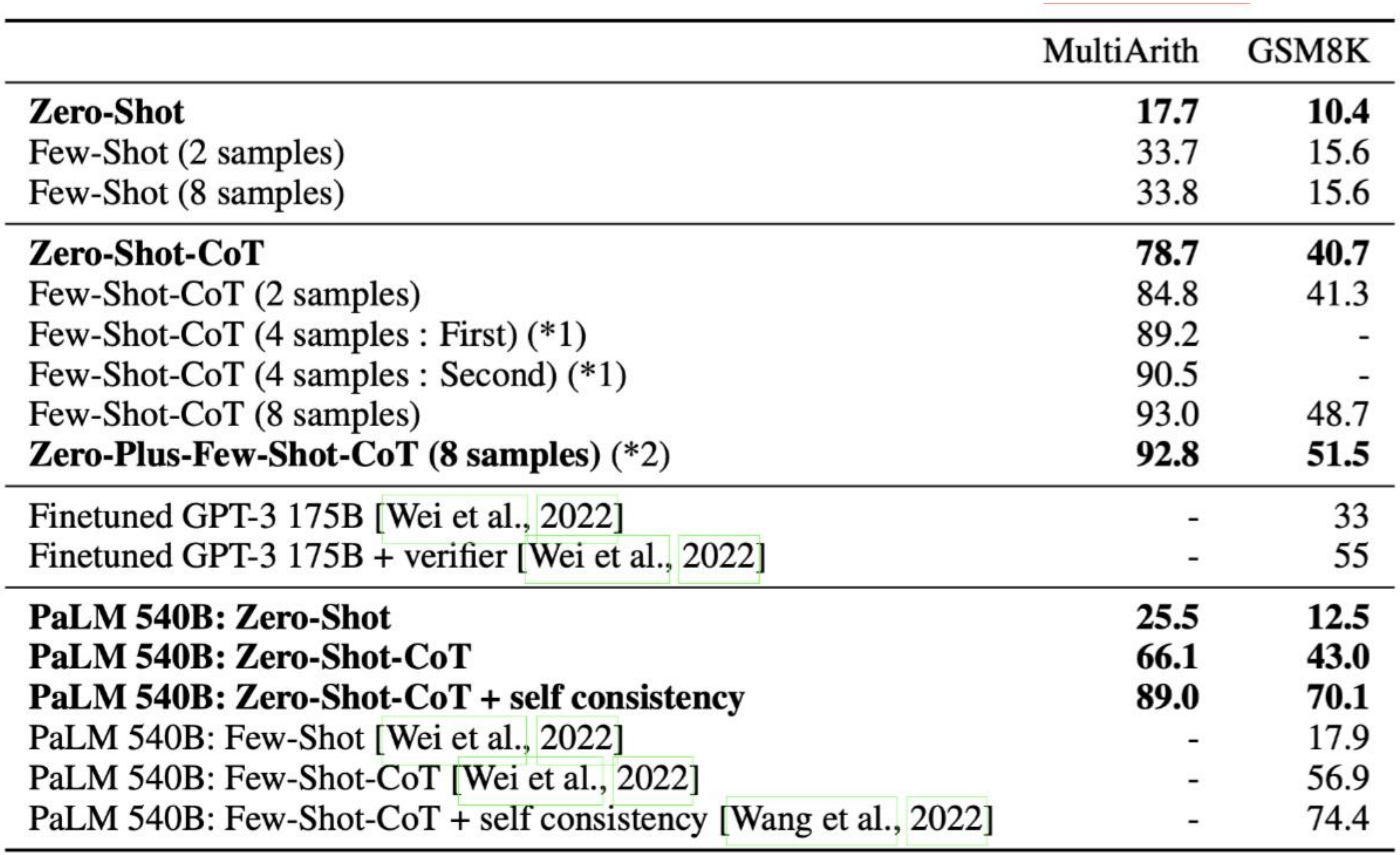
优点：不需要人工构建数据。

缺点：需要推理两次才能得出结果。

22_NIPS_Large Language Models are Zero-Shot Reasoners

思维链-零样本设定

CoT Zero-shot



	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**
Finetuned GPT-3 175B [Wei et al., 2022]	-	33
Finetuned GPT-3 175B + verifier [Wei et al., 2022]	-	55
PaLM 540B: Zero-Shot	**25.5**	**12.5**
PaLM 540B: Zero-Shot-CoT	**66.1**	**43.0**
PaLM 540B: Zero-Shot-CoT + self consistency	**89.0**	**70.1**
PaLM 540B: Few-Shot [Wei et al., 2022]	-	17.9
PaLM 540B: Few-Shot-CoT [Wei et al., 2022]	-	56.9
PaLM 540B: Few-Shot-CoT + self consistency [Wang et al., 2022]	-	74.4

结论：

1. 效果提升明显(17.7->78.7)

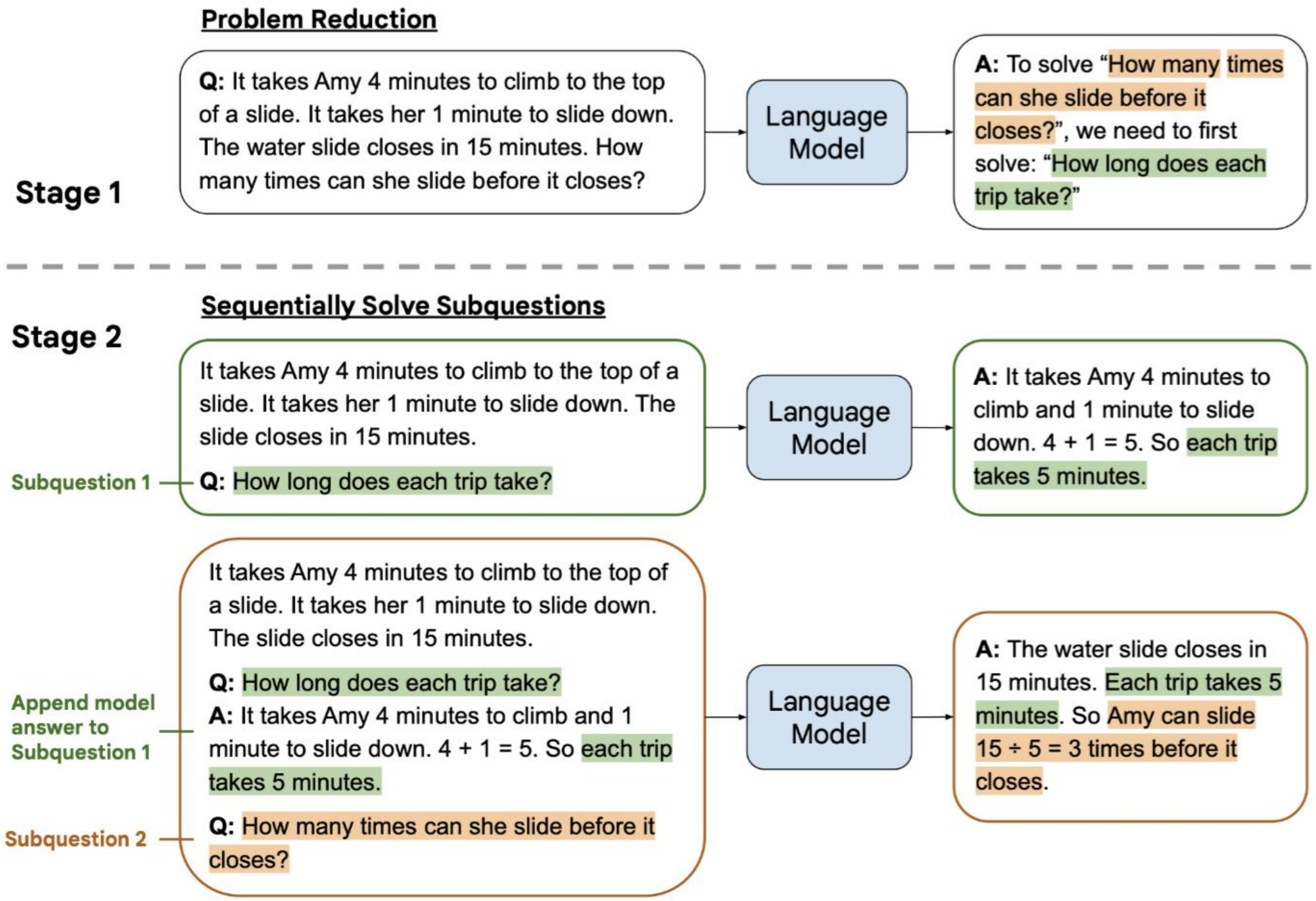
2. 在few-shot下加入魔法句在困难推理任务(GSM8K)上也有效果(48.7->51.5)。

3. Zero-shot CoT+self consistency效果提升明显(43.0->70.1)。

22_NIPS_Large Language Models are Zero-Shot Reasoners

思维链-多步推理

CoT Least-to-Most



Problem Reduction

Q: It takes Amy 4 minutes to climb to the top of a slide. It takes her 1 minute to slide down. The water slide closes in 15 minutes. How many times can she slide before it closes?

A: To solve "How many times can she slide before it closes?", we need to first solve: "How long does each trip take?"

Sequentially Solve Subquestions

It takes Amy 4 minutes to climb to the top of a slide. It takes her 1 minute to slide down. The slide closes in 15 minutes.

Q: How long does each trip take?

A: It takes Amy 4 minutes to climb and 1 minute to slide down. 4 + 1 = 5. So each trip takes 5 minutes.

It takes Amy 4 minutes to climb to the top of a slide. It takes her 1 minute to slide down. The slide closes in 15 minutes.

Q: How long does each trip take?

A: It takes Amy 4 minutes to climb and 1 minute to slide down. 4 + 1 = 5. So each trip takes 5 minutes.

Q: How many times can she slide before it closes?

A: The water slide closes in 15 minutes. Each trip takes 5 minutes. So Amy can slide 15 ÷ 5 = 3 times before it closes.

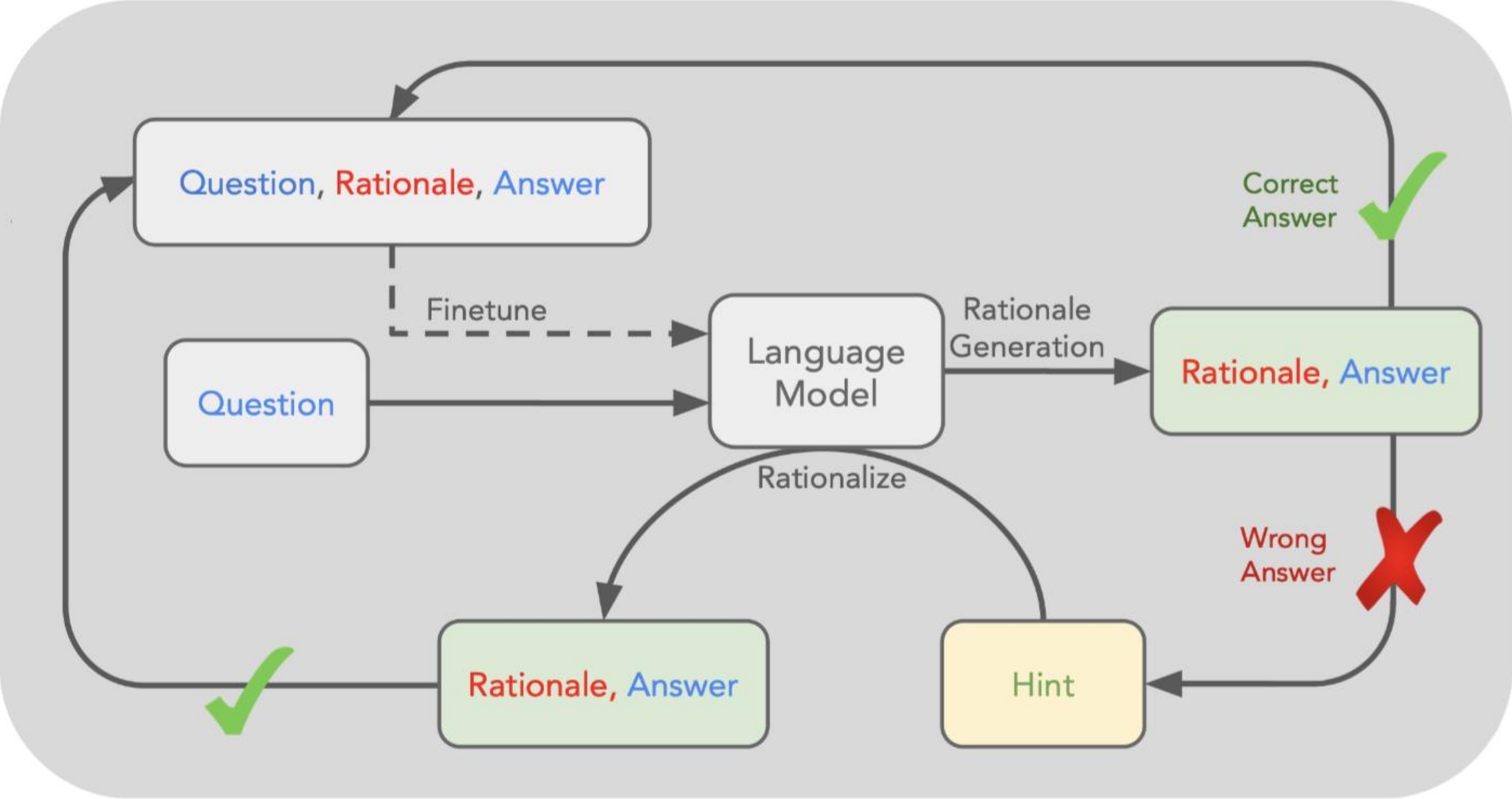
任务分解

依次处理子任务

23_ICLR_Least-to-Most Prompting Enables Complex Reasoning in Large Language Models

思维链-STaR

CoT-STaR

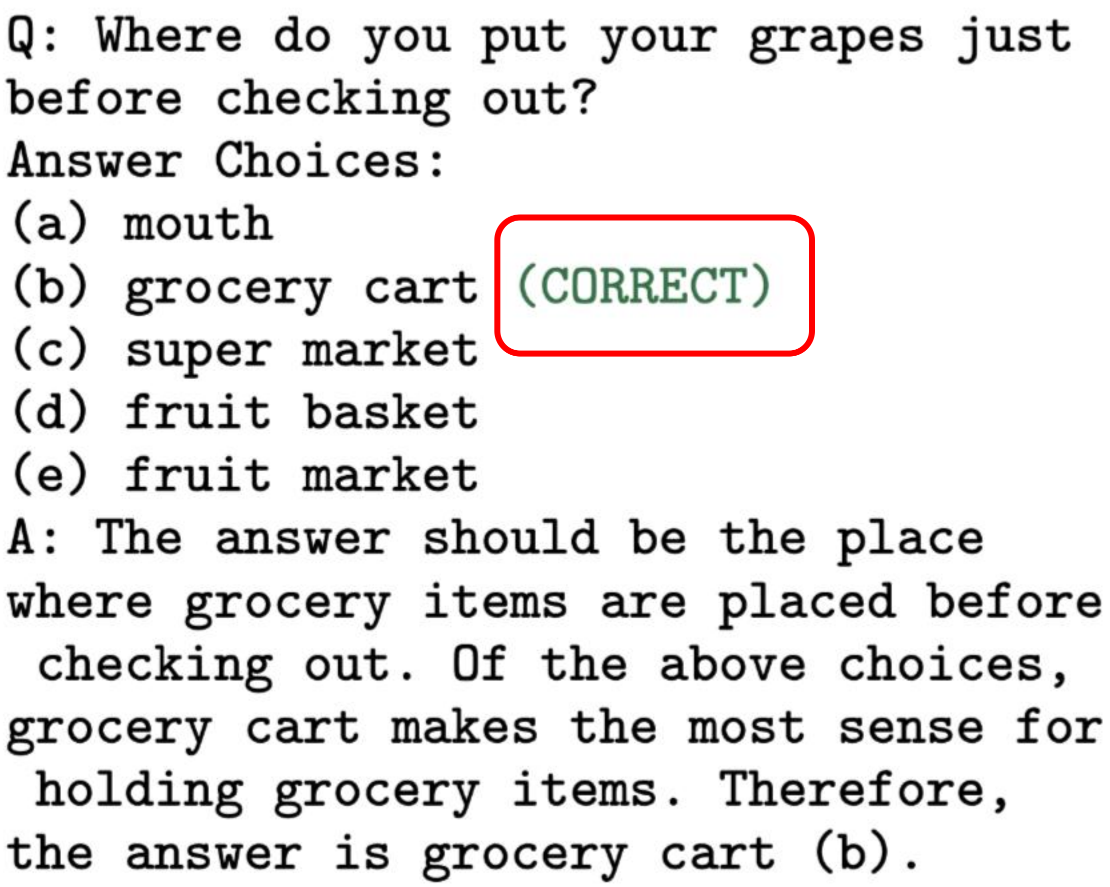


Q: What can be used to carry a small dog?
Answer Choices:
(a) swimming pool
(b) basket
(c) dog show
(d) backyard
(e) own home
A: The answer must be something that can be used to carry a small dog. Baskets are designed to hold things. Therefore, the answer is basket (b).

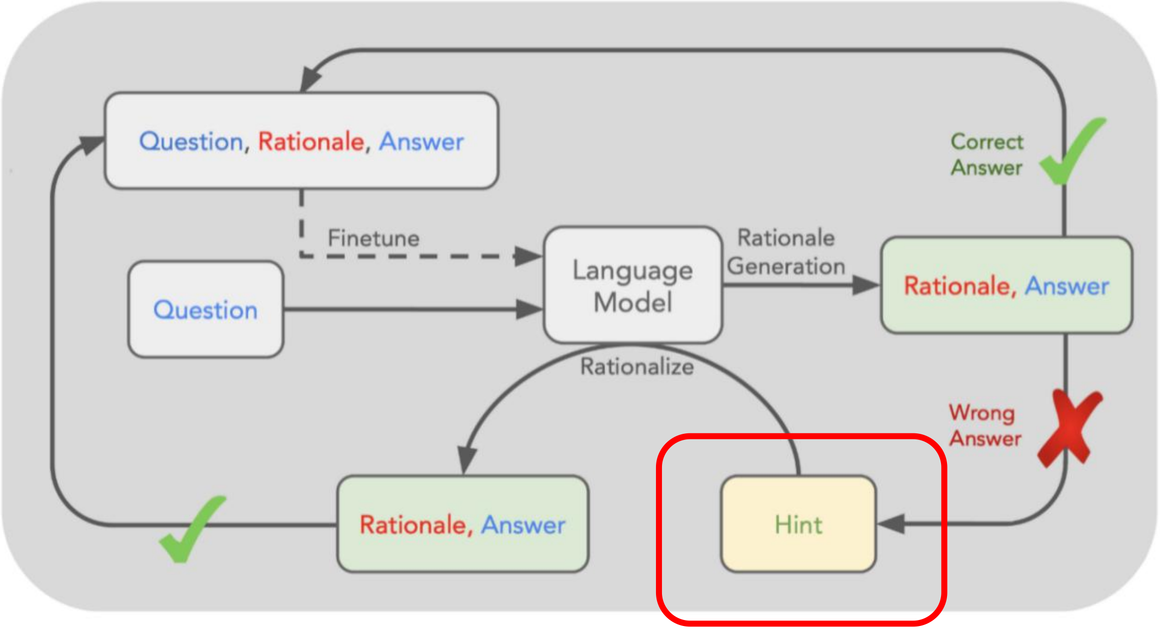
23_ICLR_Self-Consistency Improves Chain of Thought Reasoning in Language Models

思维链-STaR

CoT-STaR



Q: Where do you put your grapes just before checking out?
Answer Choices:
(a) mouth
(b) grocery cart (CORRECT)
(c) super market
(d) fruit basket
(e) fruit market
A: The answer should be the place where grocery items are placed before checking out. Of the above choices, grocery cart makes the most sense for holding grocery items. Therefore, the answer is grocery cart (b).



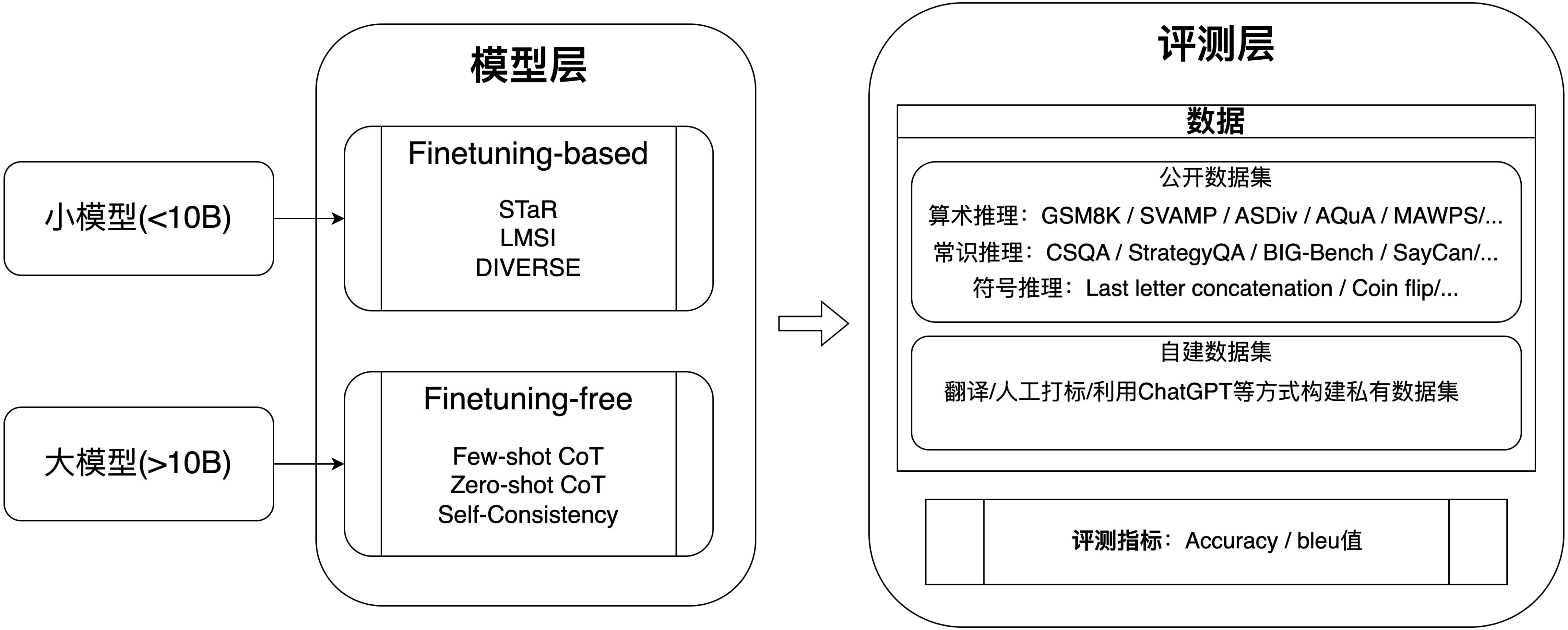
模型经过反复迭代逐渐达到稳态，对于达到稳态后模型还是不会的样本，该怎么办呢？这类样本没有为模型的训练起到任何作用。

是否可以考虑改进提示(hint)的方式进一步拓展该模型，把所有数据都用上？

23_ICLR_Self-Consistency Improves Chain of Thought Reasoning in Language Models

实施方案

Pipeline



小模型(<10B)
大模型(>10B)
模型层
Finetuning-based
STaR
LMSI
DIVERSE
Finetuning-free
Few-shot CoT
Zero-shot CoT
Self-Consistency
评测层
数据
公开数据集
算术推理：GSM8K / SVAMP / ASDiv / AQuA / MAWPS/...
常识推理：CSQA / StrategyQA / BIG-Bench / SayCan/...
符号推理：Last letter concatenation / Coin flip/...
自建数据集
翻译/人工打标/利用ChatGPT等方式构建私有数据集
评测指标：Accuracy / bleu值