

# 思维链

Chain-of-Thought (CoT)

汇报人: Chengcheng Han

2023年03月16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。

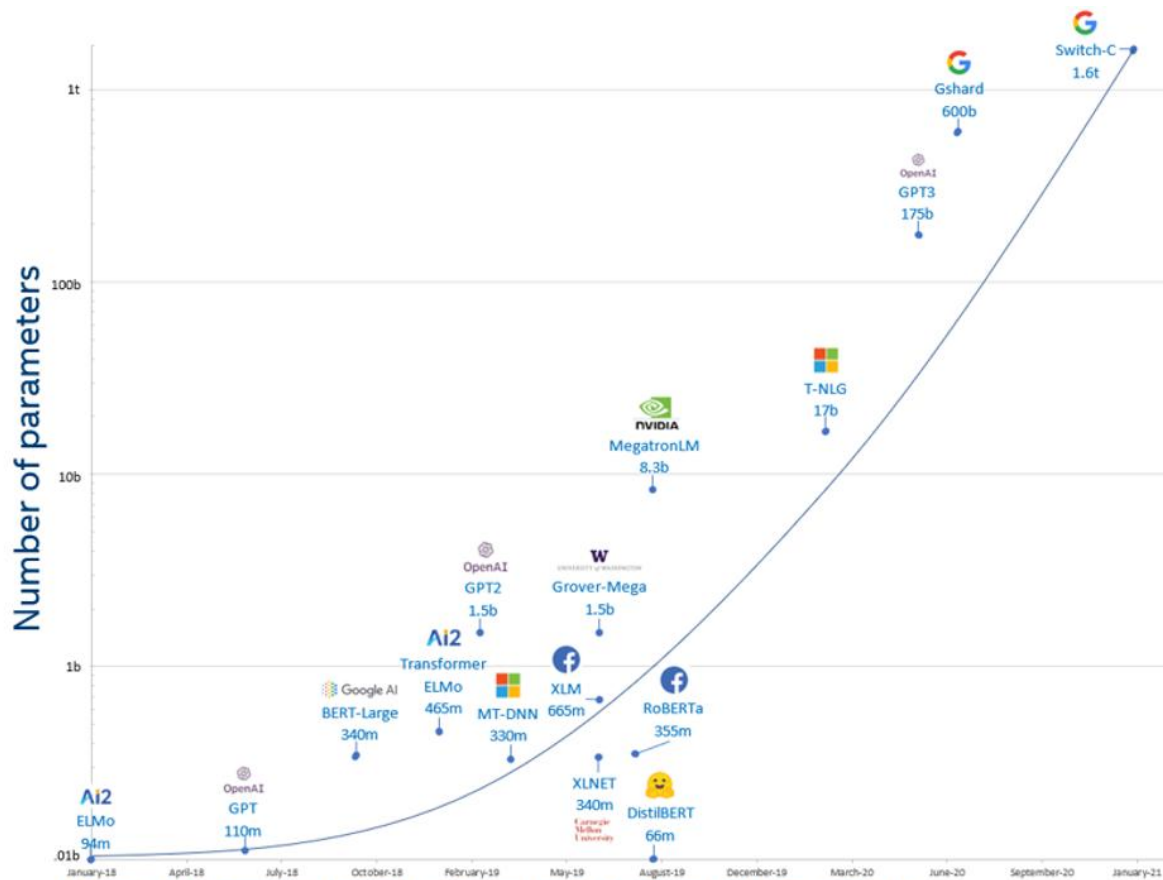
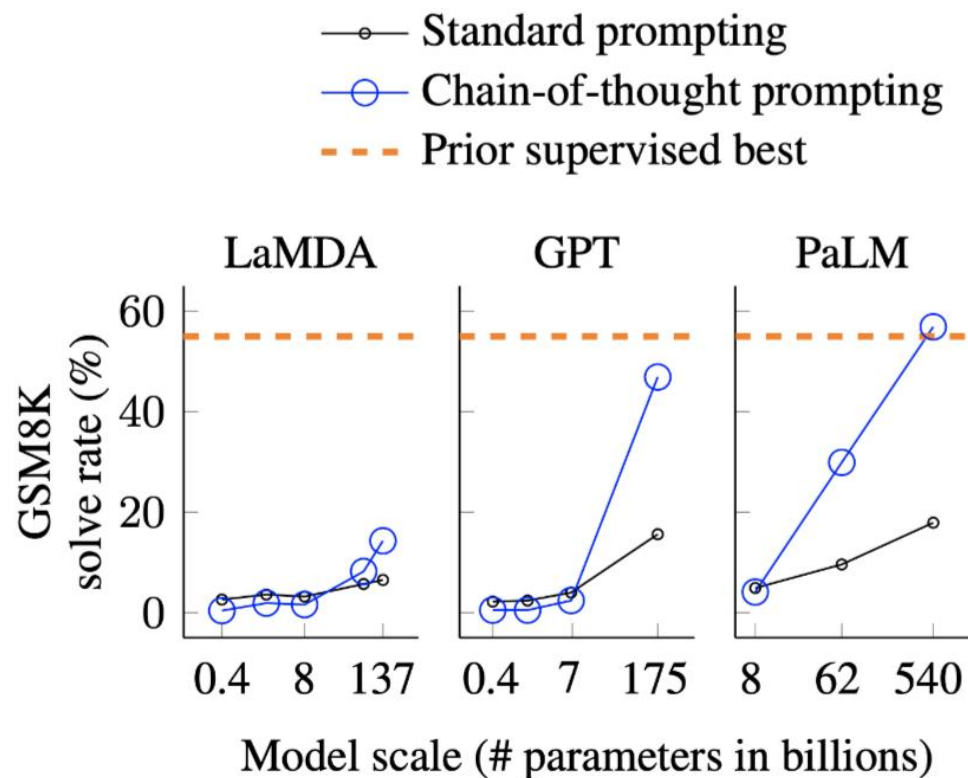


Figure 1: Exponential growth of number of parameters in DL models



# 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?

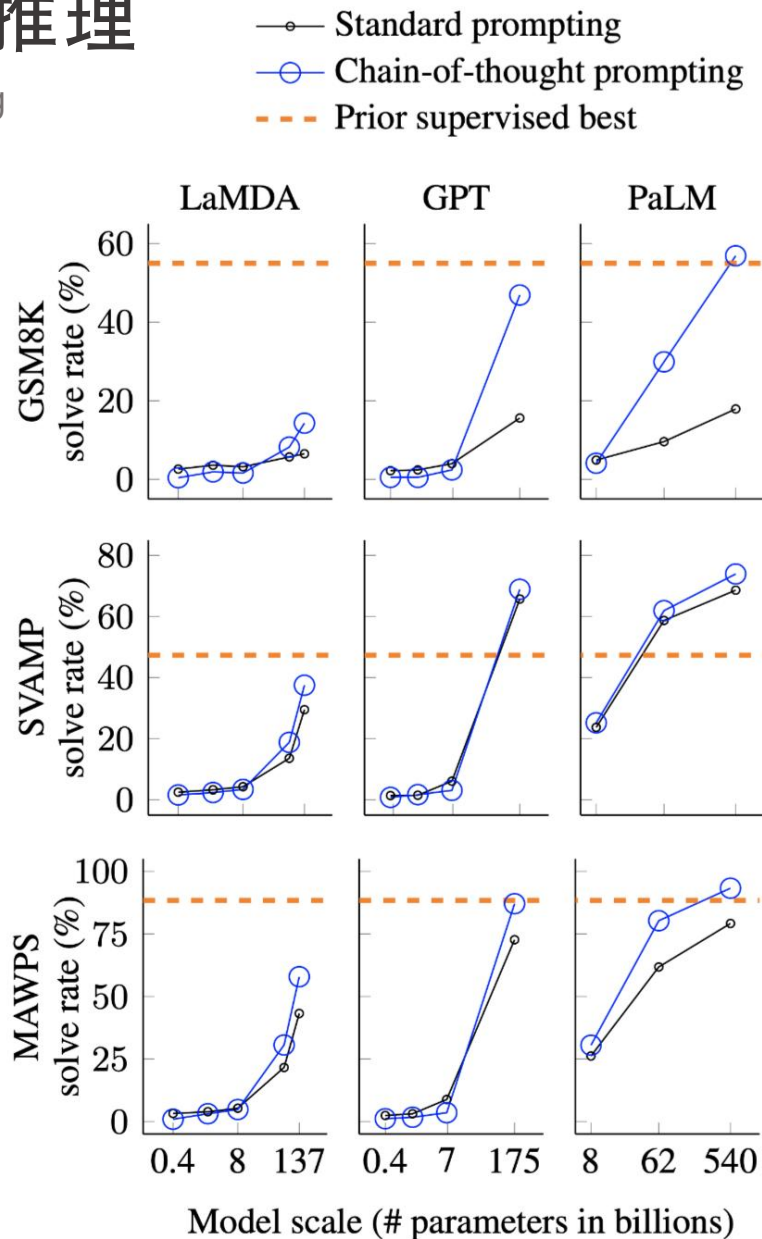
### Model Output

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

# 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。

# 思维链-分析

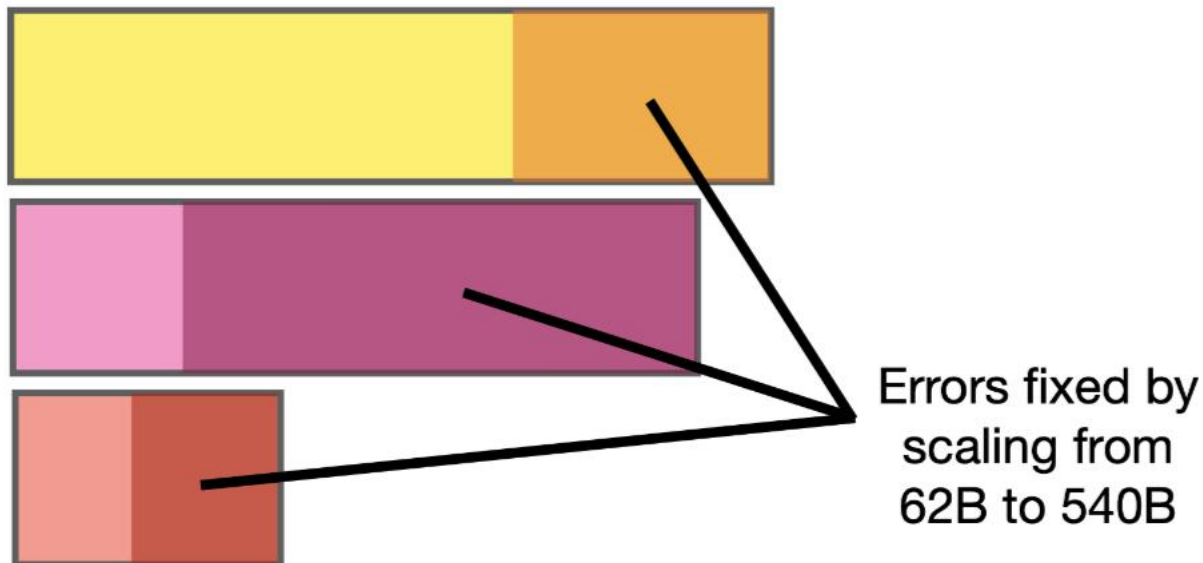
CoT-Analysis

## Types of errors made by a 62B language model:

**Semantic understanding**  
(62B made 20 errors of this type,  
540B fixes 6 of them)

**One step missing**  
(62B made 18 errors of this type,  
540B fixes 12 of them)

**Other**  
(62B made 7 errors of this type,  
540B fixes 4 of them)



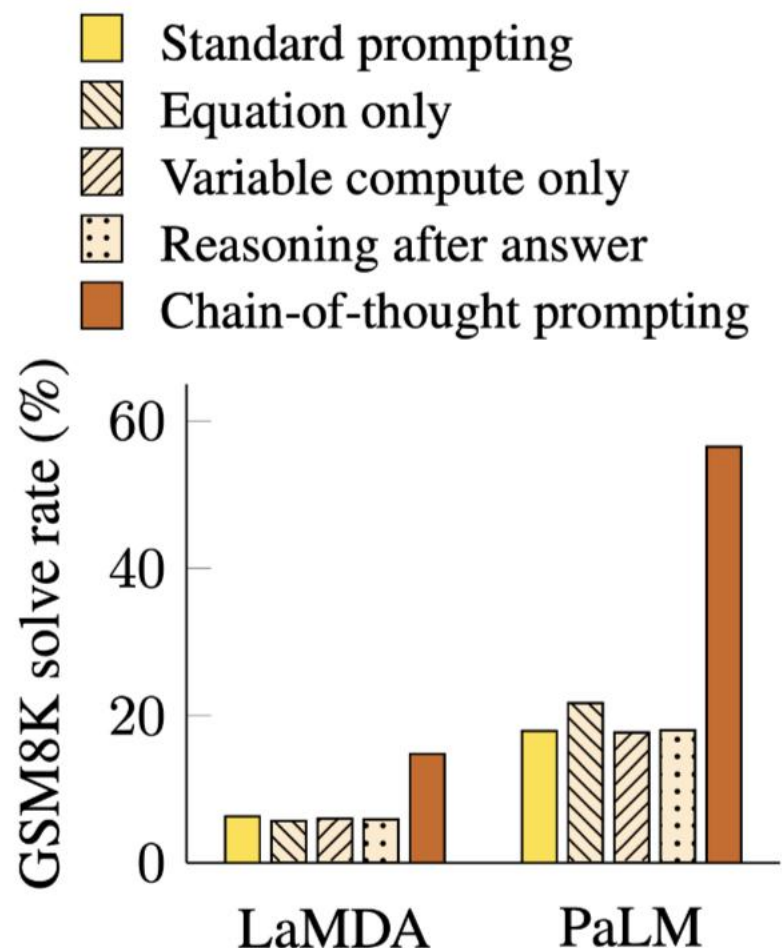
通过增加PaLM的参数量（62B->540B），模型可以修复很大一部分错误。

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。

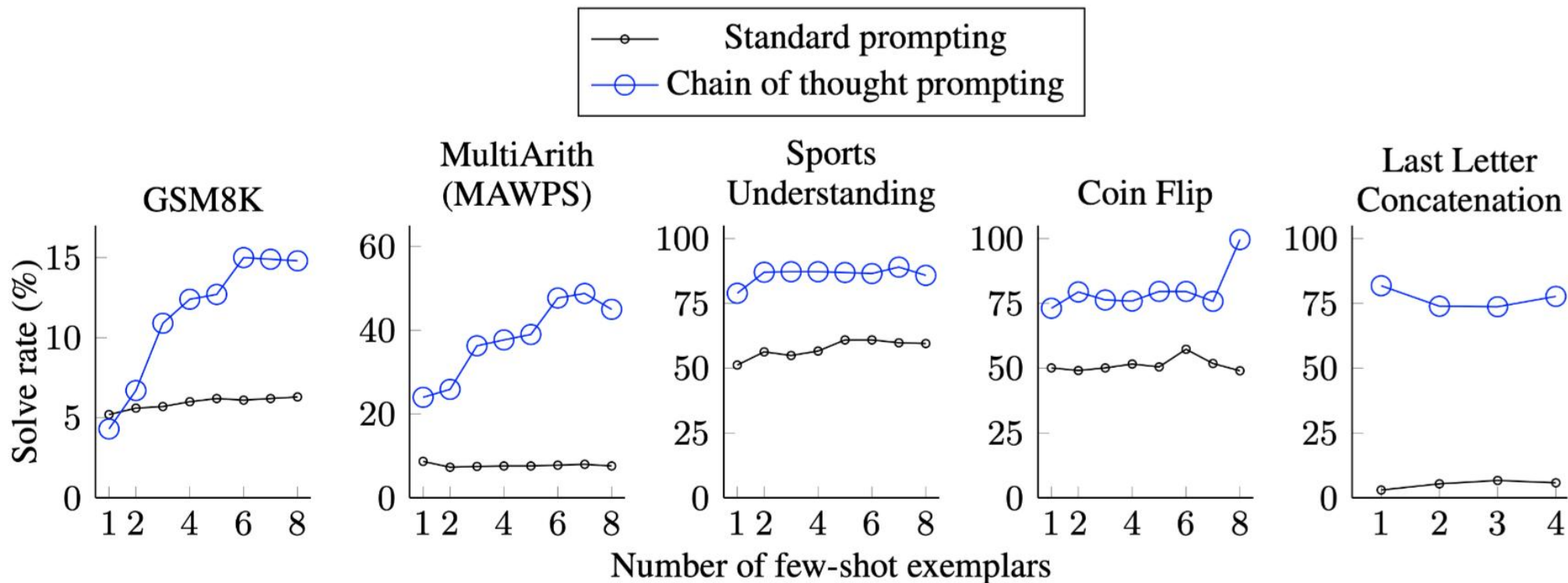
1和2说明CoT中的自然语言部分很重要。

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。



# 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。

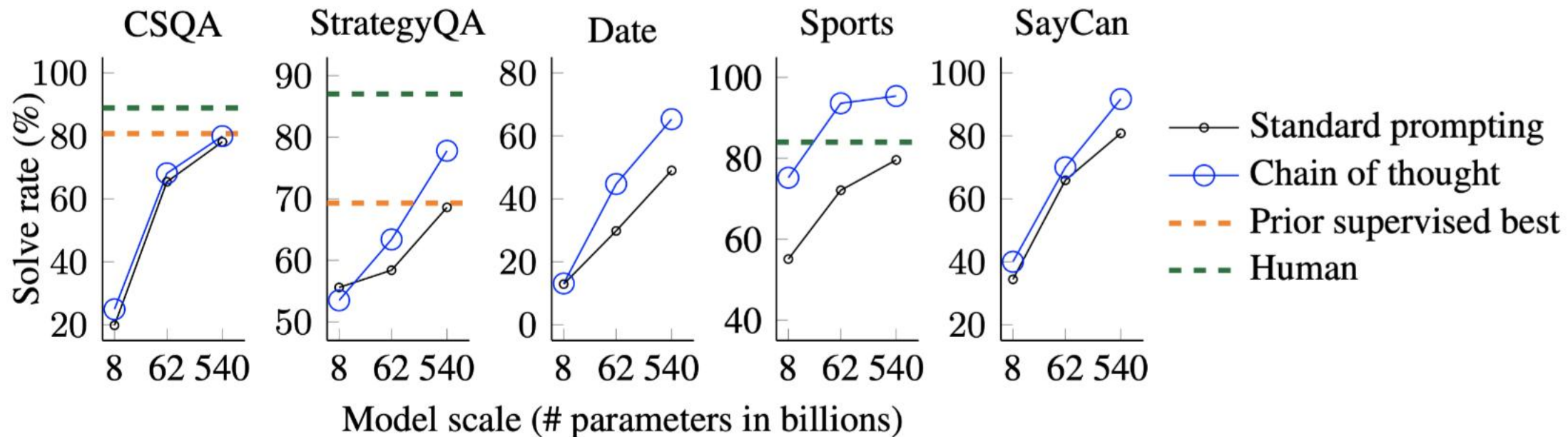
因此未来的一个方向可能是探索一种annotation的模型来得到稳健的CoT (Prompts)。

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



# 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。

# 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	<b>18.8</b>	0.0	0.2	0.0	0.0	70.4	67.1	51.6	52.2	48.7	50.4
LaMDA	420M	0.3	<b>1.6</b>	0.0	0.0	0.0	0.0	52.9	49.6	50.0	50.5	49.5	49.1
	2B	2.3	<b>6.0</b>	0.0	0.0	0.0	0.0	54.9	<b>55.3</b>	47.4	48.7	49.8	50.2
	8B	1.5	<b>11.5</b>	0.0	0.0	0.0	0.0	52.9	<b>55.5</b>	48.2	49.6	51.2	50.6
	68B	4.4	<b>52.0</b>	0.0	<b>0.8</b>	0.0	<b>2.5</b>	56.2	<b>83.2</b>	50.4	<b>69.1</b>	50.9	<b>59.6</b>
	137B	5.8	<b>77.5</b>	0.0	<b>34.4</b>	0.0	<b>13.5</b>	49.0	<b>99.6</b>	50.7	<b>91.0</b>	49.1	<b>74.5</b>
PaLM	8B	2.6	<b>18.8</b>	0.0	0.0	0.0	<b>0.2</b>	60.0	<b>74.4</b>	47.3	<b>57.1</b>	50.9	<b>51.8</b>
	62B	6.8	<b>85.0</b>	0.0	<b>59.6</b>	0.0	<b>13.4</b>	91.4	<b>96.8</b>	43.9	<b>91.0</b>	38.3	<b>72.4</b>
	540B	7.6	<b>99.4</b>	0.2	<b>94.8</b>	0.0	<b>63.0</b>	98.1	<b>100.0</b>	49.3	<b>98.6</b>	54.8	<b>90.2</b>

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。

# 思维链-局限性

CoT-Limitations

该CoT方法的局限性：

## 1. 手工构建CoT的cost很大。

小样本场景可以接受，但无法用于微调（即大规模的人工标注CoT是不可行的，该问题或许可以被自动数据生成和零样本数据生成方法解决）。

## 2. 无法保证正确的推理路径，这会导致错误的答案。

改进语言模型的事实生成(factual generations, 狭义指推理路径，广义指所有生成的事实性内容)是未来工作的一个方向(或许可以结合知识图谱)。

## 3. CoT只在大模型上有效（>10B）。

在真实世界中很多场景无法使用大型LLM，探索如何在小模型上使CoT奏效是一个未来的研究方向。换句话说，进一步的研究可以探索如何在较小的模型中诱导模型推理（生成推理过程）。

但是小模型是否真的有这样的能力呢？没有的话如何引导应该都是徒劳吧？或许小模型需要通过带有CoT的语料微调使其先习得推理的能力更合理一些？

# 思维链-一致性方法

CoT-Self-Consistency

## Greedy decode

Chain-of-thought prompting

Prompt

Language model

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  $9 \text{ eggs} * \$2 = \$18$ .

**The answer is \$18.**

Marginalize out reasoning paths to aggregate final answers

**The answer is \$18.**



# 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. ✗

(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 ✗

(d) Zero-shot-CoT (Ours)

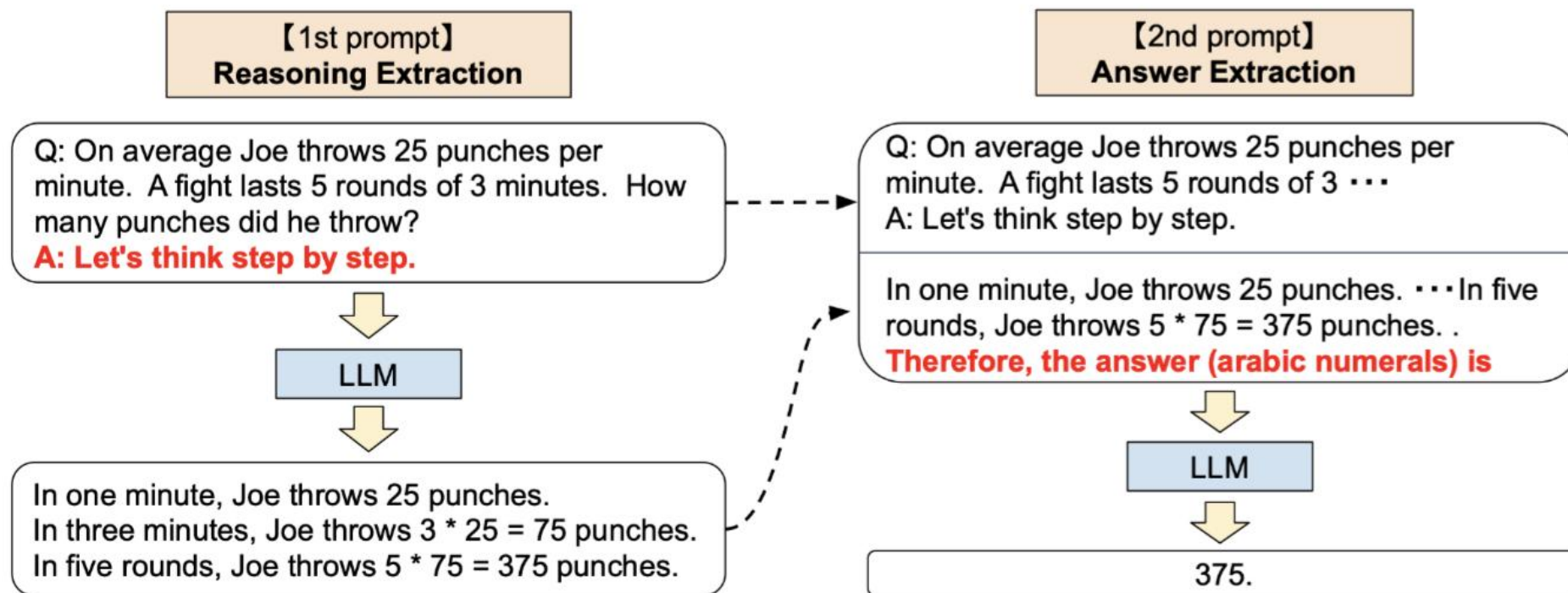
Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

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. ✓

# 思维链-零样本设定

CoT Zero-shot



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。

# 思维链-零样本设定

CoT Zero-shot

	MultiArith	GSM8K
<b>Zero-Shot</b>	<b>17.7</b>	<b>10.4</b>
Few-Shot (2 samples)	33.7	15.6
Few-Shot (8 samples)	33.8	15.6
<b>Zero-Shot-CoT</b>	<b>78.7</b>	<b>40.7</b>
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
<b>Zero-Plus-Few-Shot-CoT (8 samples) (*2)</b>	<b>92.8</b>	<b>51.5</b>
Finetuned GPT-3 175B [Wei et al., 2022]	-	33
Finetuned GPT-3 175B + verifier [Wei et al., 2022]	-	55
<b>PaLM 540B: Zero-Shot</b>	<b>25.5</b>	<b>12.5</b>
<b>PaLM 540B: Zero-Shot-CoT</b>	<b>66.1</b>	<b>43.0</b>
<b>PaLM 540B: Zero-Shot-CoT + self consistency</b>	<b>89.0</b>	<b>70.1</b>
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

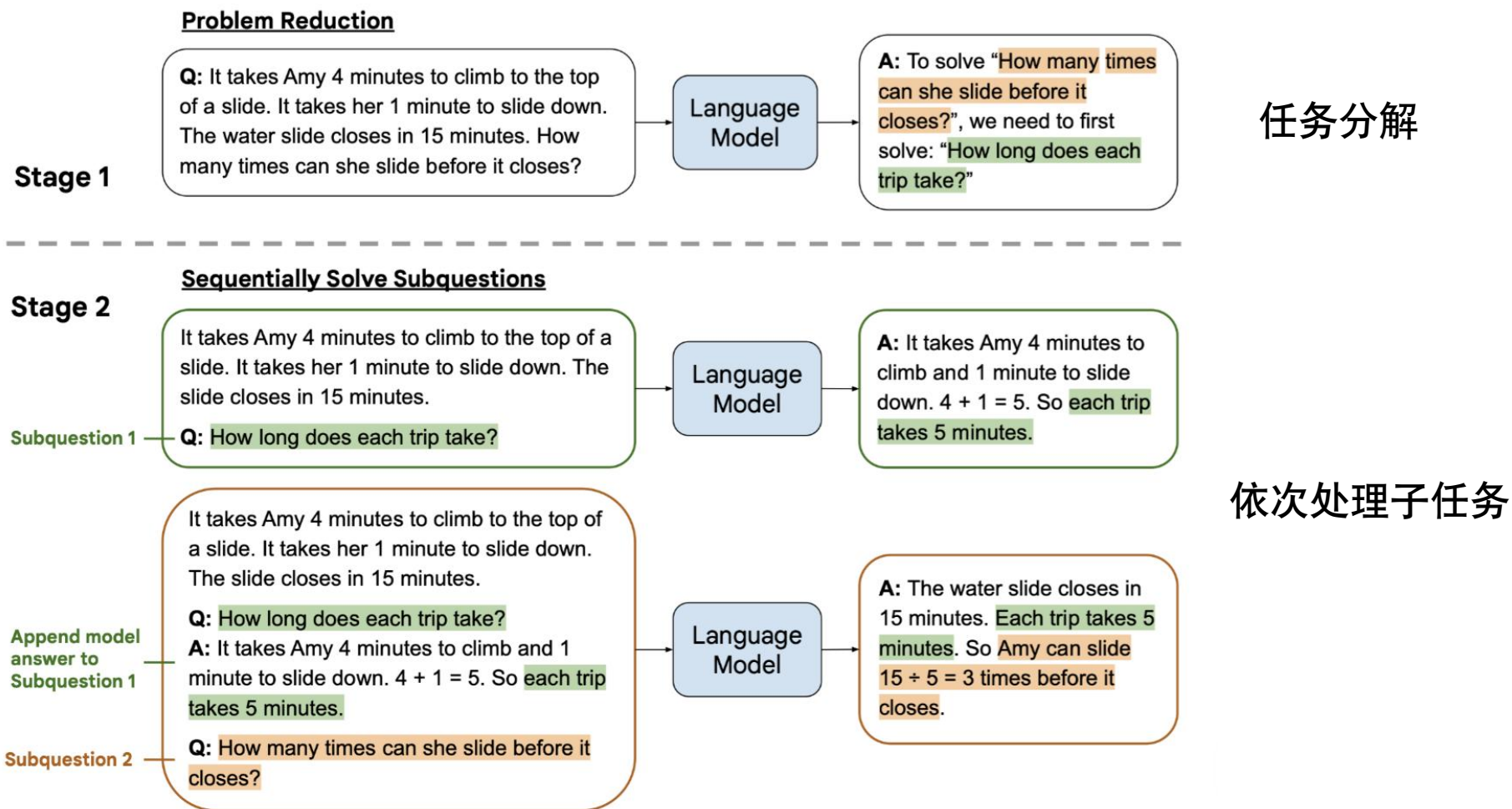
结论:

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)。



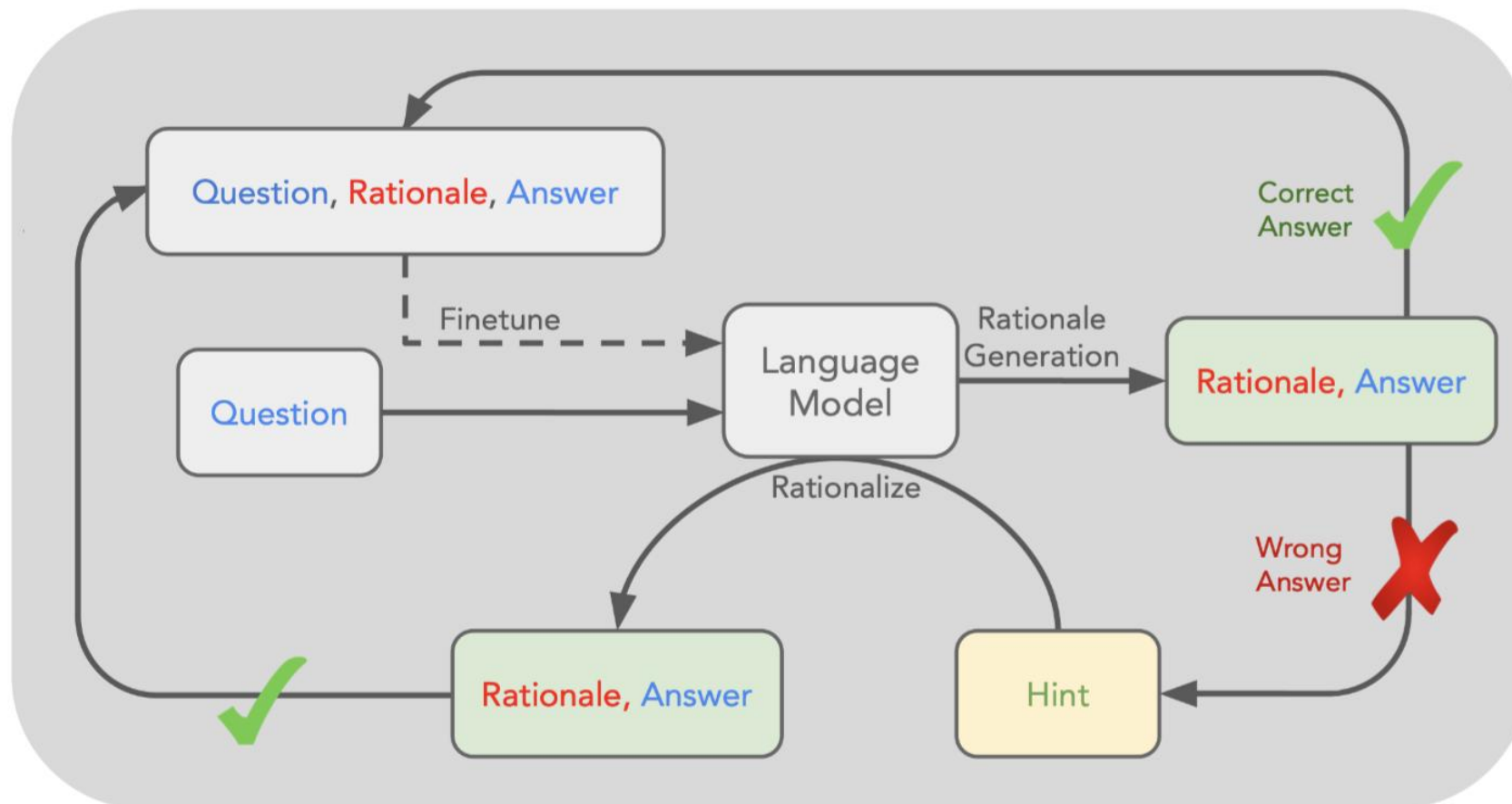
# 思维链-多步推理

CoT Least-to-Most



# 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).

# 思维链-STaR

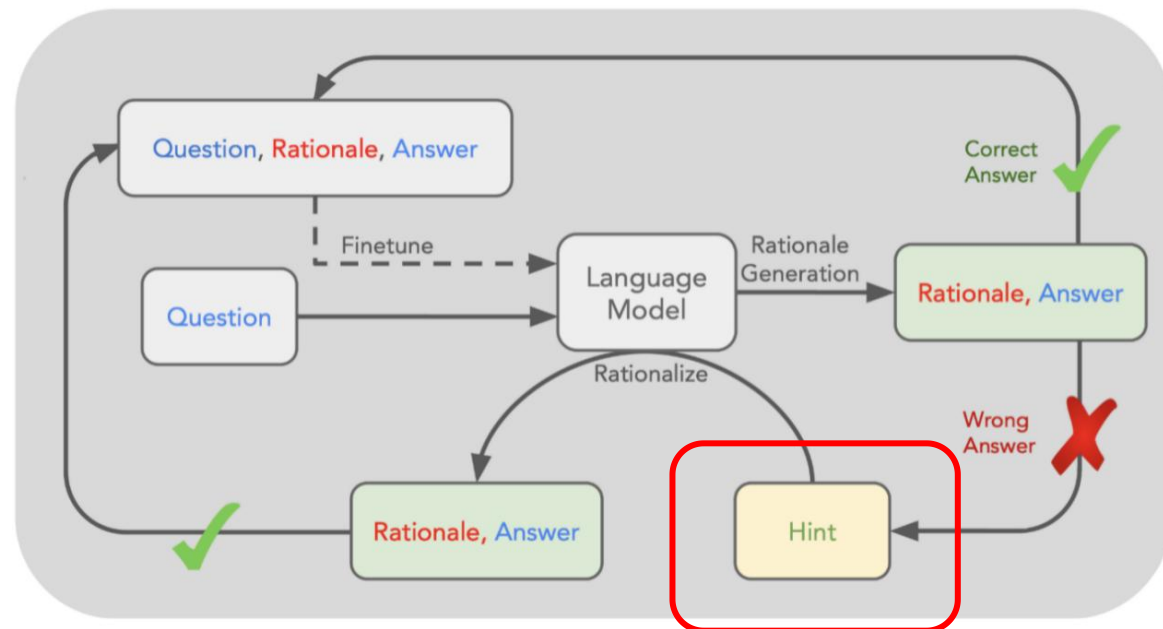
CoT-STaR

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？

# 思维链-其他

CoT-Other

**Task instruction** { Answer these questions by identifying whether the second sentence is an appropriate paraphrase of the first, metaphorical sentence.

**Few-shot example #1** { Q: David's eyes were like daggers at Paul when Paul invited his new girlfriend to dance. <- -> David had two daggers when Paul invited his new girlfriend to dance.  
choice: True  
choice: False  
A: False

**Answer explanation** { Explanation: David's eyes were not literally daggers, it is a metaphor used to imply that David was glaring fiercely at Paul.

4 more examples + explanations

**Target question** { Q: Our whole life we swim against the waves towards the green light of happiness. <- -> Our whole life we try to reach happiness.  
choice: True  
choice: False  
A:

**Context:**

wolves are afraid of mice

sheep are afraid of wolves

emily is a wolf

mice are afraid of wolves

winona is a wolf

cats are afraid of sheep

jessica is a cat

gertrude is a sheep

Question: what is emily afraid of?

**Selection:** emily is a wolf and

wolves are afraid of mice

**Inference:** emily is afraid of mice

**Context:**

julius is a swan

julius is gray

bernhard is a swan

bernhard is yellow

brian is a lion

greg is a lion

brian is white

lily is a frog

lily is gray

Question: what color is greg?

**Selection:** brian is a lion. We

know that brian is white

**Inference:** we inferred that lions are often white

**Selection:** greg is a lion. We

know that we inferred that lions are often white

**Inference:** greg is white

2204\_Can language models learn from explanations in context?

23\_ICLR\_Self-Consistency Improves Chain of Thought Reasoning in Language Models



# 思维链-未来方向

CoT-Other

1. 具体分析当前CoT无法解决的推理错误类型，探索无法解决的原因，更有针对性的对当前CoT进行改进。
2. 探索预处理数据，模型架构和优化目标(预训练任务)和CoT之间有着何种关系，以此为基础为不同的Model设计不同的CoT。
3. 探索不同的CoT的设计和构建方式/exemplars的选择/数量/顺序对模型的影响。
4. 自动生成CoT(或者其他降低人工构建CoT的成本的方法)。
5. 针对特定任务特定应用场景设计具体的CoT最大化发挥LLM的性能(通过CoT或者说Prompts将LLM引导到特定领域中)。
6. 提升模型生成的事实性内容的正确性，引入知识图谱或其他结构性知识？
7. 将CoT拓展到小型LLM上(目前的CoT只是在大型LLM(>10B)上有用)。
8. 探索其他形式的prompting构建方法以拓展LLM的任务列表及任务性能(CoT可以看作是一种prompting的构建方法)。