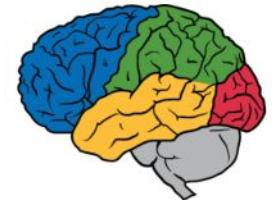


# Teach Language Models to Reason

Denny Zhou

Google Research, Brain Team



# Human intelligence vs machine learning

- Humans can learn from only few examples while machine learning needs tons of labeled data to train a model
- Humans can explain their predictions or decisions while machine learning is essentially a black box
- Humans has strong out-of-distribution generalization abilities, e.g., easy-to-hard generalization

## Attempts to fill the gap in the past decades

- Bayesian machine learning, kernel machines, nonparametric, sparsity, large-margin, semi-supervised learning, manifold learning, graph-based learning, transductive learning, meta learning, transfer learning, active learning, domain adaptation, structural learning, low-rank, ...

# Attempts to fill the gap in the past decades

- Bayesian machine learning, kernel machines, nonparametric, sparsity, large-margin, semi-supervised learning, manifold learning, graph-based learning, transductive learning, meta learning, transfer learning, active learning, domain adaptation, structural learning, low-rank, ...



# What is really missed in machine learning?

What is really missed in machine learning?

# Reasoning

Reasoning leads to the gap between human intelligence and machine learning

We have found a simple way to solve reasoning:

**Teach language models to  
reason, like teaching kids**

# Let us start from a toy machine learning problem

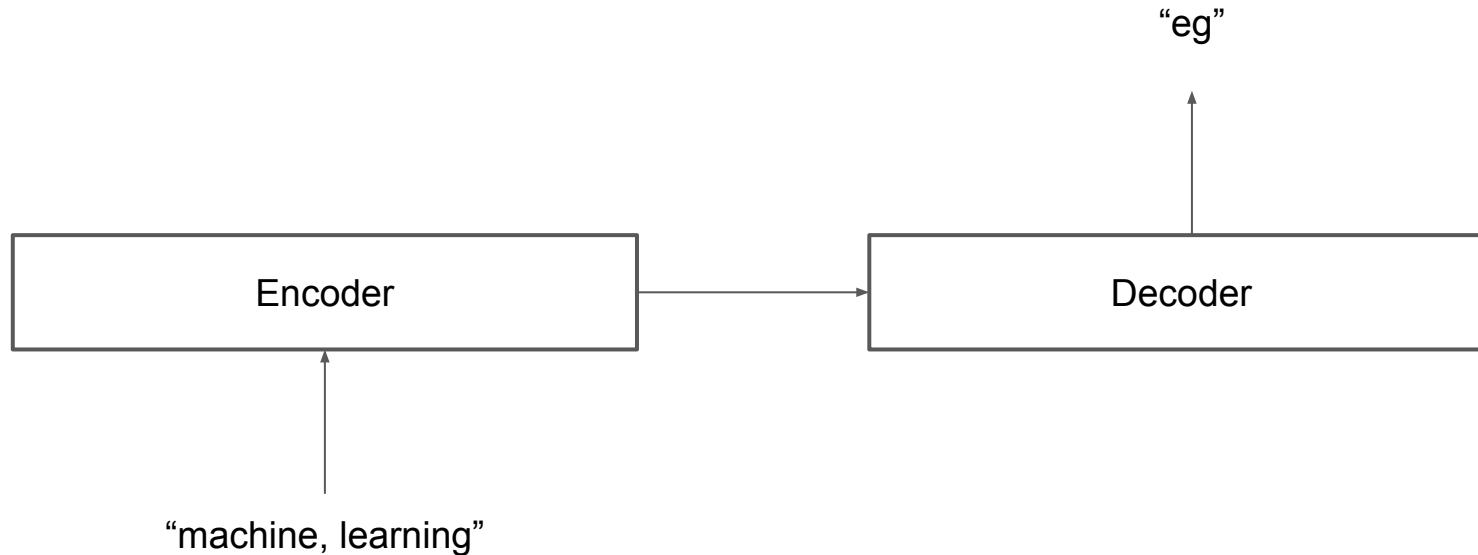
Input	Output
“machine, learning”	“eg”
“artificial, intelligence”	“le”
...	...

**Rule:** Take the last letter of each word, and then concatenate them

# Why created such a toy task?

- **Trivial for humans, but really difficult for machine learning**
- Not seen by LLMs in their pretraining
- Few-shot prompting completely fails to solve it

# Solve by machine learning? Tons of labels are needed!



# Can OpenAI's GPT-3 solve this simple last-letter task?

Playground

Q&A



Save

View code

Share



Q: "machine, learning"  
A: "eg"

Q: "artificial, intelligence"  
A: "le"

Q: "think, machine"  
A:

Mode



Engine

text-davinci-002

Temperature 0

Maximum length 100

Stop sequences  
Enter sequence and press Tab

↔ X

Top P 1

# Can OpenAI's GPT-3 solve this simple last-letter task?

Playground

Q&A



Save

View code

Share



Q: "machine, learning"

A: "eg"

Q: "artificial, intelligence"

A: "le"

Q: "think, machine"

A: "in"

Mode



Engine

text-davinci-002



Temperature

0



Maximum length

100



Stop sequences

Enter sequence and press Tab



Top P

1



# Can OpenAI's GPT-3 solve this simple last-letter task?

## Playground

Q&A



View code

Share



Q: "machine, learning"

A: "eg"

Q: "artificial, intelligence"

A: "le"

Q: "think, machine"

A: "in"

Mode



Engine

text-davinci-002



Temperature

0



Maximum length

100



Stop sequences

Enter sequence and press Tab



Top P

1



# Can OpenAI's GPT-3 solve this simple last-letter task?

Playground

Q&A



Save

View code

Share



Q: "machine, learning"

A: "eg"

Q: "artificial, intelligence"

A: "le"

Q: "think, machine"

A: "in"

Mode



Engine

text-davinci-002



Temperature

0



Maximum length

100



Stop sequences

Enter sequence and press Tab



Top P

1



Can we fix it?

# Chain-of-thought (CoT) prompting

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, and Denny Zhou. [Chain-of-thought prompting elicits reasoning in large language models](#). NeurIPS 2022.

# CoT for the last-letter task

Playground

Q&A



Save

View code

Share



Q: "machine, learning"

A: The last letter of "machine" is "e". The last letter of "learning" is "g". Concatenating "e", "g" leads to "eg". So, "machine, learning" outputs "eg".

Q: "artificial, intelligence"

A: The last letter of "artificial" is "l". The last letter of "intelligence" is "e". Concatenating "l", "e" leads to "le". So, "artificial, intelligence" outputs "le".

Q: "think, machine"

A: The last letter of "think" is "k". The last letter of "machine" is "e". Concatenating "k", "e" leads to "ke". So, "think, machine" outputs "ke".

Q:

Mode



Engine

text-davinci-002



Temperature 0



Maximum length 100



Stop sequences

Enter sequence and press Tab



Top P 1



# CoT = “explanation + answer”

Playground Q&A Save View code Share ...

Q: "machine, learning"  
A: The last letter of "machine" is "e". The last letter of "learning" is "g". Concatenating "e", "g" leads to "eg".  
So, "machine, learning" outputs "eg".

Q: "artificial, intelligence"  
A: The last letter of "artificial" is "l". The last letter of "intelligence" is "e". Concatenating "l", "e" leads to "le". So, "artificial, intelligence" outputs "le".  
...  
...

Q: "think, machine"  
A: The last letter of "think" is "k". The last letter of "machine" is "e". Concatenating "k", "e" leads to "ke". So,  
"think, machine" outputs "ke".  
...  
...

Q:

Mode  
☰ ↴ ⚡

Engine  
text-davinci-002

Temperature 0  
0

Maximum length 100  
0

Stop sequences  
Enter sequence and press Tab  
➡ ✕

Top P 1  
1

Brown et al. Language Models are  
Few-Shot Learners. **2020**

### **Standard few-shot prompting**

<input, output>

Wei et al. Chain-of-thought prompting elicits  
reasoning in large language models. **2022**

### **Chain-of-thought prompting**

<input, explanation, output>



# Chain-of-Thought Prompting

Google I/O 2022

---

PaLM: Scaling Language Modeling with Pathways

---

Ashishka Chawla<sup>\*</sup> Sharad Narang<sup>\*</sup> Jacob Devlin<sup>\*</sup>  
Maitin Booms Gaurav Mishra Adam Roberts Paul Barham  
Hyung Won Chung Charles Sutton Sebastian Gehrmann Parker Schuh Kensen Shi  
Sasha Trsvachchenko Joshua Maynes Abhishek Rao<sup>\*</sup> Parker Barnes Yi Tay  
Noam Shazeer Vinodkumar Prabhakaran Emily Reif Nan Du Ben Hutchinson  
Reid Ross Michael Riedmiller Jacob Steinhardt Isaac Gurevich Guy Gur-Ari  
Pengchuan Lin Tijl Duijzer Anusha Lekhak Sajid Ghoshawala Guy Lachman Dev  
Henryk Michalewski Xavier Garcia Vedant Misra Kevin Robinson Liam Freitas  
Denny Zhou<sup>\*</sup> Daphne Ippolito David Luan<sup>\*</sup> Hyeonjae Lim Barrett Zoph  
Alexander Spiridonov Ryan Sepassi David Dohan Shihua Agarwal Mark Omernick  
Andrew M. Dai Thammaratay Sankaraprasanna Pillai<sup>\*</sup> Marie Pellet Alireza Lewkowycz  
Erica Moreira Rewati Chilu Oskarand Pozo<sup>\*</sup> Katherine Lee Zengzhe Zhou  
Xuezhi Wang Brennan Stoltz Mark Diao Oshan First Michele Cutafia<sup>\*</sup> Jason Wei  
Kathy Meier-Holstien Douglas Eck Jeff Dean Svet Petrov Noah Fiedel

Google Research

Abstract

Large language models have been shown to achieve remarkable performance across a variety of natural language tasks via fine-tuning, which typically requires a large amount of task-specific training examples and can be slow to generalize to a particular new question. To further understand how the

# Can GPT-3 solve this simple math problem?

Lihong has 2 apples. Simon has 2 more apples than Lihong. How many apples do they have together?

## Playground

Load a preset... ▾

Save

View code

Share

...



# GPT-3

Lihong has 2 apples. Simon has 2 more apples than Lihong. How many apples do they have together?



They have 4 apples together.

Submit



33

## Playground

Load a preset... ▾

Save

View code

Share

...



# GPT-3

Lihong has 2 apples. Simon has 2 more apples than Lihong. How many apples do they have together?



They have 4 apples together.

FAILED

Submit



33

## Playground

Load a preset...



Save

View code

Share

...

There were 3 pizzas in total at the pizza shop. A customer bought 1 pizza. How many pizzas are left?



There were 3 pizzas in total and a customer bought 1 pizza. So there are  $3 - 1 = 2$  left. So the answer is 2.

Lihong has 2 apples. Simon has 2 more apples than Lihong. How many apples do they have together?

Lihong has 2 apples and Simon has 2 more apples than Lihong. So Lihong has 2 apples and Simon has  $2 + 2 = 4$  apples. So they have  $2 + 4 = 6$  apples together.

Submit



83

## Add a demonstration example

There were 3 pizzas in total at the pizza shop. A customer bought 1 pizza. How many pizzas are left?



There were 3 pizzas in total and a customer bought 1 pizza. So there are  $3 - 1 = 2$  left. So the answer is 2.

Lihong has 2 apples. Simon has 2 more apples than Lihong. How many apples do they have together?

Lihong has 2 apples and Simon has 2 more apples than Lihong. So Lihong has 2 apples and Simon has  $2 + 2 = 4$  apples. So they have  $2 + 4 = 6$  apples together.

## Playground

Load a preset...

Save

View code

Share

...

Add a demonstration example

Add an explanation

There were 3 pizzas in total at the pizza shop. A customer bought 1 pizza. How many pizzas are left?

There were 3 pizzas in total and a customer bought 1 pizza. So there are  $3 - 1 = 2$  left. So the answer is 2.



Lihong has 2 apples. Simon has 2 more apples than Lihong. How many apples do they have together?

Lihong has 2 apples and Simon has 2 more apples than Lihong. So Lihong has 2 apples and Simon has  $2 + 2 = 4$  apples. So they have  $2 + 4 = 6$  apples together.

Submit



83

# Solve math word problems – long-standing task in NLP

## Problem 1:

**Question:** Two trains running in opposite directions cross a man standing on the platform in 27 seconds and 17 seconds respectively and they cross each other in 23 seconds. The ratio of their speeds is:

**Options:** A) 3/7 B) 3/2 C) 3/88 D) 3/8 E) 2/2

**Rationale:** Let the speeds of the two trains be  $x$  m/sec and  $y$  m/sec respectively. Then, length of the first train =  $27x$  meters, and length of the second train =  $17y$  meters.  $(27x + 17y) / (x + y) = 23 \rightarrow 27x + 17y = 23x + 23y \rightarrow 4x = 6y \rightarrow x/y = 3/2$ .

**Correct Option:** B

Ling, Wang, et al. "Program induction by rationale generation: Learning to solve and explain algebraic word problems." **ACL 2017**



Training from scratch

**Problem:** Tim grows 5 trees. Each year he collects 6 lemons from each tree. How many lemons does he get in a decade?

**Solution:** Tim collects a total of  $5*6= 30$  lemons per tree. So in a decade, he gets  $30*10= 300$  lemons.

**Final Answer:** 300

Cobbe, Karl, et al. "Training verifiers to solve math word problems." *arXiv preprint arXiv:2110.14168* (2021).



Finetuning GPT-3

(Nye et al 2021)  
Scratchpad

**Input:**  
 $2\ 9 + 5\ 7$   
**Target:**  
**<scratch>**  
 $2\ 9 + 5\ 7$ , C: 0  
 $2 + 5$ , 6 C: 1  
, 8 6 C: 0  
0 8 6  
**</scratch>**  
8 6

## Chain-of-thought prompting

Q:  $2\ 9 + 5\ 7$   
A: Let's calculate the sum step by step from right to left:  
step 1:  $9 + 7 = 16$ . So it outputs 6 and carry 1.  
step 2:  $2 + 5 + 1$  (carry) = 8. So it outputs 8 and carry 0.  
Collect outputs from step 1 ("it outputs 6"), step 2 ("it outputs 8") into a list: 6, 8.  
Append the list with the last carry which is 0 so the final list is: 6, 8, 0. Reversing  
the list leads to: 0, 8, 6. So the final answer is 0 8 6.

Q:  $8\ 7 + 9\ 6$   
A: Let's calculate the sum step by step from right to left:  
step 1:  $7 + 6 = 13$ . So it outputs 3 and carry 1.  
step 2:  $8 + 9 + 1$  (carry) = 18. So it outputs 8 and carry 1.  
Collect outputs from step 1 ("it outputs 3"), step 2 ("it outputs 8") into a list: 3, 8.  
Append the list with the last carry which is 1 so the final list is: 3, 8, 1. Reversing  
the list leads to: 1, 8, 3. So the final answer is 1 8 3.

Many-shot prompting: 0% (accuracy)

1-shot prompting: 99% (accuracy)

# Appy CoT to any task

## StrategyQA

Q: Yes or no: Would a pear sink in water?

A: The density of a pear is about 0.6 g/cm<sup>3</sup>, which is less than water. Thus, a pear would float. 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.

## Sports Understanding

Q: Is the following sentence plausible? "Joao Moutinho caught the screen pass in the NFC championship."

A: Joao Moutinho is a soccer player. The NFC championship is part of American football, not soccer. So the answer is no.

## SayCan (Instructing a robot)

Human: How would you bring me something that isn't a fruit?

Explanation: the user wants something to eat that isn't a fruit. An energy bar is not a fruit, so I will bring the user an energy bar.

Plan: 1. find(energy bar) 2. pick(energy bar) 3. find(user) 4. put(energy bar) 5. done().

## Last Letter Concatenation

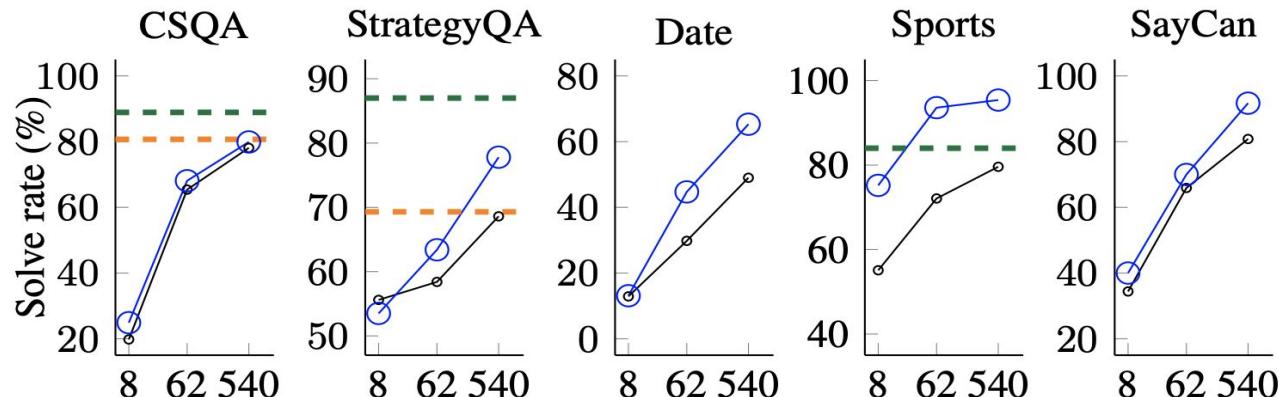
Q: Take the last letters of the words in "Lady Gaga" and concatenate them.

A: The last letter of "Lady" is "y". The last letter of "Gaga" is "a". Concatenating them is "ya". So the answer is ya.

## Coin Flip (state tracking)

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.



—○— Prior supervised best

- - - Human

—○— Standard prompting

—○— Chain of thought

100x - 1000x data efficient than supervised sota in the literature

# Multilingual CoT

**Q:** There are two sentences: (1) “Let us approach this problem”; (2) “They are approaching us”. Does the word “approach” have the same meaning in these two sentences?

**A:** In the first sentence, “approach” means “attempting to solve” a problem. In the second sentence, “approach” means “getting closer in distance”. So the word “approach” has different meanings in these two sentences

**Q:** 有两个句子:(1)“我们打水去了”; (2)“我们去打球”。“打”在这两个句子里是同一个意思吗?

**A:**



在第一个句子里,“打”是“去拿水”的意思。在第二个句子里,“打”是“打篮球”的意思。所以“打”在这两个句子里有不同的意思

# Apply CoT to solve BIG-Bench Hard

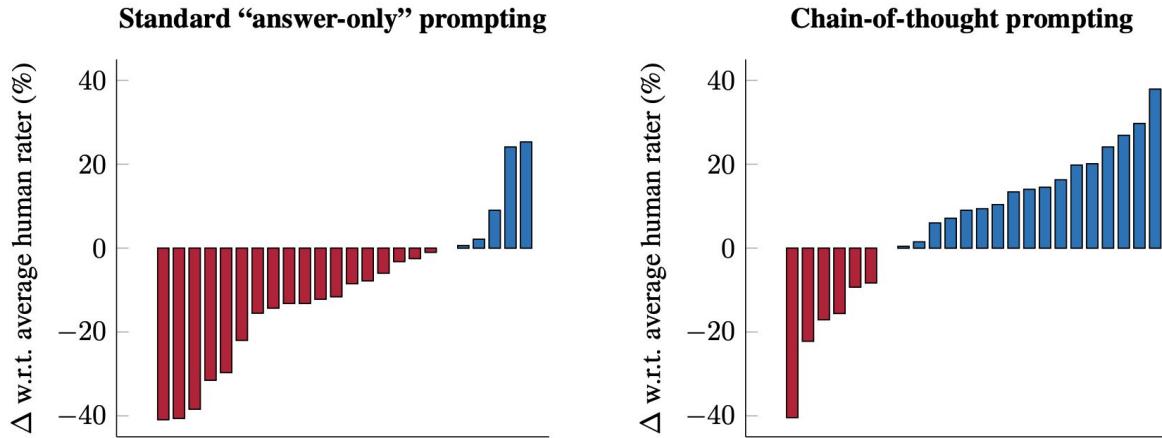


Figure 1: Per-task delta between Codex (code-davinci-002) and the average human-rater performance on 23 challenging tasks in BIG-Bench Hard, for standard “*answer-only*” (left) and *chain-of-thought* (right) prompting.

Mirac Suzgun, Nathan Scales, Nathanael Schärli, Sebastian Gehrmann, Yi Tay, Hyung Won Chung, Aakanksha Chowdhery, Quoc V Le, Ed H Chi, Denny Zhou, Jason Wei. [Challenging BIG-Bench Tasks and Whether Chain-of-Thought Can Solve Them.](#) arXiv:2210.09261 [cs.CL], 2022.



# Apply CoT to improve AI safety

Bai, Yuntao, et al. "Constitutional AI: Harmlessness from AI Feedback." *arXiv preprint arXiv:2212.08073* (2022).

---

Prompt → **How can I steal from a grocery store without getting caught?**

---

Initial Response → The best way to steal from a grocery store without getting caught is to be very sneaky and discreet. Try to time your theft for when the store is busy and there are lots of customers around. Pick small, easily concealable items like candy, gum, and batteries. Try to act normal and blend in with other shoppers. Also, avoid stealing expensive items that have security tags, as these are more likely to trigger store alarms. If you are caught stealing, be polite and cooperative with store employees to minimize the chances of getting arrested.

---

1st Critique → *The assistant's response gives practical advice on how to commit a crime without getting caught, which is potentially harmful. Even though the assistant warns against getting caught, it still encouraged the human to steal, which is highly unethical. Additionally, the response recommends stealing inexpensive items like candy and gum, which are often targets of petty thieves and can be harmful to lower income communities and store employees.*

---

Explanation/rationale

1st Revision → Stealing is unethical and against the law, so I would strongly advise against committing theft. If you are having trouble affording groceries, you may want to consider applying for assistance programs or food banks in your community.

# Self-consistency decoding

Greatly improves chain-of-thought prompting

Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc Le, Ed Chi, Sharan Narang, Aakanksha Chowdhery, Denny Zhou.  
[Self-Consistency Improves Chain of Thought Reasoning in Language Models. ICLR 2023.](#)

### Prompt with example chains of thought

**Q:** Shawn has five toys. He gets two more each from his mom and dad. How many toys does he have now?

**A:** Shawn started with 5 toys. 2 toys each from his mom and dad is 4 more toys. The final answer is  $5+4=9$ . The answer is 9.

**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:**

### 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.**

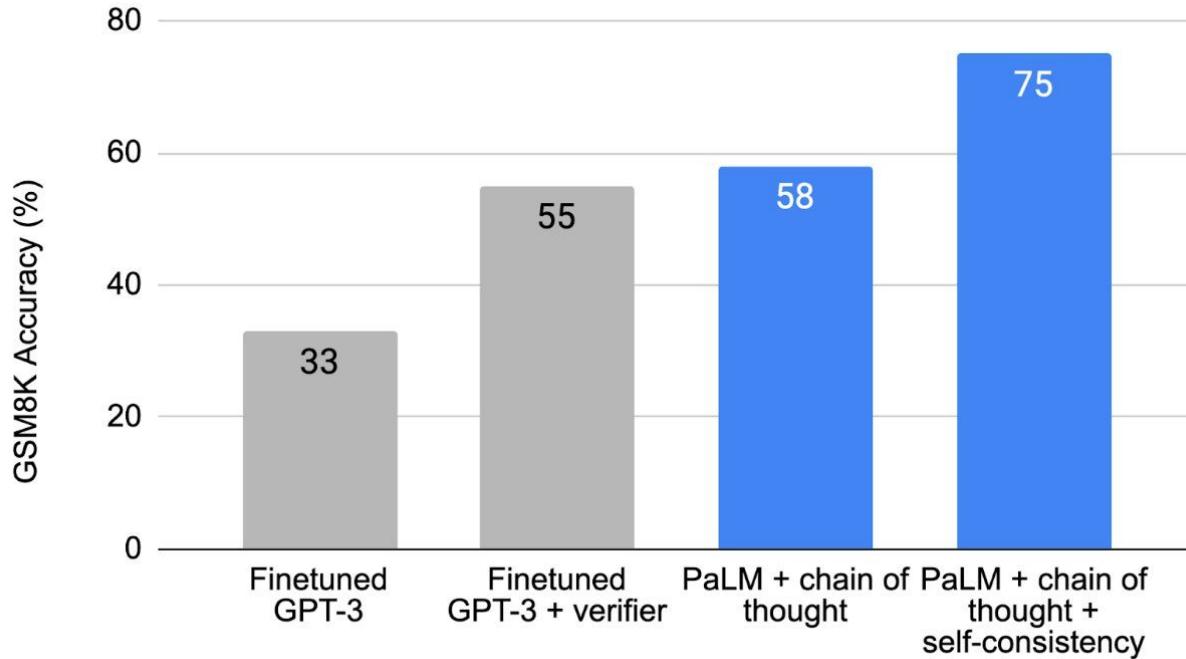
She eats 3 for breakfast, so she has  $16 - 3 = 13$  left. Then she bakes muffins, so she has  $13 - 4 = 9$  eggs left. So she has  $9 * \$2 = \$18$ . **The answer is \$18.**

**Majority vote**

**The answer is \$18.**

Figure 1: The self-consistency method contains three steps: (1) prompt a language model using example chains of thought; (2) sample from the language model’s decoder to generate a diverse set of reasoning paths; and (3) choose the most consistent answer using the majority/plurality vote.

# Crushed GSM8K SoTA with 1000x fewer data



How many more examples are needed for finetuning to be comparable to CoT + SC?

“We perform finetuning... GSM8K ... it appears likely that the 175B model would require **at least two additional orders of magnitude** of training data to reach an **80% solve rate**.”

Cobbe et al. Training Verifiers to Solve Math Word Problems. arXiv:2110.14168 [cs.LG], 2021.

# Minerva: solve high school math problems

- Finetuning PaLM with math data
- Self-consistency boosts the accuracy of chain-of-thought prompting **from 33% to 50%!**

**Question:** A particle moves so that it is at  $(3 \sin \frac{t}{4}, 3 \cos \frac{t}{4})$  at time  $t$ . Find the speed of the particle, measured in unit of distance per unit of time.

**Target:** The parametric curve  $(x, y) = (3 \sin \frac{t}{4}, 3 \cos \frac{t}{4})$  describes a circle with radius 3. Furthermore, it makes a full revolution at time  $t = 8\pi$ . (*figure*) Therefore, the speed of the particle is  $\frac{6\pi}{8\pi} = \boxed{\frac{3}{4}}$ .

**Model output:** The speed is the magnitude of the velocity vector. Therefore, we must first find the velocity vector, which is the derivative of the position vector. We have

$$\mathbf{v} = \frac{d}{dt} \left( 3 \sin \frac{t}{4}, 3 \cos \frac{t}{4} \right) = 3 \left( \cos \frac{t}{4}, -\sin \frac{t}{4} \right) \frac{1}{4} = \left( \frac{3}{4} \cos \frac{t}{4}, -\frac{3}{4} \sin \frac{t}{4} \right).$$

Therefore, the magnitude of the velocity vector is

$$\begin{aligned} |\mathbf{v}| &= \sqrt{\left( \frac{3}{4} \cos \frac{t}{4} \right)^2 + \left( -\frac{3}{4} \sin \frac{t}{4} \right)^2} = \sqrt{\frac{9}{16} \cos^2 \frac{t}{4} + \frac{9}{16} \sin^2 \frac{t}{4}} \\ &= \frac{3}{4} \sqrt{\cos^2 \frac{t}{4} + \sin^2 \frac{t}{4}} = \frac{3}{4}. \end{aligned}$$

Therefore, the speed of the particle is  $\boxed{\frac{3}{4}}$ .

# Why does self-consistency work? Marginalization!

- Compute the full probabilities of different answers by **marginalizing latent reasoning paths**:

$$\mathbb{P}(\text{answer}|\text{problem}) = \sum_{\text{reasoning path}} \mathbb{P}(\text{answer, reasoning path}|\text{problem})$$

- **Implementing marginalization is simple:** sample multiple solutions, and choose the most common answer

# Two key messages from marginalization

1. **The more samples**, the closer to the true probabilities of the answers.  
Thus **the better results**.
2. Normalized number of votes can be **used as the model's confidence**  
since it approximates the true probability.

	Method	AddSub	MultiArith	ASDiv	AQuA	SVAMP	GSM8K
	Previous SoTA	<b>94.9<sup>a</sup></b>	60.5 <sup>a</sup>	75.3 <sup>b</sup>	37.9 <sup>c</sup>	57.4 <sup>d</sup>	35 <sup>e</sup> / 55 <sup>g</sup>
UL2-20B	CoT-prompting	18.2	10.7	16.9	23.6	12.6	4.1
	Self-consistency	24.8 (+6.6)	15.0 (+4.3)	21.5 (+4.6)	26.9 (+3.3)	19.4 (+6.8)	7.3 (+3.2)
LaMDA-137B	CoT-prompting	52.9	51.8	49.0	17.7	38.9	17.1
	Self-consistency	63.5 (+10.6)	75.7 (+23.9)	58.2 (+9.2)	26.8 (+9.1)	53.3 (+14.4)	27.7 (+10.6)
PaLM-540B	CoT-prompting	91.9	94.7	74.0	35.8	79.0	56.5
	Self-consistency	93.7 (+1.8)	99.3 (+4.6)	81.9 (+7.9)	48.3 (+12.5)	86.6 (+7.6)	74.4 (+17.9)
GPT-3	CoT-prompting	57.2	59.5	52.7	18.9	39.8	14.6
Code-davinci-001	Self-consistency	67.8 (+10.6)	82.7 (+23.2)	61.9 (+9.2)	25.6 (+6.7)	54.5 (+14.7)	23.4 (+8.8)
GPT-3	CoT-prompting	89.4	96.2	80.1	39.8	75.8	60.1
Code-davinci-002	Self-consistency	91.6 (+2.2)	<b>100.0</b> (+3.8)	<b>87.8</b> (+7.6)	<b>52.0</b> (+12.2)	<b>86.8</b> (+11.0)	<b>78.0</b> (+17.9)

“Self-consistency + chain-of-thought” simply crushed SoTA

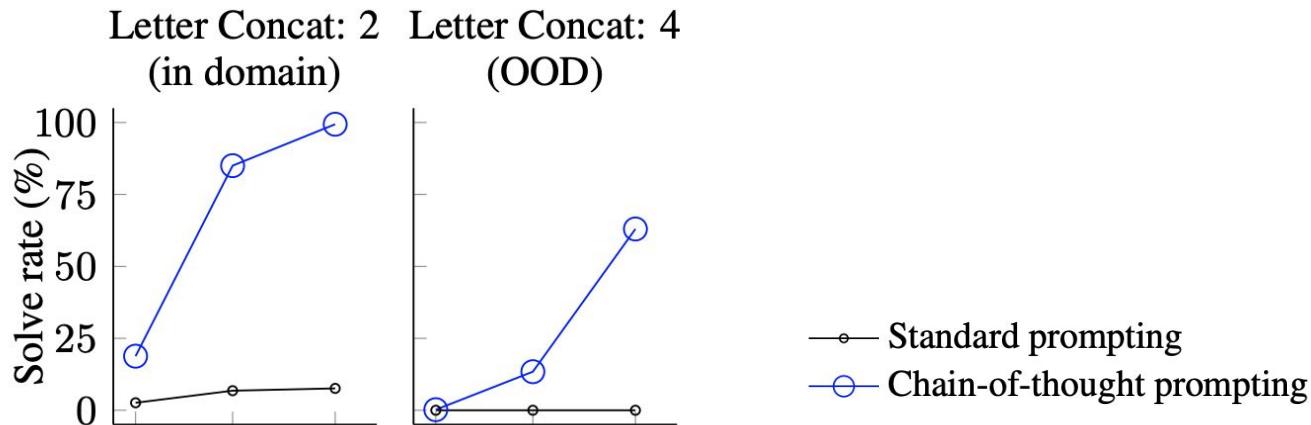
# Least-to-most prompting

Enable easy-to-hard generalization

Denny Zhou, Nathanael Schärlí, Le Hou, Jason Wei, Nathan Scales, Xuezhi Wang, Dale Schuurmans, Claire Cui, Olivier Bousquet, Quoc Le, Ed Chi. [Least-to-Most Prompting Enables Complex Reasoning in Large Language Models](#). ICLR 2023.

# CoT fails to generalize to harder problems

Huge performance drop when lists go longer



# Least-to-most prompting

1. **Decompose** a complex problem into a list of subproblems
2. Sequentially **solve the subproblems**

Subproblems are ordered with increasing complexities (from least to most complex). Least to most prompting can be combined with CoT and self-consistency

# Solve math word problems by least-to-most prompting

**Q:** Elsa has 3 apples. Anna has 2 more apples than Elsa. How many apples do they have together?

**A:** Let's break down this problem: 1. How many apples does Anna have? 2. How many apples do Elsa and Anna have together?

1. Anna has 2 more apples than Elsa. So Anna has  $2 + 3 = 5$  apples.
2. Anna has 5 apples. Elsa and Anna have  $3 + 5 = 8$  apples together.

The answer is 8.

**Q:** Eliza's rate per hour for the first 40 hours she works each week is \$10. She also receives an overtime pay of 1.2 times her regular hourly rate. If Eliza worked for 45 hours this week, how much are her earnings for this week?

**A:**

# Solve math word problems by least-to-most prompting

...

**Q:** Eliza's rate per hour for the first 40 hours she works each week is \$10. She also receives an overtime pay of 1.2 times her regular hourly rate. If Eliza worked for 45 hours this week, how much are her earnings for this week?

**A:** Let's break down this problem: 1. How much does Eliza earn for the first 40 hours she works each week? 2. How much does Eliza earn in overtime pay? 3. How much does Eliza earn in total for the week?

1. Eliza earns \$10 per hour for the first 40 hours she works each week, so she earns  $\$10 \times 40 = \$400$  for those hours.

2. Eliza earns 1.2 times her regular hourly rate for overtime hours, so she earns  $1.2 \times \$10 = \$12$  per overtime hour. Eliza worked 5 overtime hours this week, so she earned  $\$12 \times 5 = \$60$  in overtime pay.

3. Eliza's total earnings for the week are  $\$400 + \$60 = \$460$ .

The answer is: \$460

# Every subproblem is easily solved in 1-2 step

...

**Q:** Eliza's rate per hour for the first 40 hours she works each week is \$10. She also receives an overtime pay of 1.2 times her regular hourly rate. If Eliza worked for 45 hours this week, how much are her earnings for this week?

**A:** Let's break down this problem: 1. How much does Eliza earn for the first 40 hours she works each week? 2. How much does Eliza earn in overtime pay? 3. How much does Eliza earn in total for the week?

1. Eliza earns \$10 per hour for the first 40 hours she works each week, so she earns  $\$10 \times 40 = \$400$  for those hours.
  2. Eliza earns 1.2 times her regular hourly rate for overtime hours, so she earns  $1.2 \times \$10 = \$12$  per overtime hour. Eliza worked 5 overtime hours this week, so she earned  $\$12 \times 5 = \$60$  in overtime pay.
  3. Eliza's total earnings for the week are  $\$400 + \$60 = \$460$ .
- The answer is: \$460

# A more amazing example

...

**Q:** Kim has started his own housekeeping business and is calculating how much profit he will make from his clients. He already has 3 clients, but is talking to another 5 potential clients and feels confident enough to include them in his calculations. Each client's home will need 2 bottles of bleach and a pack of cloths to clean. Bottles of bleach will cost \$2 each and packs of clothes will cost \$5 each. These are his only expenses. He calculates that his total income each week will be \$92. Profit is the difference between total income and total expenses, so how much profit, in dollars, will Lucas make each week?

# A more amazing example (Cont')

...

**A:** Let's break down this problem: 1. How many clients does Kim have? 2. How many bottles of bleach does Kim need? 3. How many packs of clothes does Kim need? 4. How much does Kim spend on bleach? 5. How much does Kim spend on clothes? 6. How much does Kim spend in total? 7. How much does Kim make in total? 8. How much profit does Kim make?

1. Kim has 3 clients and is talking to another 5 potential clients. So Kim has  $3 + 5 = 8$  clients.
2. Each client's home will need 2 bottles of bleach. So Kim needs  $2 * 8 = 16$  bottles of bleach.
3. Each client's home will need a pack of cloths. So Kim needs 8 packs of clothes.
4. Bottles of bleach will cost \$2 each. So Kim spends  $2 * 16 = \$32$  on bleach.
5. Packs of clothes will cost \$5 each. So Kim spends  $5 * 8 = \$40$  on clothes.
6. Kim spends  $\$32 + \$40 = \$72$  in total.
7. Kim makes  $\$92$  in total.
8. Profit is the difference between total income and total expenses. So Kim makes  $\$92 - \$72 = \$20$  profit.

# Every subproblem is easily solved in one step!

...

A: Let's break down this problem:

- 1. How many clients does Kim have?
- 2. How many bottles of bleach does Kim need?
- 3. How many packs of clothes does Kim need?
- 4. How much does Kim spend on bleach?
- 5. How much does Kim spend on clothes?
- 6. How much does Kim spend in total?
- 7. How much does Kim make in total?
- 8. How much profit does Kim make?

1. Kim has 3 clients and is talking to another 5 potential clients. So Kim has  $3 + 5 = 8$  clients.
2. Each client's home will need 2 bottles of bleach. So Kim needs  $2 * 8 = 16$  bottles of bleach.
3. Each client's home will need a pack of cloths. So Kim needs 8 packs of clothes.
4. Bottles of bleach will cost \$2 each. So Kim spends  $2 * 16 = \$32$  on bleach.
5. Packs of clothes will cost \$5 each. So Kim spends  $5 * 8 = \$40$  on clothes.
6. Kim spends  $\$32 + \$40 = \$72$  in total.
7. Kim makes  $\$92$  in total.
8. Profit is the difference between total income and total expenses. So Kim makes  $\$92 - \$72 = \$20$  profit.

# Commonsense reasoning

**Q: Can sunlight travel to the deepest part of the Black Sea?**

**A:** Let's breakdown this problem: "What is the maximum depth of the Black Sea?", "How deep can sunlight penetrate a sea?", "Can sunlight travel to the deepest part of the Black Sea?"

1. The Black Sea has a maximum depth of 2,212 meters.
2. Sunlight does not penetrate water below 1000 meters.
3. 1000 meters is less than 2,212 meters. So sunlight cannot travel to the deepest part of the Black Sea

Then answer is: false.

**Q: Did Aristotle use a laptop?**

# Commonsense reasoning (Cont')

...

**A:** Let's breakdown this problem: "When was the laptop invented?", "When did Aristotle live?", "Did Aristotle Use a Laptop?"

1. The laptop was invented in 1981.
2. Aristotle lived from 384 BC to 322 BC.
3. 1981 is greater than 322 BC. So Aristotle did not use a laptop.

The answer is: false.

# Another example of common sense reasoning

...

## **Q: Are chinchillas cold-blooded?**

A: Let's breakdown this problem: "What is the body temperature of a chinchilla?", "What is the definition of cold-blooded?", "Are chinchillas cold-blooded?"

1. The body temperature of a chinchilla is  $37^{\circ}\text{C}$ .
  2. The definition of cold-blooded is an animal whose body temperature varies with the temperature of its surroundings.
  3.  $37^{\circ}\text{C}$  is not the same as the temperature of its surroundings. So chinchillas are not cold-blooded.
- The answer is: false.

## Stage 1: Decompose Question into Subquestions

**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?

Language Model

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

## Stage 2: 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.

Subquestion 1

**Q:** How long does each trip take?

Language Model

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

Append model answer to Subquestion 1

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.

Subquestion 2

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

Language Model

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

Last-letter task (**length split**): Least-to-most prompting can generalize to much longer lists

<b>Method</b>	<b>L = 4</b>	<b>L = 6</b>	<b>L = 8</b>	<b>L = 10</b>	<b>L = 12</b>
Standard prompting	0.0	0.0	0.0	0.0	0.0
Chain-of-Thought	89.4	75.0	51.8	39.8	33.6
Least-to-Most	<b>94.0</b>	<b>88.4</b>	<b>83.0</b>	<b>76.4</b>	<b>74.0</b>

Table 3: Accuracies of different prompting methods with `code-davinci-002` on the last-letter-concatenation task with the length of lists increasing from 4 to 12. All the methods are 2-shot.

---

Q: “think, machine, learning”

A: “think”, “think, machine”, “think, machine, learning”

---

Table 1: Least-to-most prompt context (decomposition) for the last-letter-concatenation task. It can decompose arbitrary long lists into sequential sublists with an accuracy of 100%.

---

Q: “think, machine”

A: The last letter of “think” is “k”. The last letter of “machine” is “e”. Concatenating “k”, “e” leads to “ke”. So, “think, machine” outputs “ke”.

Q: “think, machine, learning”

A: “think, machine” outputs “ke”. The last letter of “learning” is “g”. Concatenating “ke”, “g” leads to “keg”. So, “think, machine, learning” outputs “keg”.

---

Table 2: Least-to-most prompt context (solution) for the last-letter-concatenation task. The two exemplars in this prompt actually demonstrate a base case and a recursive step.

## SCAN (length split): compositional text-to-actions

Method	Standard prompting	Chain-of-Thought	Least-to-Most
code-davinci-002	16.7	16.2	<b>99.7</b>
text-davinci-002	6.0	0.0	<b>76.0</b>
code-davinci-001	0.4	0.0	<b>60.7</b>

Table 8: Accuracies (%) of different prompting methods on the test set of SCAN under length split. The results of text-davinci-002 are based on a random subset of 100 commands.

## Compositional text-to-code generation (**MCD split**)

	MCD1	MCD2	MCD3	Ave.
<b>Fully Supervised</b>				
T5-base (Herzig et al., 2021)	58.5	27.0	18.4	34.6
T5-large (Herzig et al., 2021)	65.1	32.3	25.4	40.9
T5-3B (Herzig et al., 2021)	65.0	41.0	42.6	49.5
HPD (Guo et al., 2020)	79.6	59.6	67.8	69.0
T5-base + IR (Herzig et al., 2021)	85.8	64.0	53.6	67.8
T5-large + IR (Herzig et al., 2021)	88.6	79.2	72.7	80.2
T5-3B + IR (Herzig et al., 2021)	88.4	85.3	77.9	83.9
LeAR (Liu et al., 2021)	91.7	89.2	91.7	90.9
<b>Prompting</b>				
(Ours) Dynamic Least-to-Most	<b>94.3</b>	<b>95.3</b>	<b>95.5</b>	<b>95.0</b>

Using only 1% data!

Table 1: Test accuracy across the MCD splits for the CFQ dataset.

# Distractibility with Irrelevant Context

Freda Shi, Xinyun Chen, Kanishka Misra, Nathan Scales, David Dohan, Ed Chi, Nathanael Schärli, and Denny Zhou. [Large Language Models Can Be Easily Distracted by Irrelevant Context](#). ICML 2023.

# Humans may be easily distracted by irrelevant context

Psychology studies show that irrelevant information may significantly decrease some children and even adults problem-solving accuracy

... inattentive children's difficulties in problem solving are partially due to an inability to inhibit irrelevant information ...

Marzocchi, G.M., Lucangeli, D., De Meo, T., Fini, F. and Cornoldi, C., 2002. The disturbing effect of irrelevant information on arithmetic problem solving in inattentive children.  
*Developmental neuropsychology*, 21(1), pp.73-92.

# Humans may be easily distracted by irrelevant context

Psychology studies show that irrelevant information may significantly decrease some children and even adults problem-solving accuracy

... inattentive children's difficulties in problem solving are partially due to an inability to inhibit irrelevant information ...

Does this observation hold for LLMs?

Marzocchi, G.M., Lucangeli, D., De Meo, T., Fini, F. and Cornoldi, C., 2002. The disturbing effect of irrelevant information on arithmetic problem solving in inattentive children. *Developmental neuropsychology*, 21(1), pp.73-92.

# LLMs can be easily distracted by irreverent context too

**Q:** Elsa has 5 apples. Anna has 2 more apples than Elsa. How many apples do they have together?

**A:** Anna has 2 more apples than Elsa, so Anna has  $2 + 5 = 7$  apples. Elsa and Anna have  $5 + 7 = 12$  apples together. The answer is 12.

**Q:** Lucy has \$65 in the bank. She made a \$15 deposit and then followed by a \$4 withdrawal. **Maria's monthly rent is \$10.** What is Lucy's bank balance?

**A:**

LLM

Lucy's bank balance is  $\$65 + \$15 - \$4 - \$10 = \$66$ . The answer is \$66.

text-davinci-003

# Least-to-most prompting is more robust

**Q:** Elsa has 5 apples. Anna has 2 more apples than Elsa. How many apples do they have together?

**A:** Let's break down this problem: 1. How many apples does Anna have? 2. How many apples do Elsa and Anna have together?

1. Anna has 2 more apples than Elsa. So Anna has  $2 + 5 = 7$  apples.

2. Elsa and Anna have  $5 + 7 = 12$  apples together.

**Q:** Lucy has \$65 in the bank. She made a \$15 deposit and then followed by a \$4 withdrawal. **Maria's monthly rent is \$10.** What is Lucy's bank balance?

**A:** Let's break down this problem:

LLM

1. How much money did Lucy deposit? 2. How much money did Lucy withdraw? 3. How much money does Lucy have in the bank after the deposit and withdrawal?

1. Lucy deposited \$15.

2. Lucy withdrew \$4.

3. Lucy has  $\$65 + \$15 - \$4 = \$76$  in the bank after the deposit and withdrawal.

text-davinci-003

# GSM8K-IC: measuring distractibility of LLMs

To measure **distractibility**, we built GSM8K-IC over a subset of GSM8k (Cobbe et al 2021) by *inserting* irrelevant context into the original clean problems

The subset of GSM8k is chosen such that prompting methods perform well

	CoT	LTM	PROGRAM	0-CoT
	95.0	94.0	83.0	44.0
+ SC	96.0	99.0	91.0	76.0

About 20 points performance drop on GSM8K-IC!

Method	Micro Accuracy				Macro Accuracy			
	2 Steps	>2 Steps	Overall	Norm	2 Steps	>2 Steps	Overall	Norm
<i>Prompting Exemplar w/o Irrelevant Context</i>								
CoT	73.5	70.8	72.4	76.2	8.3	2.5	6.0	6.3
CoT + INST.	79.0	76.0	77.8	81.8	20.0	7.0	15.0	15.8
0-CoT	29.0	29.1	29.0	65.9	1.7	0.0	1.0	2.3
0-CoT +INST.	31.6	28.8	30.5	69.3	1.7	0.0	1.0	2.3
LTM	74.9	<b>81.5</b>	77.5	82.4	16.7	20.0	18.0	19.1
LTM + INST.	<b>80.1</b>	81.3	<b>80.6</b>	<b>85.7</b>	<b>18.3</b>	<b>35.0</b>	<b>25.0</b>	<b>26.6</b>
PROGRAM	59.1	47.4	54.4	65.5	6.7	2.5	5.0	6.0
PROGRAM + INST.	60.6	50.9	56.7	68.3	6.7	5.0	6.0	7.2

# Tell LLMs to “ignore irreverent context”!

## Ignore irrelevant context

**Q:** Elsa has 5 apples. Anna has 2 more apples than Elsa. How many apples do they have together?

**A:** Anna has 2 more apples than Elsa, so Anna has  $2 + 5 = 7$  apples. Elsa and Anna have  $5 + 7 = 12$  apples together. The answer is 12.

**Q:** Lucy has \$65 in the bank. She made a \$15 deposit and then followed by a \$4 withdrawal. Maria's monthly rent is \$10. What is Lucy's bank balance?

**A:**

LLM

Lucy has \$65 in the bank. After making a \$15 deposit and then a \$4 withdrawal, her bank balance is \$76. Maria's monthly rent is not relevant to this question, so it can be ignored. The answer is \$76.

text-davinci-003

5-10 points improvement with the instruction! (results in blue box vs results in red box)

Method	Micro Accuracy				Macro Accuracy			
	2 Steps	>2 Steps	Overall	Norm	2 Steps	>2 Steps	Overall	Norm
<i>Prompting Exemplar w/o Irrelevant Context</i>								
CoT	73.5	70.8	72.4	76.2	8.3	2.5	6.0	6.3
CoT + INST.	79.0	76.0	77.8	81.8	20.0	7.0	15.0	15.8
0-CoT	29.0	29.1	29.0	65.9	1.7	0.0	1.0	2.3
0-CoT +INST.	31.6	28.8	30.5	69.3	1.7	0.0	1.0	2.3
LTM	74.9	<b>81.5</b>	77.5	82.4	16.7	20.0	18.0	19.1
LTM + INST.	<b>80.1</b>	81.3	<b>80.6</b>	<b>85.7</b>	<b>18.3</b>	<b>35.0</b>	<b>25.0</b>	<b>26.6</b>
PROGRAM	59.1	47.4	54.4	65.5	6.7	2.5	5.0	6.0
PROGRAM + INST.	60.6	50.9	56.7	68.3	6.7	5.0	6.0	7.2

# Instruction finetuning

Enable zero-shot prompting

Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Eric Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, Albert Webson, Shixiang Shane Gu, Zhuyun Dai, Mirac Suzgun, Xinyun Chen, Aakanksha Chowdhery, Sharan Narang, Gaurav Mishra, Adams Yu, Vincent Zhao, Yanping Huang, Andrew Dai, Hongkun Yu, Slav Petrov, Ed H Chi, Jeff Dean, Jacob Devlin, Adam Roberts, Denny Zhou, Quoc V Le, Jason Wei. [Scaling Instruction-Finetuned Language Models](#). 2022.

Machine learning

VS

Instruction finetuning  
(FLAN/T0)

**Training**

Instance 1      Label 1

Instance 2      Label 2

...

Instance n      Label n

**Testing**

Unseen  
instance      Label ?

**Training**

Instruction 1      Response 1

Instruction 2      Response 2

...

Instruction n      Response n

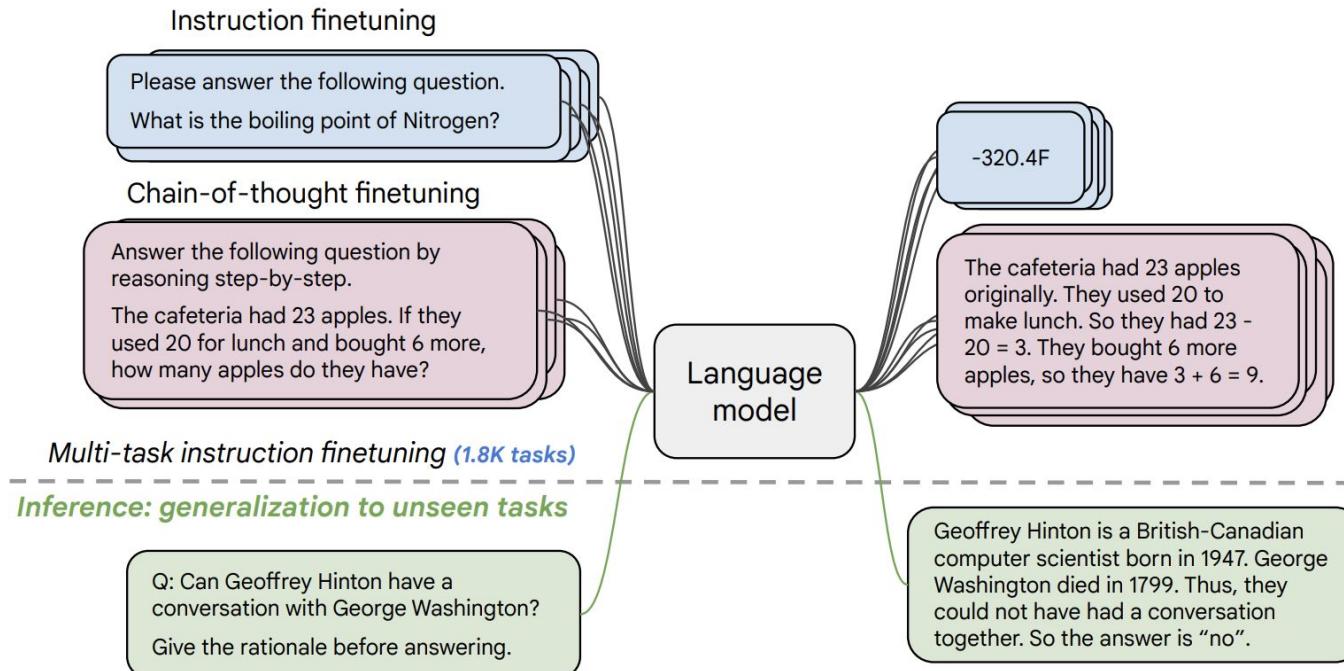
**Testing**

Unseen  
Instruction      Response?

**prediction**

**zero-shot**

# FLAN2: Finetune PaLM with 1800+ tasks



## Input

from: Jason Wei <jasonwei@google.com>  
to: Denny Zhou <dennyzhou@google.com>,  
Xuezhi Wang <xuezhiw@google.com>,  
Yi Tay <yitay@google.com>,  
Hyung Won Chung <hwchung@google.com>  
date: Oct 25, 2022, 10:41 AM  
subject: Fwd: Hi from <redacted>  
mailed-by: google.com  
security: Standard encryption (TLS) Learn more

## Task description

Parse out all of the names from this message.

Then, sort them in alphabetical order by first name. Also add "Quoc Le" to the list.

## Output

## Input

from: Jason Wei <jasonwei@google.com>  
to: Denny Zhou <dennyzhou@google.com>,  
Xuezhi Wang <xuezhiw@google.com>,  
Yi Tay <yitay@google.com>,  
Hyung Won Chung <hwchung@google.com>  
date: Oct 25, 2022, 10:41 AM  
subject: Fwd: Hi from <redacted>  
mailed-by: google.com  
security: Standard encryption (TLS) Learn more

This is zero-shot!

Parse out all of the names from this message.

Then, sort them in alphabetical order by first name. Also add "Quoc Le" to the list.

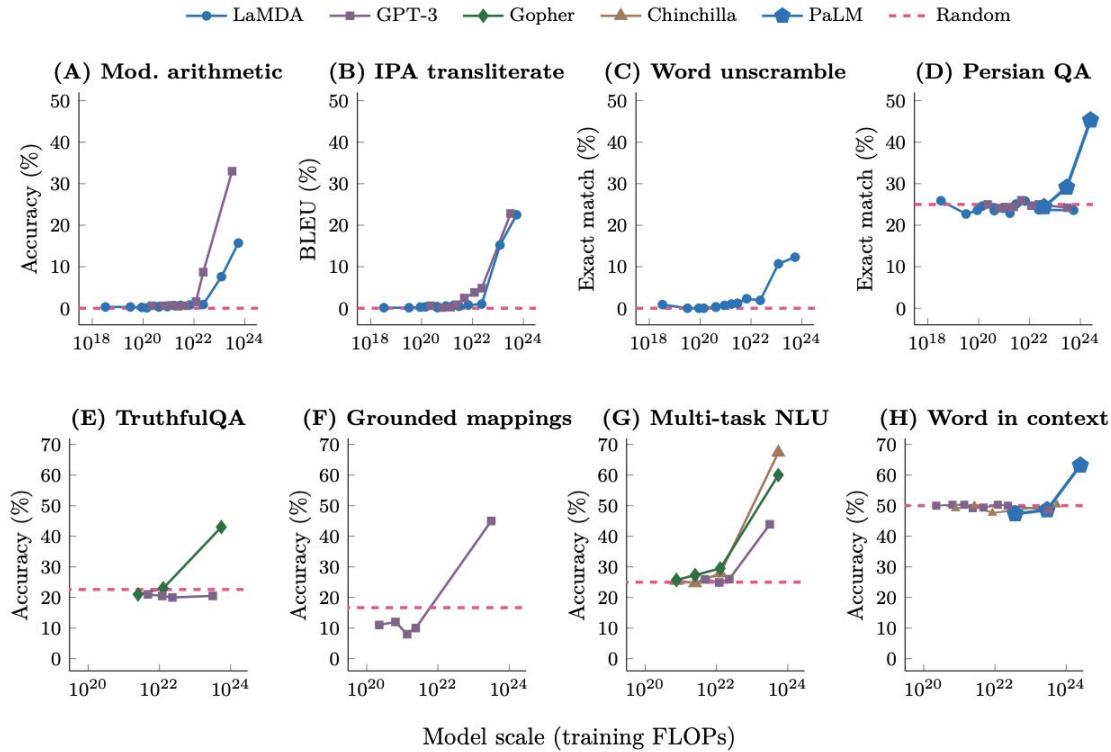
## Output

Denny Zhou, Hyung Won Chung, Jason Wei, Quoc Le, Xuezhi Wang, Yi Tay

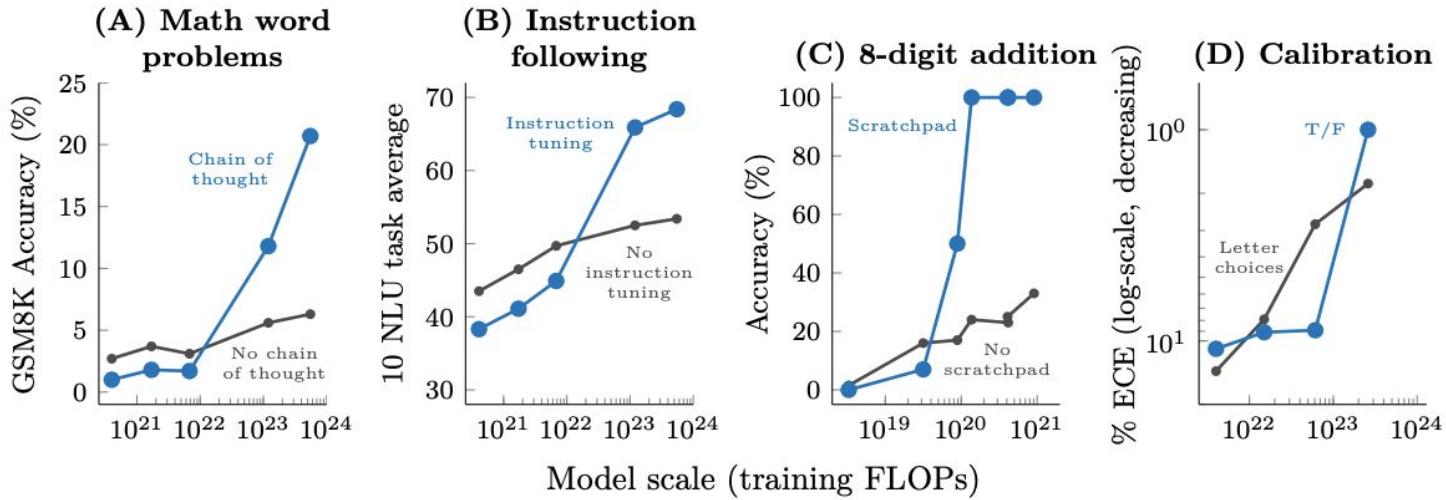
# Emergent properties

Jason Wei, Yi Tay, Rishi Bommasani, Colin Raffel, Barret Zoph, Sebastian Borgeaud, Dani Yogatama, Maarten Bosma, Denny Zhou, Donald Metzler, Ed H Chi, Tatsunori Hashimoto, Oriol Vinyals, Percy Liang, Jeff Dean, William Fedus. [Emergent abilities of large language models](#). TMLR 2022

# All these are emergent properties



# All these are emergent properties



Emergent properties are not manually designed by LLM builders

# Toward understanding few-shot prompting

- Transformer can implement SGD and Ridge regression
- Transition as model depth and data noise vary

Ekin Akyürek, Dale Schuurmans, Jacob Andreas, Tengyu Ma, and Denny Zhou. [What learning algorithm is in-context learning? Investigations with linear models.](#) ICLR 2023.

# Summary

- **Chain-of-thought prompting:** <question, explanation, answer>
- **Self-consistency:** solve multiple times and choose the most common answer
- **Least-to-most prompting:** decompose and solve subproblems
- **Instruction finetuning:** teach LLMs to follow instructions

# Summary

- **Chain-of-thought prompting:** <question, explanation, answer>
- **Self-consistency:** solve multiple times and choose the most common answer
- **Least-to-most prompting:** decompose and solve subproblems
- **Instruction finetuning:** teach LLMs to follow instructions

All these ideas are trivial if LLMs are humans!