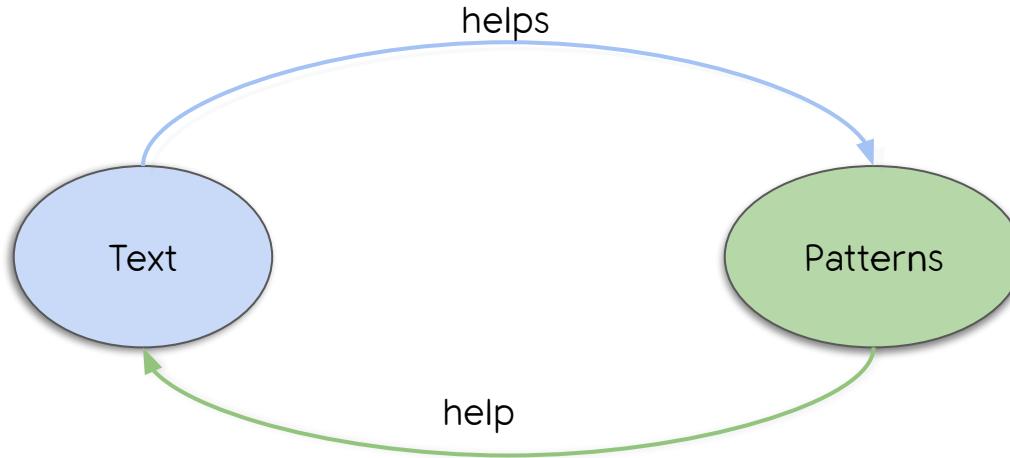


Text and Patterns: For Effective Chain of Thought, It Takes Two to Tango



Aman Madaan* and **Amir Yazdanbakhsh***
Carnegie Mellon University *Google Research, Brain Team
amadaan@cs.cmu.edu, ayazdan@google.com
(Equal Contribution)

[Paper](#)

Background

Solving middle-school math problems

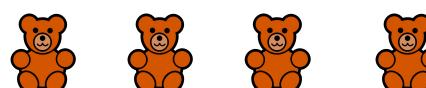
Q: *Shawn has 5 toys. For Christmas, he got 2 toys each from his mom and dad. How many toys does he have now?*



+



=

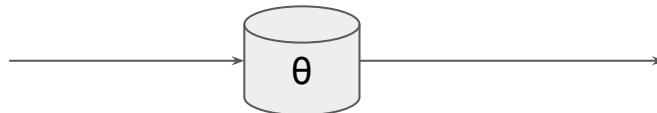


A: The answer is 9 toys



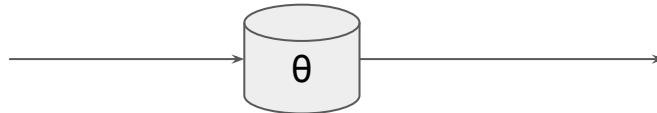
Fine-tuning

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



A: The answer is 5 cars.

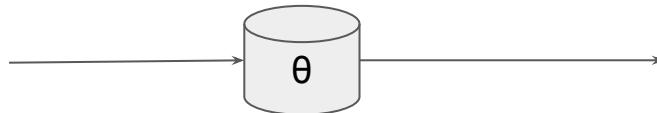
Q: Leah had 32 chocolates and her sister had 42. If they ate 35, how many pieces do they have left in total?



A: The answer is 39 pieces.

Train/Fine-tune

Q: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?



A: The answer is 9 toys

Test

Few-shot prompting (in-context learning/autocomplete)

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

A: The answer is 5 cars.

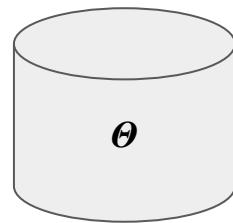
Q: Leah had 32 chocolates and her sister had 42. If they ate 35, how many pieces do they have left in total?

A: The answer is 39 pieces.

Q: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?

A:

Prompt



The answer is 9 toys

Design of prompt
(prompt engineering) is critical

Chain of thought prompting

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Ed Chi, Quoc Le, and Denny Zhou. "Chain of thought prompting elicits reasoning in large language models." *arXiv preprint arXiv:2201.11903* (2022).

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

Thought (T): There are originally 3 cars. 2 more cars arrive. $3 + 2 = 5$.

A: The answer is 5 cars.

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

Thought (T): Originally, Leah had 32 chocolates. Her sister had 42. So in total they had $32 + 42 = 74$. After eating 35, they had $74 - 35 = 39$.

A: The answer is 39 pieces.

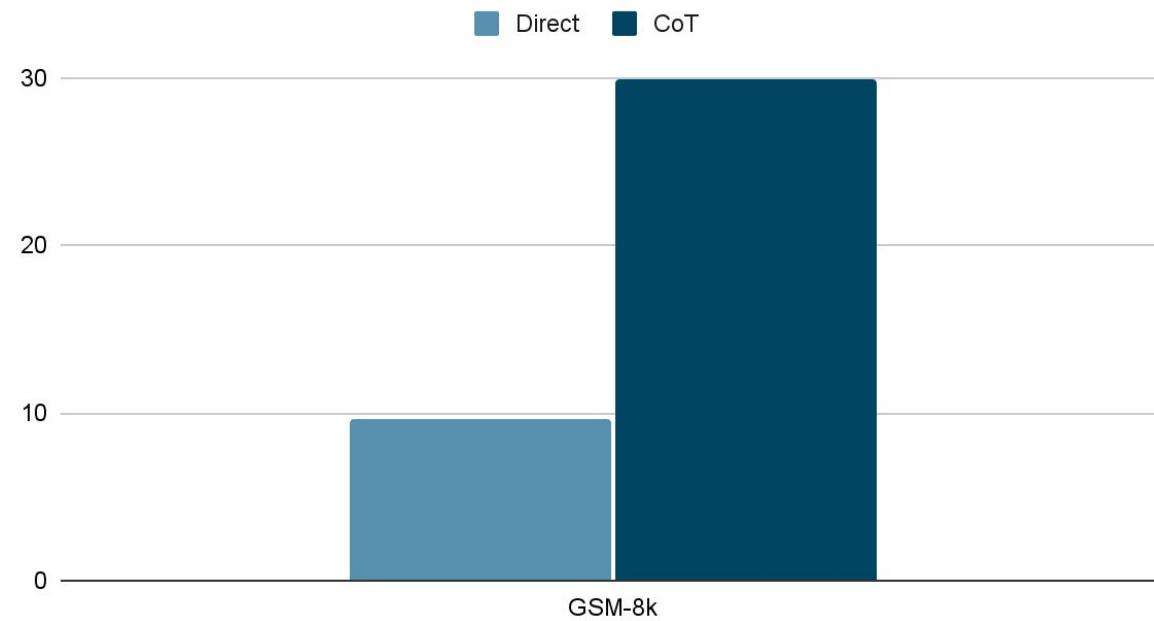
Q: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?

T:

Adds a thought to the prompt that explains the answer - *the thought process*.

Chain of thought prompting is extremely effective

PaLM 62B



How PaLM-SayCan Works

Large Language Models are Zero-Shot Reasoners

Takeshi Kojima
The University of Tokyo
t.kojima@weblab.t.u-tokyo.ac.jp

Shixiang Shane Gu
Google Research, Brain Team

Machel Reid
Google Research*

Yutaka Matsuo
The University of Tokyo

Yusuke Iwasawa
The University of Tokyo

This demo shows a PaLM-enabled helper robot performing a series of complex tasks using chain of thought prompting and the step-by-step solution needed to carry out the requests.

- 1 Select a task for the helper robot
- 2 See how PaLM-SayCan interprets the task
- 3 See how the helper robot executes the task
- 4 Learn the step-by-step solution behind PaLM-SayCan

Jump to tasks

<https://say-can.github.io/>

Google Research

LEAST-TO-MOST PROMPTING ENABLES COMPLEX REASONING IN LARGE LANGUAGE MODELS

Denny Zhou* Nathanael Schärlí Le Hou Jason Wei Nathan Scales Xuezhi Wang

Dale Schuurmans Claire Cui Olivier Bousquet Quoc Le Ed Chi

Google Research

SELF-CONSISTENCY IMPROVES CHAIN OF THOUGHT REASONING IN LANGUAGE MODELS

Xuezhi Wang^{†‡}, Jason Wei[†], Dale Schuurmans[†], Quoc Le[†], Ed H. Chi[†], Sharan Narang[†], Aakanksha Chowdhery[†], Denny Zhou^{†§}

[†]Google Research, Brain Team

[‡]xuezhiw@google.com, [§]dennyyzhou@google.com

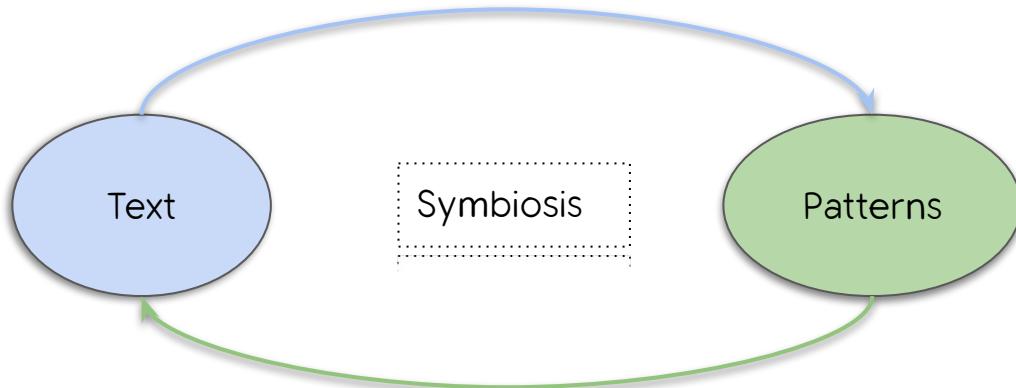
What makes chain of thought prompting so effective?

What makes the chain of thought prompting so effective?

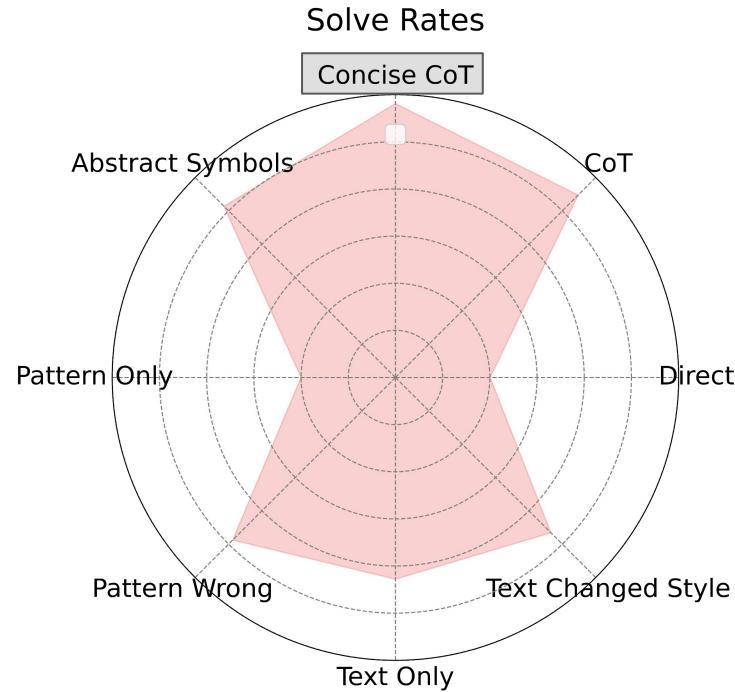
- *The thought makes the model think about the problem?*
- *The thought helps the model learn better*
- *The thought serves as an additional example of the task*
- *The thought helps the model remind of the task*
- *The thought helps extract relevant information for solving the task*

What makes chain of thought prompting so effective

Text: Extracting common sense, copy mechanism



Patterns: Task understanding, final answer generation



Approach

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

T: There are originally 3 cars. 2 more cars arrive. $3 + 2 = 5$.

A: The answer is 5 cars.

Symbols

Patterns

Text

- ***Counterfactual prompting:***

- Change one *knob* at a time (symbol, patterns, text)



What if? prompting (counterfactual prompting)

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

Thought (T): There are originally 3 cars. 2 more cars arrive. $3 + 2 = 5$.

A: The answer is 5 cars.

What if we don't have actual numbers?

What if the prompt is misleading?

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

Thought (T): There are originally α cars. β more cars arrive. $\alpha + \beta = \lambda$.

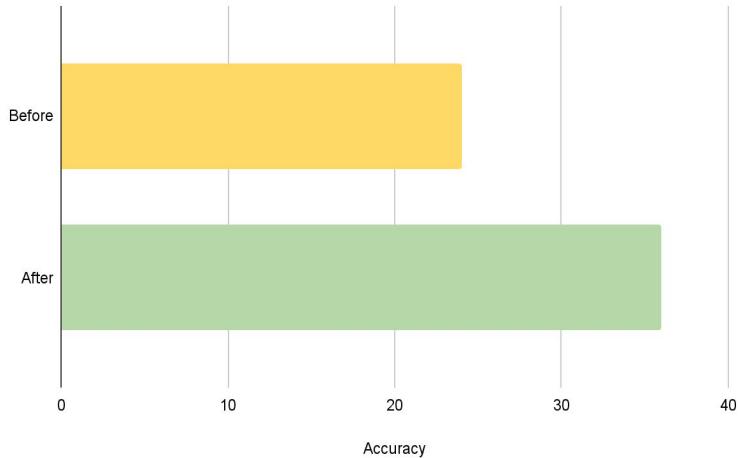
A: The answer is λ cars.

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

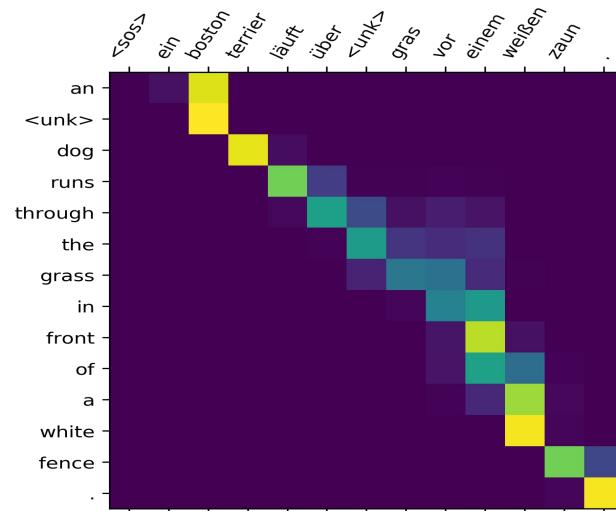
Thought (T): There are originally 3 cars. 2 more cars arrive. $3 + 2 = 7$.

A: The answer is 5 cars.

Evaluating counterfactual prompts



Empirical results: difference in final outcome



Attention patterns: difference in mechanism

Experimental Setup

Setup

- Models: PaLM-62B, PaLM-540B, CODEX, GPT-3
- Run each experiment with three seeds, report average etc.
- Results here are for PaLM-62B

Tasks

◀ MATHEMATICAL ▶

► Q → Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?

◆ T → Shawn started with 5 toys. If he got 2 toys each from his mom and dad, then that is 4 more toys. $5 + 4 = 9$.

◆ A → The answer is 9.

◀ COMMONSENSE ▶ (SPORTS)

► Q → Is the following sentence plausible? “Jamal Murray was perfect from the line.”

◆ T → Jamal Murray is a basketball player. Being perfect from the line is part of basketball.

◆ A → The answer is yes.

◀ COMMONSENSE ▶ (DATE)

► Q → It is 4/19/1969 today. What is the date 24 hours later in MM/DD/YYYY?

◆ T → Today is 04/19/1969. 24 hours later is one day after today, which would be 04/20/1969.

◆ A → The answer is 04/20/1969.

◀ SYMBOLIC ▶ (SORTING)

► Q → 2, 4, 3, 8, 9, 6, 7, 1.

◆ T → 1 < 2 < 3 < 4 < 5 < 6 < 7 < 8 < 9.

◆ A → The answer is 1, 2, 3, 4, 5, 6, 7, 8, 9.

Symbols, Patterns, Text

◀ MATHEMATICAL ▶

Question: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?

Thought: Shawn started with **5** toys. If he got **2** toys each from his mom and dad, then that is **4** more toys.
 $5 + 4 = 9$.

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

Thought: Jason started with **20** lollipops. Then he had **12** after giving some to Denny. So he gave Denny
 $20 - 12 = 8$.

◀ COMMONSENSE ▶ (SPORTS)

Question: Is the following sentence plausible? "Jamal Murray was perfect from the line."

Thought: **Jamal Murray** is a **basketball** player. **Being perfect from the line** is part of **basketball**.

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

Thought: **Joao Moutinho** is a **soccer** player. **The NFC championship** is part of American football, **not soccer**.

◀ COMMONSENSE ▶ (DATE)

Question: It is **4/19/1969** today. What is the date 24 hours later in MM/DD/YYYY?

Thought: **⟨calculation⟩** Today is **04/19/1969**. 24 hours later is one day after today, which would be **04/20/1969**. **⟨output⟩** The answer is **04/20/1969**.

Question: 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?

Thought: **⟨calculation⟩** One day after **06/01/1943** is **06/02/1943**, so today is **06/02/1943**. **⟨output⟩** 10 days before today is **05/23/1943**.

◀ SYMBOLIC ▶ (SORTING)

Question: 3, 1, 2, 7, 8, 5, 6, 9, 4

Thought: **1 < 2 ... < 9**

Part 1: Symbols

*What if we replace all the numbers
with greek alphabets?*

Abstract symbols

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

Thought (T): There are originally 3 cars. 2 more cars arrive.
 $3 + 2 = 5$.

A: The answer is 5 cars.

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

Thought (T): Originally, Leah had 32 chocolates. Her sister had 42. So in total they had $32 + 42 = 74$. After eating 35, they had $74 - 35 = 39$.

A: The answer is 39 pieces.

Q: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?

T:



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

T: There are originally α cars. β more cars arrive. $\alpha + \beta = \lambda$,

A: The answer is λ cars.

Q: Leah had α chocolates and her sister had β . If they ate λ , how many pieces do they have left in total?

T: Originally, Leah had α chocolates. Her sister had β . So in total they had $\alpha + \beta = \pi$. After eating λ , they had $\pi - \lambda = \mu$.

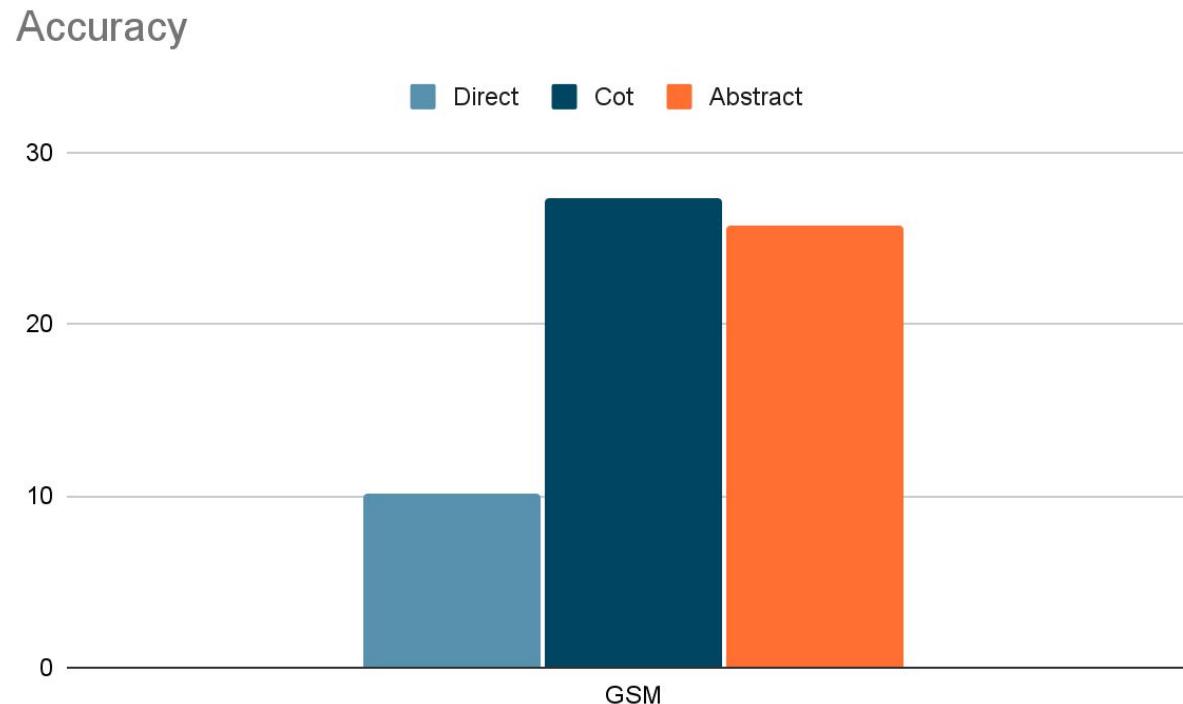
A: The answer is μ pieces.

Q: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?

T:

The test questions are not changed.

Abstract symbols results (outcome)



Performance does not change!

Abstract symbols results (mechanism)

Q : There are 5 trees the grove . Grove workers will plant trees in grove today . After they are done , there will be 2 1 trees . How many trees did the grove workers plant today ? A : There are 5 trees originally . Then there were 2 1 trees after more were planted . So there must have been $2 - 5 =$. 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 3 2 chocolates and her sister had 4 2 . If they ate 3 , how many pieces do they have left total ? A : Originally , Leah had 3 2 chocolates . Her sister had 4 2 . So total they had $3 + 4 = 7$. After eating 3 5 they had $4 - 3 = 3$. The answer is 3 9 . Q : Jason had 2 0 lollipops . He gave Denny some lollipops . Now Jason has 1 2 lollipops . How many lollipops did Jason give to Denny ? A : Jason started with 2 lollipops . Then he had 2 after giving some to Denny . So he gave Denny $2 - 1 = 1$. The answer is 1 8 . Q : Shawn has five toys . For Christmas , he got two toys each from his mom and dad . How many toys does he have now ? A : Shawn started with 5 toys . If he got 2 toys each from his mom and dad , then that is 4 more toys . $5 + 4 = 9$. The answer is 9 . Q : There were nine computers in the server room . Five more computers were installed each day , from monday to thursday . How many computers are now in the server room ? A : There were originally 9 computers . For each of 4 days , 5 more computers were added . So $5 * 4 = 20$ computers were added . $9 + 20 = 29$. The answer is 2 9 . Q : Michael had 5 8 golf balls . On tuesday , he lost 2 3 golf balls . On wednesday , he lost 2 more . How many golf balls did he have at the end of wednesday ? A : Michael started with 5 8 golf balls . After losing 2 3 on tuesday , he had $5 - 2 = 3$. After losing 2 more , he had $3 - 2 = 1$ golf ball . The answer is 1 3 . Q : Olivia has \$ 2 3 . She bought five bagels for \$ 3 each . How much money does she have left ? A : Olivia had 2 3 dollars . 5 bagels for 3 dollars each will be $5 * 3 = 15$ dollars . So she has $23 - 15 = 8$ dollars left . The answer is 8 .

Identical mechanism!

Q : There are α trees in the grove . Grove workers will plant trees in the grove today . After they are done , there will be β one trees . How many trees did the grove workers plant today ? A : There are α trees originally . Then there were β one trees after some more were planted . So there must have been $\beta - \alpha = \lambda$. The answer is λ . Q : If there are α cars in the parking lot and β more cars arrive , how many cars are the parking lot ? A : There are originally α cars . β more cars arrive . $\alpha + \beta = \lambda$. The answer is λ . Q : Leah had α chocolates and her sister had β . If they ate λ , how many pieces do they have left in total ? A : Originally , Leah had α chocolates . Her sister had β . So in total they had $\alpha + \beta = \pi$. After eating λ , they had $\pi - \lambda = \mu$. The answer is μ . Q : Jason had α lollipops . He gave Denny some . Now Jason has β lollipops . How many lollipops did Jason give to Denny ? A : Jason started with α lollipops . Then he had β after giving some to Denny . So he gave Denny $\alpha - \beta = \lambda$. The answer is λ . Q : Shawn has α toys . For Christmas , he got β toys each from his mom and dad . How many toys does he have now ? A : Shawn started with α toys . If he got β toys each from his mom and dad , then that is λ more toys . $\alpha + \lambda = \pi$. The answer is π . Q : There were α computers in the server room . β more computers were installed each day , from monday to thursday . How many computers are now in the server room ? A : There were originally α computers . For each of four days , β more computers were added . So $\beta * 4 = \lambda$ computers were added . $\alpha + \lambda = \pi$. The answer is π . Q : Michael had α golf balls . On tuesday , he lost β golf balls . On wednesday , he lost λ more . How many golf balls did he have at the end of wednesday ? A : Michael started with α golf balls . After losing β on tuesday , he had $\alpha - \beta = \mu$. After losing λ more , he had $\mu - \lambda = \nu$ golf balls . The answer is ν . Q : Olivia has \$ α . She bought five bagels for \$ β each . How much money does she have left ? A : Olivia had α dollars . 5 bagels for β dollars each will be $5 * \beta = \lambda$ dollars . So she has $\alpha - \lambda$ dollars left . $\alpha - \lambda = \pi$. The answer is π .

Other tasks

Thought: **Jamal Murray** is a basketball player. **Being perfect from the line** is part of basketball.

SPORTS

Thought: **PERSON** is a basketball player. **Being perfect from the line** is part of basketball.

Thought: **Jamal Murray** is a basketball player. **Being ACTIVITY** is part of basketball.

Thought: Today is **04/19/1969**. 24 hours later is one day after today, which would be **04/20/1969**.

DATE

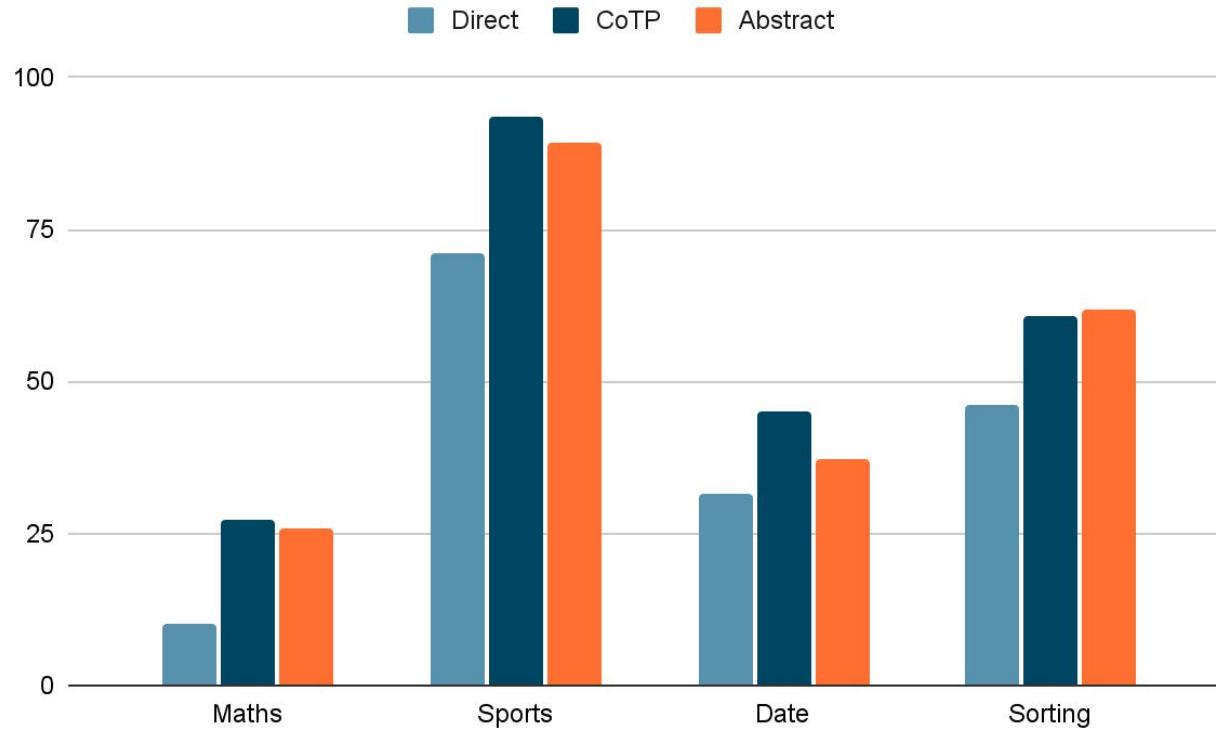
Thought: Today is **DATE**. 24 hours later is one day after today, which would be **DATE**.

Thought: **1 < 2 < 3 < 4 < 5 < 6 < 7 < 8 < 9**

SORTING

Thought: **$\varsigma < \phi < \gamma < \delta < \zeta < \chi < \epsilon < \pi < \upsilon$**

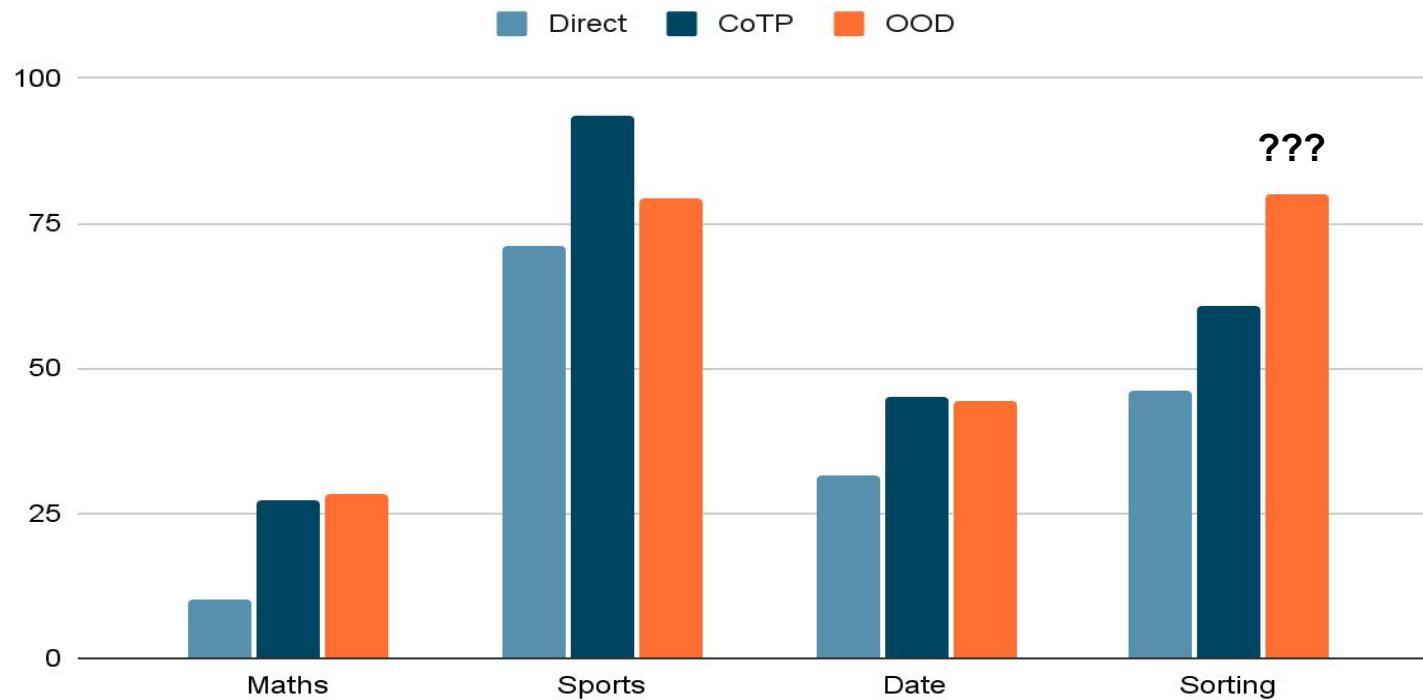
Abstract Symbols



OOD Symbols

- Instead of replacing symbols with abstract outputs, replace with “OOD” symbols
 - 5 toys → 5.5 toys
 - Fractions don’t appear in GSM
 - Jamal Murray → Adair Foster
 - Randomly generated name
 - 04/19/1969 → 04/30/3069
 - Date in the future
 - $1 < 2 \rightarrow 11 < 23$
 - The task involves sorting integers

OOD Symbols



Few-shot learning?

◀ SYMBOLIC ▶ (SORTING)

Question: 7 , 8 , 4 , 1 , 2 , 9 , 3 , 6 , 5	DIRECT	46.0%
Thought: 1 < 2 < 3 < 4 < 5 < 6 < 7 < 8 < 9	COT (Table 32)	60.6%
Thought: $\varsigma < \phi < \gamma < \delta < \zeta < \chi < \epsilon < \pi < \nu$	$C_{\text{symb_abs}}(p)$ (Table 35)	61.8%
Thought: 11 < 23 < 34 < 48 < 56 < 63 < 72 < 85 < 95	$C_{\text{symb_ood}}(p)$ (Table 42)	80.0%

- With standard prompt, the model tends to generate the count 1, 2, ..., 9
- OOD prompt is better at “reminding” the model of the task

Part 2: Patterns

Symbols, Patterns, Text

◀ MATHEMATICAL ▶

Question: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?

Thought: Shawn started with **5** toys. If he got **2** toys each from his mom and dad, then that is **4** more toys.
 $5 + 4 = 9$.

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

Thought: Jason started with **20** lollipops. Then he had **12** after giving some to Denny. So he gave Denny
 $20 - 12 = 8$.

◀ COMMONSENSE ▶ (SPORTS)

Question: Is the following sentence plausible? "Jamal Murray was perfect from the line."

Thought: **Jamal Murray** is a **basketball** player. **Being perfect from the line** is part of **basketball**.

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

Thought: **Joao Moutinho** is a **soccer** player. **The NFC championship** is part of American football, **not soccer**.

◀ COMMONSENSE ▶ (DATE)

Question: It is **4/19/1969** today. What is the date 24 hours later in MM/DD/YYYY?

Thought: **⟨calculation⟩** Today is **04/19/1969**. 24 hours later is one day after today, which would be **04/20/1969**. **⟨output⟩** The answer is **04/20/1969**.

Question: 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?

Thought: **⟨calculation⟩** One day after **06/01/1943** is **06/02/1943**, so today is **06/02/1943**. **⟨output⟩** 10 days before today is **05/23/1943**.

◀ SYMBOLIC ▶ (SORTING)

Question: 3, 1, 2, 7, 8, 5, 6, 9, 4

Thought: **1 < 2 ... < 9**

What if all the math is wrong?

Abstract symbols

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

Thought (T): There are originally 3 cars. 2 more cars arrive. $3 + 2 = 5$.

A: The answer is 5 cars.

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

Thought (T): Originally, Leah had 32 chocolates. Her sister had 42. So in total they had $32 + 42 = 74$. After eating 35, they had $74 - 35 = 39$.

A: The answer is 39 pieces.

Q: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?

T:



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

Thought (T): There are originally 3 cars. 2 more cars arrive. $3 + 2 = 7$.

A: The answer is 5 cars.

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

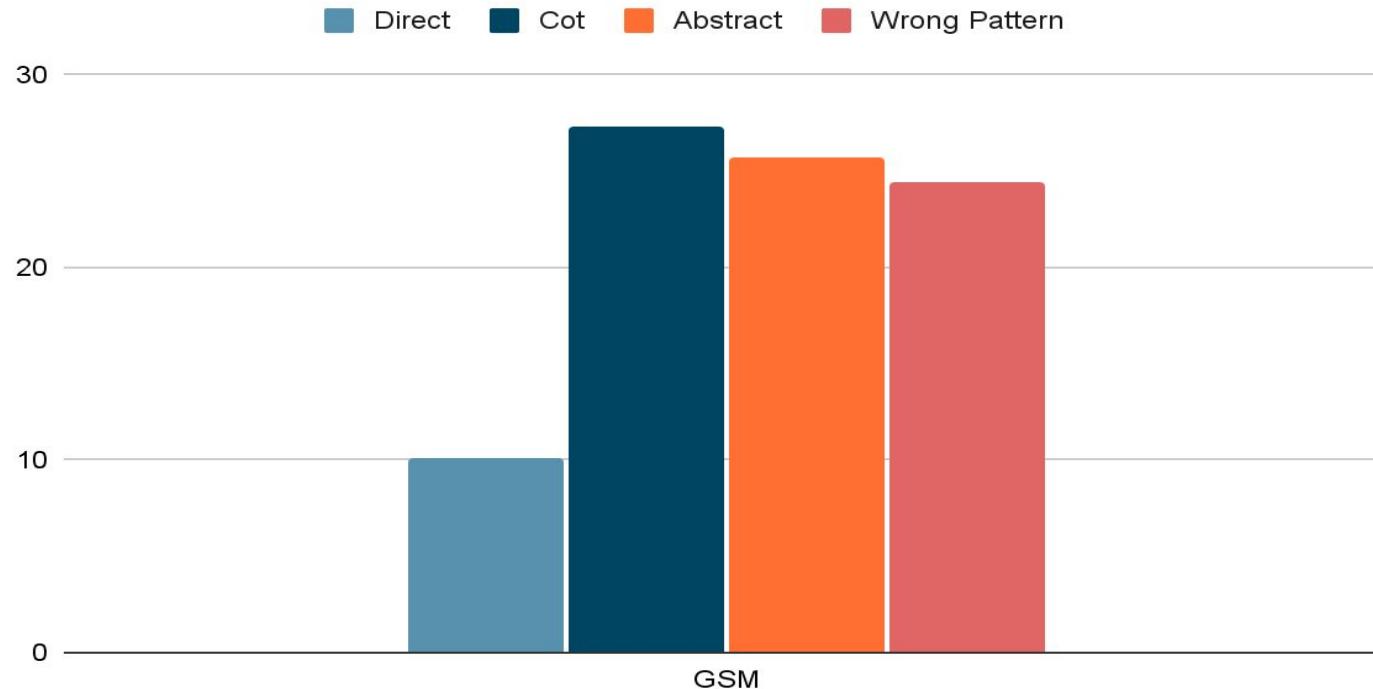
Thought (T): Originally, Leah had 32 chocolates. Her sister had 42. So in total they had $32 + 42 = 50$. After eating 35, they had $74 - 35 = 25$.

A: The answer is 39 pieces.

Q: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?

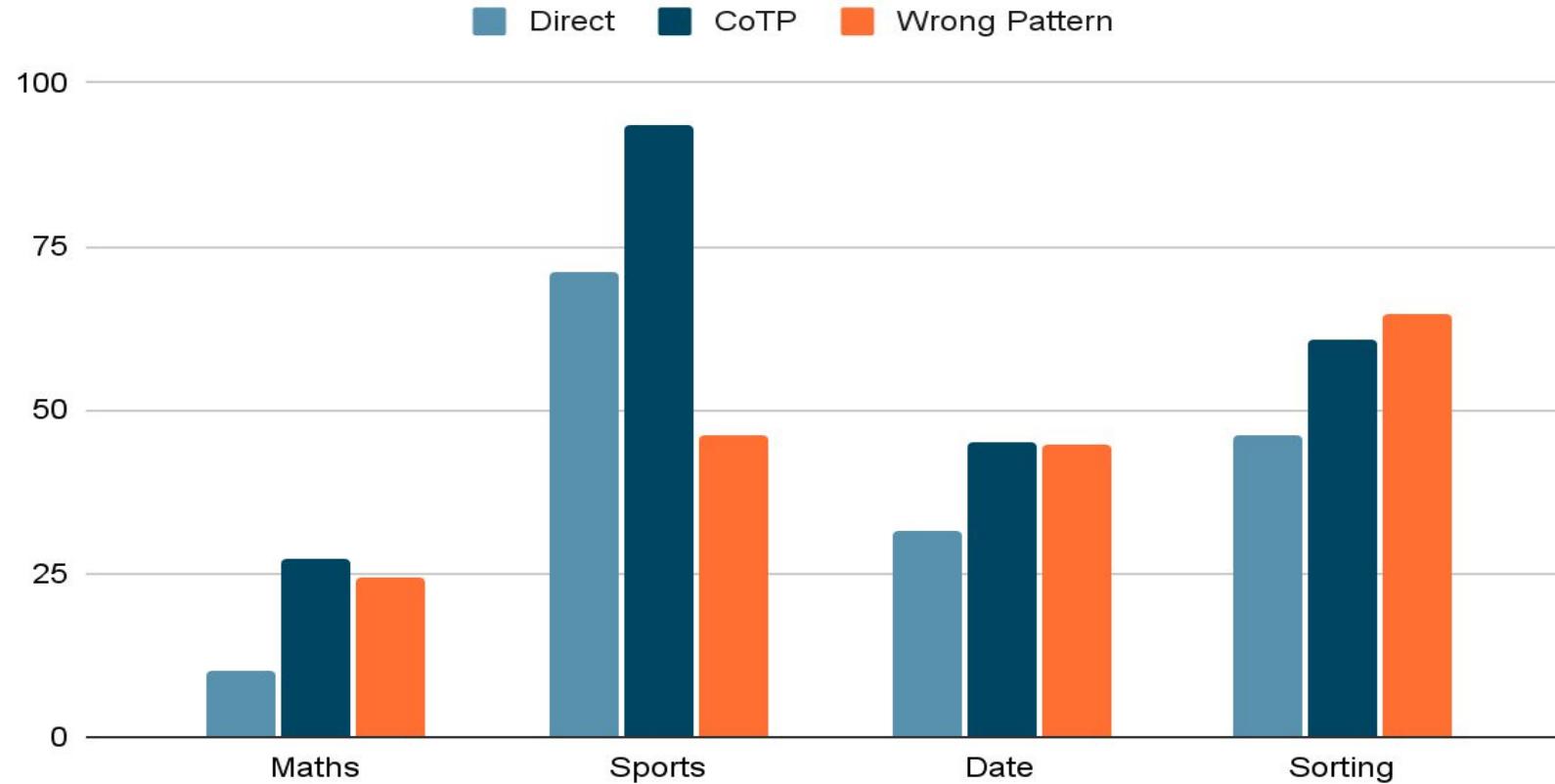
T:

Wrong math



Performance does not change!

Other Tasks



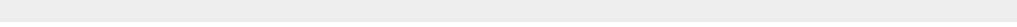
Wrong Patterns Drastically Impact SPORTS

Thought: Jamal Murray **is** a **basketball** player. Being perfect from the line **is** part of CoT (Table 31) 93.67%

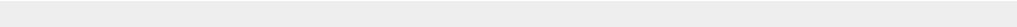
Thought: Jamal Murray **is** a **soccer** player. Being perfect from the line **is** part of $C_{\text{pat_wrong}}(p)$ (Table 55) 46.02%

- 50% is random baseline – the model is intentionally trying to generate misleading outputs!?
- The role of label correctness is task dependent – as long as the model can be reminded of the task

Kim, Junyeob, Hyuhng Joon Kim, Hyunsoo Cho, Hwiyeol Jo, Sang-Woo Lee, Sang-goo Lee, Kang Min Yoo, and Taeuk Kim. "Ground-Truth Labels Matter: A Deeper Look into Input-Label Demonstrations." *arXiv preprint arXiv:2205.12685* (2022).



***What if we remove
all the patterns?***



No patterns

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

Thought (T): There are originally 3 cars. 2 more cars arrive. $3 + 2 = 5$.

A: The answer is 5 cars.

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

Thought (T): Originally, Leah had 32 chocolates. Her sister had 42. So in total they had $32 + 42 = 74$. After eating 35, they had $74 - 35 = 39$.

A: The answer is 39 pieces.

Q: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?

T:



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

T: There are originally 3 cars. 2 more cars arrive.

A: The answer is 5 cars.

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

T: Originally, Leah had 32 chocolates. Her sister had 42. So in total they had 74. After eating 35, they had 39.

A: The answer is 39 pieces.

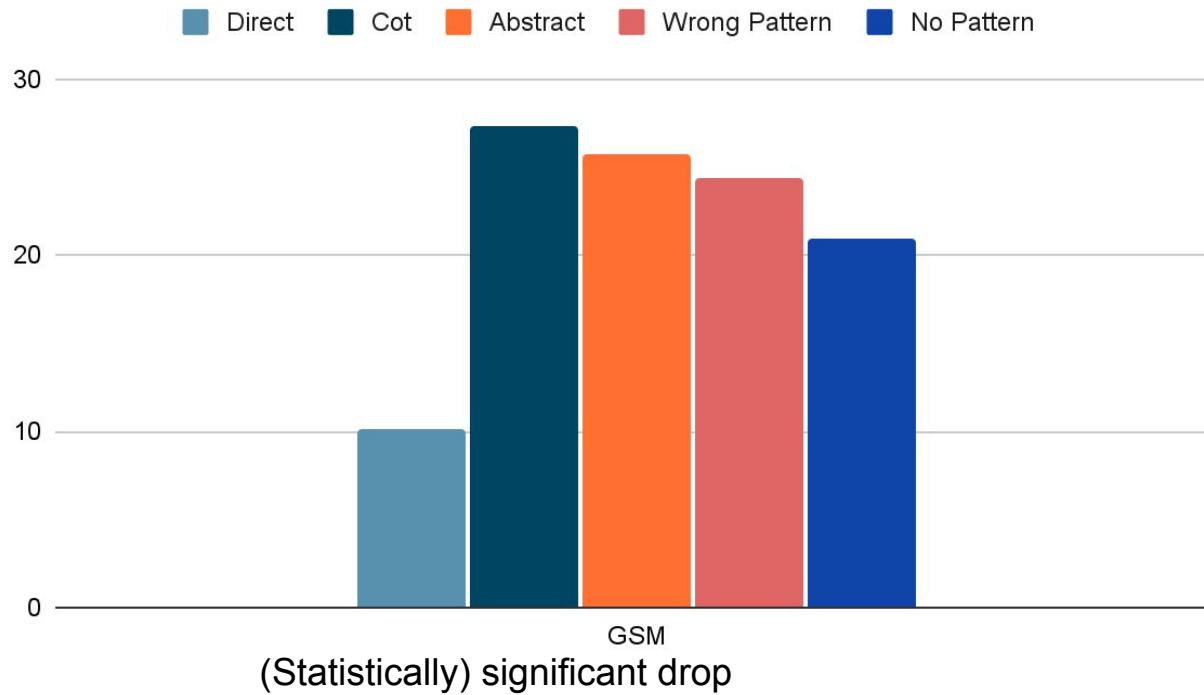
Q: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?

T:

The test questions are not changed.

No patterns results (outcome)

Accuracy no pattern





**What if we ~~delete the text~~ and
just keep the patterns?**

No Text

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

Thought (T): There are originally 3 cars. 2 more cars arrive. $3 + 2 = 5$.

A: The answer is 5 cars.

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

Thought (T): Originally, Leah had 32 chocolates. Her sister had 42. So in total they had $32 + 42 = 74$. After eating 35, they had $74 - 35 = 39$.

A: The answer is 39 pieces.

Q: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?

T:



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

T: $3 + 2 = 5$.

A: The answer is 5 cars.

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

T: $32 + 42 = 74$. $74 - 35 = 39$.

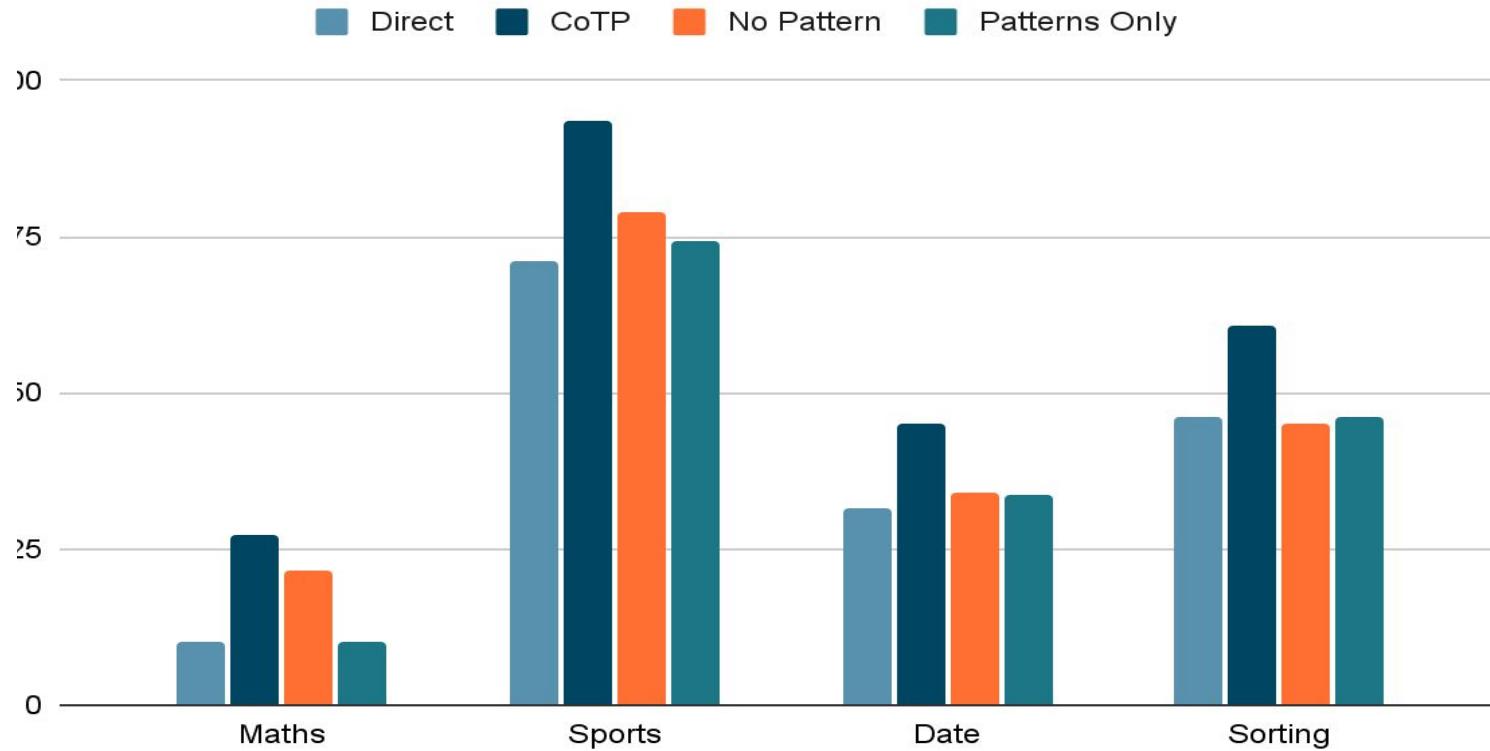
A: The answer is 39 pieces.

Q: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?

T:

The test questions are not changed.

Patterns are necessary but not sufficient



Correctness is task dependent – as long as they can remind the model of the task!

Part 3: Text

**Change the grammatical
style of the text, we will**

Question / Thought	Prompt Type	Solve Rate
◀ MATHEMATICAL ▶		
Question: Shawn has five toys. For Christmas, he got two toys each from his mom and dad. How many toys does he have now?	DIRECT	10.11%
Thought: Shawn started with 5 toys. If he got 2 toys each from his mom and dad, then that is 4 more toys. $5 + 4 = 9$.	CoT (Table 29)	27.37%
Thought: Teddy started with 5 cookies. If he got 2 cookies each from his Jenna and Rehan, then that is 4 more cookies. $5 + 4 = 9$.	$C_{\text{text_diff_entities}}(p)$ (Table 59)	16.60%
Thought: Capacity of one bus is 198 passengers / 9 buses = 22 passengers in one bus. Thus, 5 buses can fit $22 * 5 = 110$ passengers.	$C_{\text{text_rand}}(p)$ (Table 61)	2.98%
-	$C_{\text{text_yoda_question}}(p)$ (Table 73)	27.09%
Thought: With 5 toys, Shawn started. 2 toys each from his mom and dad, if he got, then that is 4 more toys. $5 + 4 = 9$.	$C_{\text{text_yoda_thought}}(p)$ (Table 70)	23.22%
Thought: With 5 toys, Shawn started. 2 toys each from his mom and dad, if he got, then that is 4 more toys. $5 + 4 = 9$.	$C_{\text{text_yoda_both}}(p)$ (Table 73)	22.19%
Thought: dad, he got 5 toys. then started mom 2 each is more than from If his and toys. toys 4 with Shawn $5 + 4 = 9$.	$C_{\text{text_inter_shuf}}(p)$ (Table 67)	10.84%
Thought: with Shawn toys 5 started. dad, from more 2 his toys then is toys he mom got that each 4 and If. $5 + 4 = 9$.	$C_{\text{text_intra_shuf}}(p)$ (Table 64)	17.01%

Performance \propto change

Changing the question doesn't affect performance

Analogical Reasoning

Part 4: Symbiosis

Symbiosis of Text and Patterns Enables effective CoT

→ Q3 → A pet store currently has 5 dogs, 2 cats, and 10 birds. How many legs in total do the pets in the store have?

◆ $\mathcal{T}^r[\text{CoT}(p)] \mapsto$ There are 5 dogs, 2 cats, and 10 birds. 5 dogs have 4 legs each, so $5 * 4 = 20$ legs. 2 cats have 4 legs each, so $2 * 4 = 8$ legs. 10 birds have 2 legs each, so $10 * 2 = 20$ legs. So the total number of legs is $20 + 8 + 20 = 48$.
The answer is 48.✓

◆ $\mathcal{T}^r[C_{\text{pat_only}}(p)] \mapsto 5 + 2 + 10 = 17.$ The answer is 17.✗

→ Q6 → Today is Christmas Eve of 1937. What is the date one week ago from today in MM/DD/YYYY?

◆ $\mathcal{T}^r[\text{CoT}(p)] \mapsto$ Today is 12/24/1937. One week ago from today is 12/17/1937. So the answer is 12/17/1937.✓

◆ $\mathcal{T}^r[C_{\text{pat_only}}(p)] \mapsto 12/20/1937$ ✗

Symbiosis of Text and Patterns Enables effective CoT

→ Q3 → A pet store currently has 5 dogs, 2 cats, and 10 birds. How many legs in total do the pets in the store have?

◆ $\mathcal{T}[\text{CoT}(p)] \rightarrow$ There are 5 dogs, 2 cats, and 10 birds. 5 dogs have 4 legs each, so $5 * 4 = 20$ legs. 2 cats have 4 legs each, so $2 * 4 = 8$ legs. 10 birds have 2 legs each, so $10 * 2 = 20$ legs. So the total number of legs is $20 + 8 + 20 = 48$.
The answer is 48. ✓

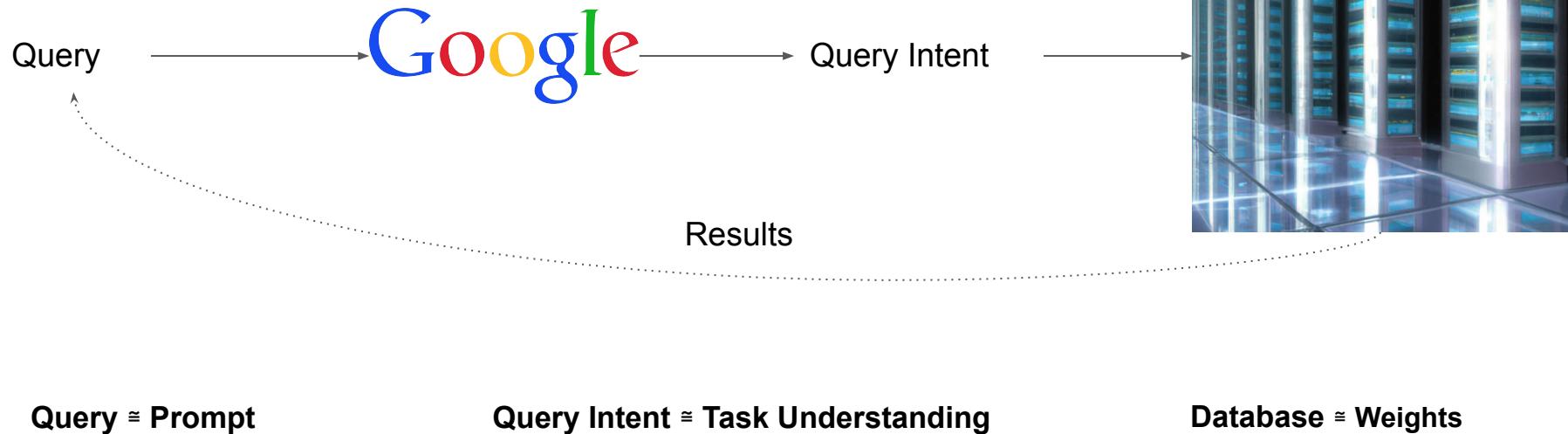
◆ $\mathcal{T}[C_{\text{pat_none}}(p)] \rightarrow$ The pet store has 5 dogs, 2 cats, and 10 birds. So the total number of legs is $5 + 2 + 10 = 17$.
The answer is 17. ✗

→ Q4 → Yes or no: Is the following sentence plausible? “Robert Woods killed the powerplay”.

◆ $\mathcal{T}[\text{CoT}(p)] \rightarrow$ Robert Woods is a football player. Killing the powerplay is part of hockey, not football. So the answer is no. ✓

◆ $\mathcal{T}[C_{\text{pat_none}}(p)] \rightarrow$ Robert Woods is a football player. Killing the powerplay is a part of hockey. So the answer is yes. ✗

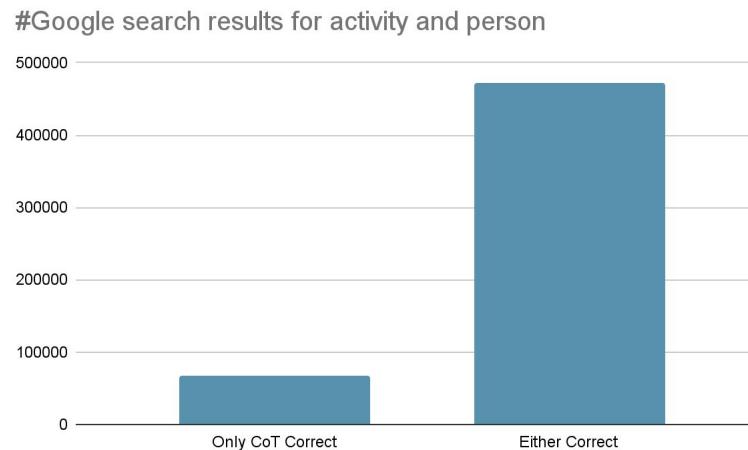
The Search Engine Analogy



Also see: <https://ai.stanford.edu/blog/understanding-incontext/>

The Search Engine Analogy

- For SPORTS, measure the popularity of activities and sports person for cases where CoT is exclusively correct
- CoT is especially helpful for rare entities



Concise Chain of Thought

- Helps us to identify the bare minimum necessary for good performance

Task	PaLM-62B		PaLM-540B	
	CoT	CCoT (ours)	CoT	CCoT (ours)
GSM-8K	27.4%	29.1%	53.2%	56.2%
DATE	44.7%	51.3%	65.3%	69.1%
SPORTS	93.7%	94.6%	95.4%	97.4%
SORTING	55.3%	60.2%	71.2%	88.6%

- **1.8x fewer tokens**

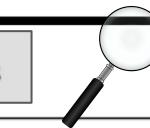
Not Included in the Presentation

- Results on GPT-3, CODEX, PaLM-540B
 - Similar trends across models
- Detailed prompts
- Error analysis

What makes the chain of thought prompting so effective?

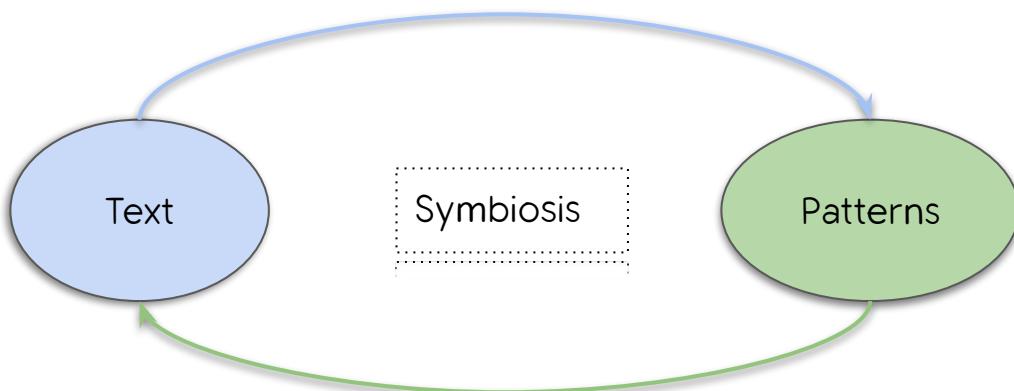
- *The thought makes the model think about the problem?*
- *The thought helps the model learn better*
- *The thought serves as an additional example of the task* ~
- *The thought helps the model remind of the task* ~
- *The thought helps extract relevant information for solving the task* ~

Language Models are Few-Shot Learners



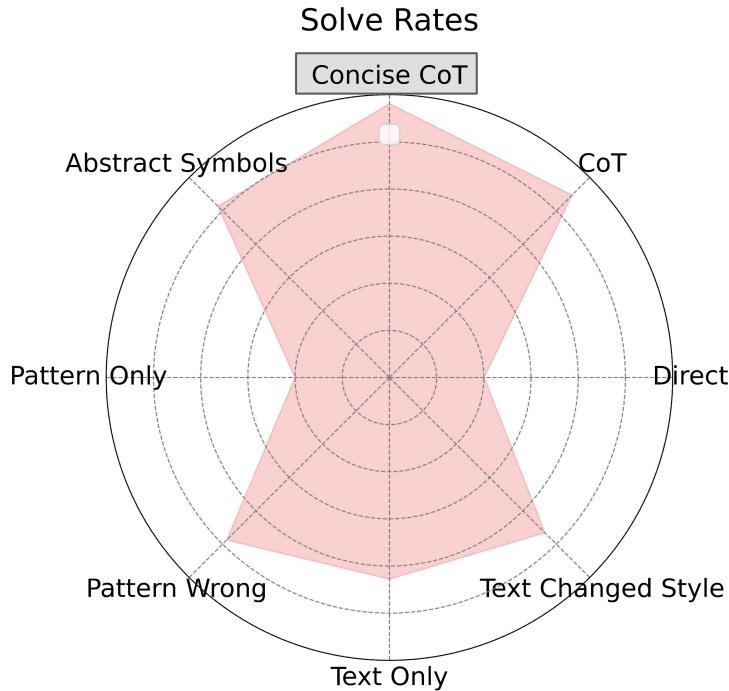
What makes chain of thought prompting so effective?

Text: Extracting common sense, copy mechanism



Patterns: Task understanding, final answer generation

Introduce text in your problem, make sure the prompt conveys the task (can remind the model of the task!)



[Paper](#)

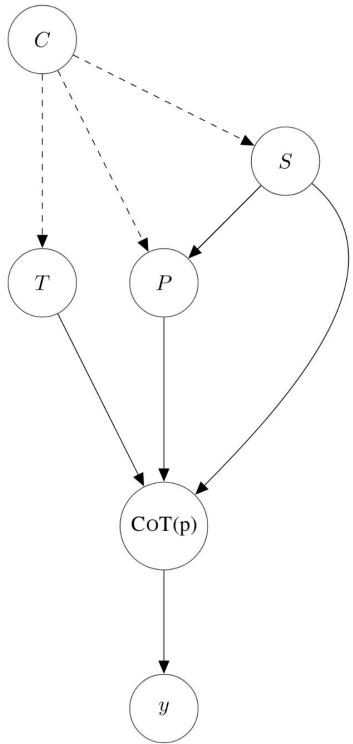


Figure 5: Counterfactual graph for our work: symbols S , patterns P , text T are the key components of our work. The outcome is binary y (accuracy). We make a simplifying assumption that these are identifiable. S and P may be confounded by an unidentifiable C , and this is a limitation of our work.

$$ATE = \mathbb{E}_D[y(p) - y(C_z(p))]$$