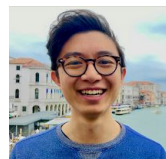
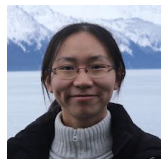


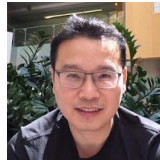
# Self-Consistency Improves Chain of Thought Reasoning in Language Models

Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc Le, Ed Chi,  
Sharan Narang, Aakanksha Chowdhery, Denny Zhou



Google Brain

[xuezhiw@google.com](mailto:xuezhiw@google.com)



# Recap: Chain-of-thought (CoT) prompting

## Standard Prompting

### Input

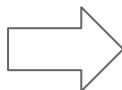
Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: The answer is 11.

Q: 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

### 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. ✅

[Chain of thought prompting elicits reasoning in large language models](#) (NeurIPS 2022).

J. Wei, X. Wang, D. Schuurmans, M. Bosma, B. Ichter, F. Xia, E. Chi, Q. Le, & D. Zhou.

# CoT prompting greedily decodes the optimal reasoning path

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: 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

## 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 final answer is incorrect due to wrong reasoning paths
- **Different people think differently:** can we do better by generating multiple possible reasoning paths?

# Self-consistency: step 1 - prompt

## Prompt with chain of thought

**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:** 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

# Self-consistency: step 2 - sample decode

## Prompt with chain of thought

**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:** 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 decode with diverse 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 uses  $3 + 4 = 7$  eggs every day. So in total she sells  $7 * \$2 = \$14$  per day.

**The answer is \$14.**

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.**

Note the reasoning paths are optional, so they can be marginalized out

# Self-consistency: step 3 - majority vote

## Prompt with chain of thought

**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:** 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 decode with diverse 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 uses  $3 + 4 = 7$  eggs every day. So in total she sells  $7 * \$2 = \$14$  per day.

The answer is \$14.

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.

Majority vote on the answers

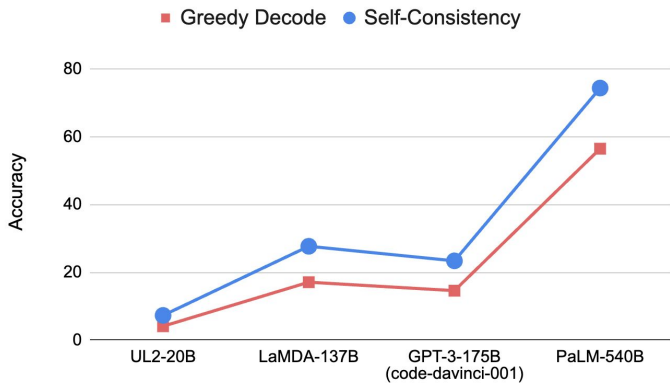
The answer is \$18.

# Self-consistency is simple but effective

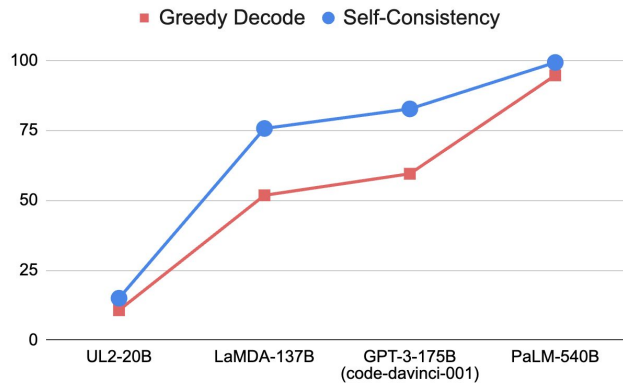
- Simple compared to many other existing works
  - No fine-tuning, no human annotation, no additional modules like a verifier or a re-ranker
- **Striking performance gains** across:
  - Four LLMs with varying scales: UL2-20B, LaMDA-137B, PaLM-540B, GPT-3 (175B)
- **SoTA performance** across:
  - Various reasoning benchmarks: arithmetic, commonsense, and symbolic

# Arithmetic reasoning

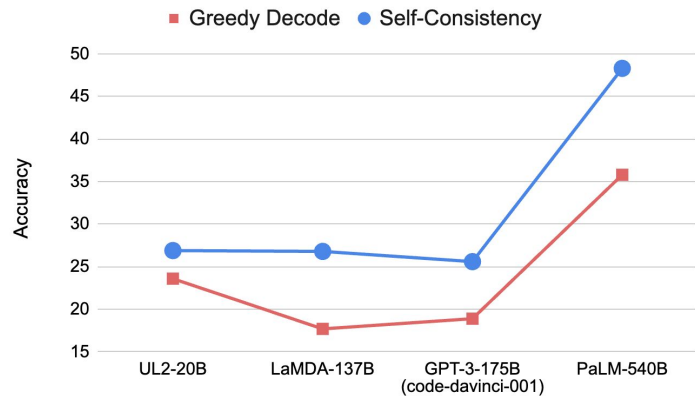
## GSM8K



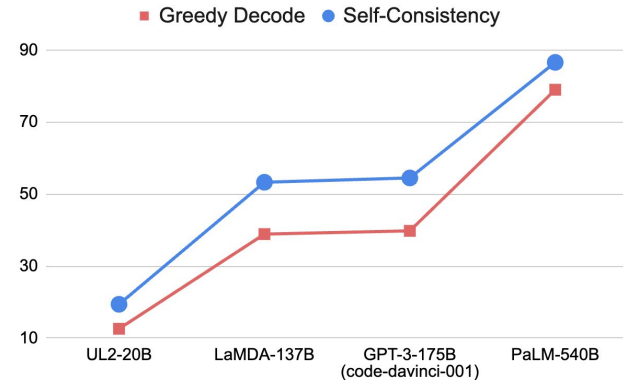
## MultiArith



## AQuA



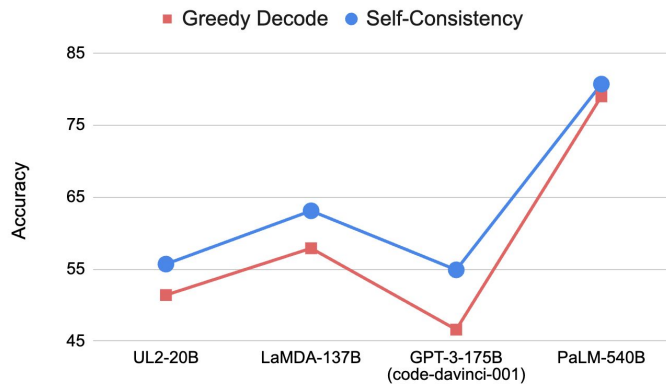
## SVAMP



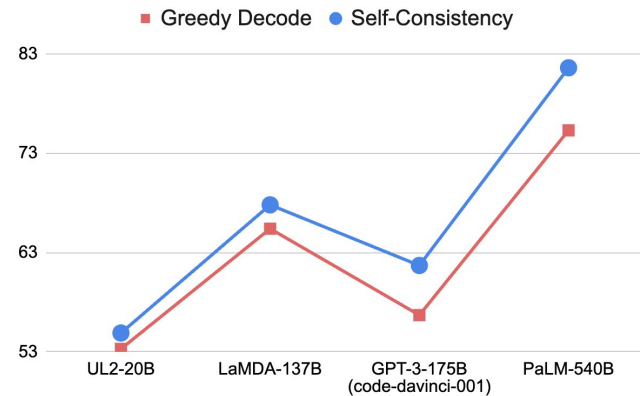


# Commonsense reasoning

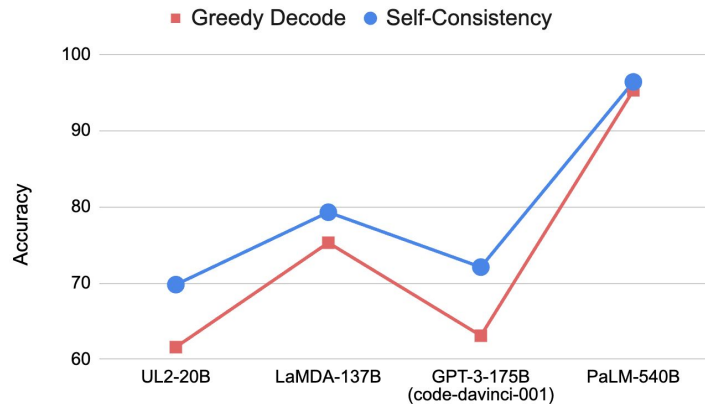
### CommonsenseQA



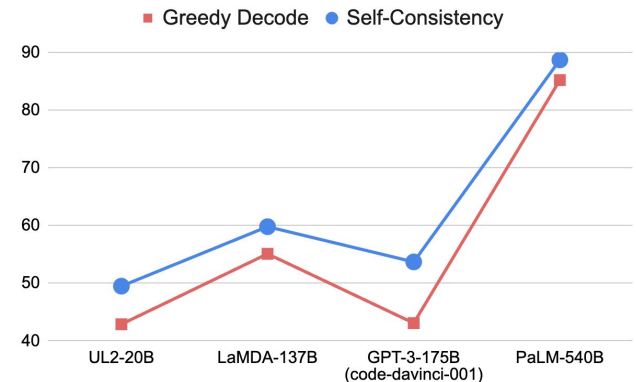
### StrategyQA



### ARC-easy



### ARC-challenge



# Self-consistency works with zero-shot CoT

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. ✓*

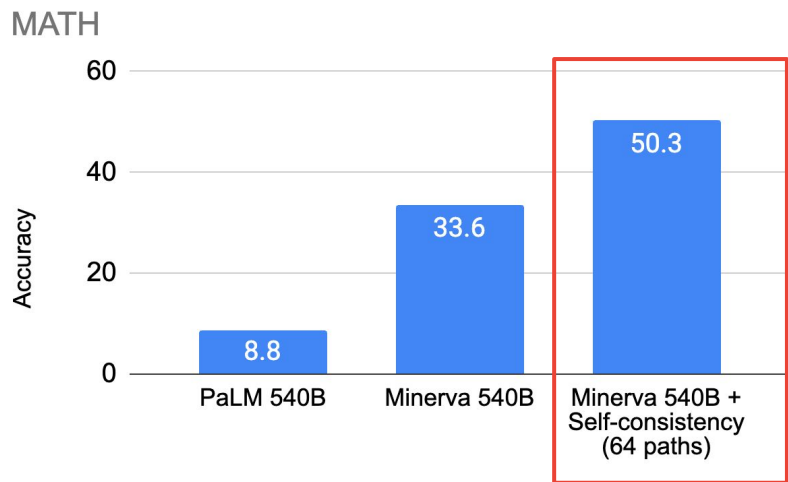
Kojima et al. [Large Language Models are Zero-Shot Reasoners](#). NeurIPS 2022.

		GSM8K accuracy
PaLM-540B	Zero-shot CoT (Kojima et al., 2022)	43.0
	+ Self-consistency (40 paths)	<b>69.2</b>

Self-consistency works with “let’s think step by step” too!



# Self-consistency achieves SoTA in Minerva and Flan-PaLM



Lewkowycz et al. [Solving Quantitative Reasoning Problems with Language Models](#). 2022.

## MMLU

-	Random	25.0
-	Average human rater	34.5
May 2020	GPT-3 5-shot	43.9
Mar. 2022	Chinchilla 5-shot	67.6
Apr. 2022	PaLM 5-shot	69.3
Oct. 2022	<b>Flan-PaLM 5-shot</b>	<b>72.2</b>
-	<b>Flan-PaLM 5-shot: CoT + SC</b>	<b>75.2</b>
-	Average human expert	89.8

Chung et al. [Scaling Instruction-Finetuned Language Models](#). 2022.

# Check out our paper!

- ID 11718: Self-Consistency Improves Chain of Thought Reasoning in Language Models
- <https://arxiv.org/abs/2203.11171>
- Questions: xuezhiw@google.com