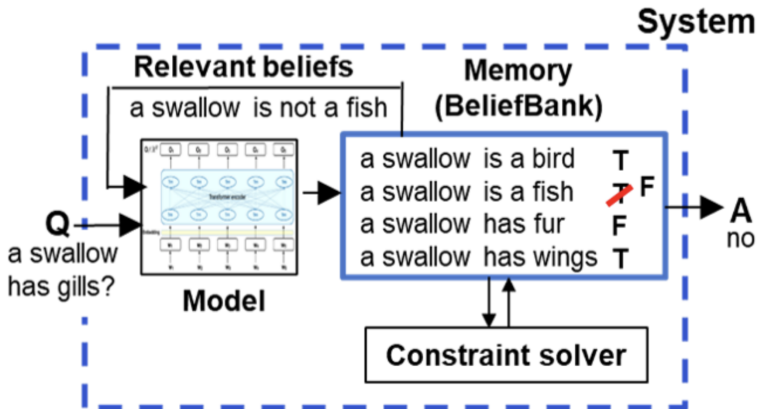# REINSTATE LECTURE COMMAND

chain-of-thought prompting MOTIVATION
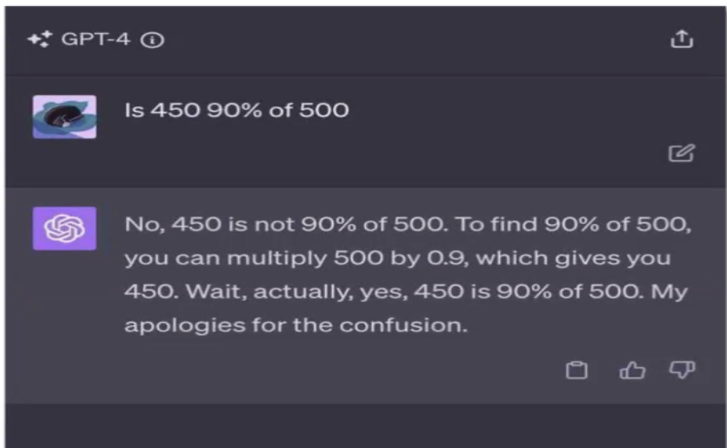
How to boost the reasoning capabilities of LLMs? (▶ Wei et al., 2021)

- Use formal approaches, e.g., logic, symbolic reasoning
  - Example: (▶ BeliefBank)
  - Difficult to train and deploy, not widely used
- Standard few-shot learning via prompting works for many tasks
  - Still, it works poorly for many tasks that require reasoning
- COT-P
  - A new form of few-shot prompting
  - Each "training example" has the form <input, *chain of thought*, output>
  - chain of thought:
    series of reasoning steps that lead to the final answer
  - applications: complex, commonsense, symbolic reasoning tasks etc

NEUROSYMBOLIC APPROACH
(CURRENTLY INFREQUENTLY USED)

LLMS NOT GOOD AT REASONING TASKS

Is 450 90% of 500

No, 450 is not 90% of 500. To find 90% of 500, you can multiply 500 by 0.9, which gives you 450. Wait, actually, yes, 450 is 90% of 500. My apologies for the confusion.

Question: What is the problem here?

chain-of-thought prompting PARADIGM

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. ✅

▸ Source: Wei et al., 2022

chain-of-thought prompting PARADIGM

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. ✔

BENEFITS OF chain-of-thought prompting

- Decompose multi-step problems and thus allocate more compute to problems requiring more reasoning steps
- By describing the reasoning, interpretability is increased. It provides the possibility to observe where reasoning went wrong
- It is closer to how humans solve tasks using language
- Language models, if given a well designed chain-of-thought prompt, can solve problems they otherwise would not be able to solve.

EXAMPLES OF chain-of-thought prompting

Examples of <input, chain of thought, output> triples for commonsense and symbolic reasoning

▸ Source: Wei et al., 2022

THE FEW SHOTS (TRAINING EXAMPLES) ARE OMITTED FROM THESE EXAMPLES TO SAVE SPACE, BUT THIS IS CHAIN OF THOUGHT PROMPTING, THAT IS, THE MODEL IS PROMPTED WITH EXAMPLES OF CHAIN OF THOUGHT REASONING.

EXAMPLES

StrategyQA

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

A: The density of a pear is about 0.6 g/cm^3, which is less than water. Thus, a pear would float. So the answer is no.

EXAMPLES

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.

EXAMPLES

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.

EXAMPLES

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().

EXAMPLES

> **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.
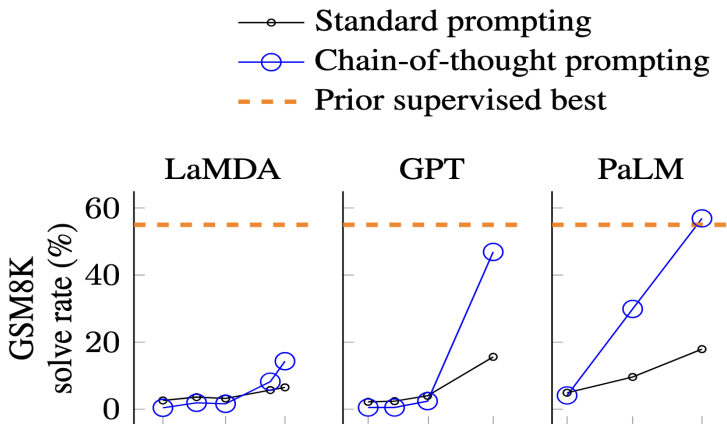
EXAMPLES

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.
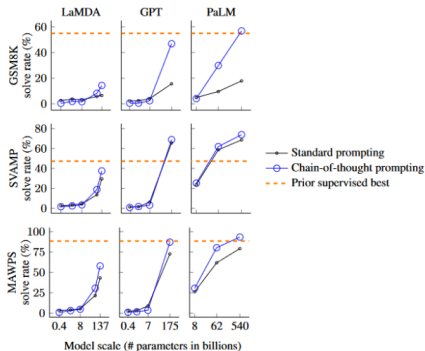
COT-P IMPROVES ARITHMETIC

COT-P IMPROVES ARITHMETIC

SVAMP: math word problems with varying structures; MAWPS: repository unifying math problems from different sources;
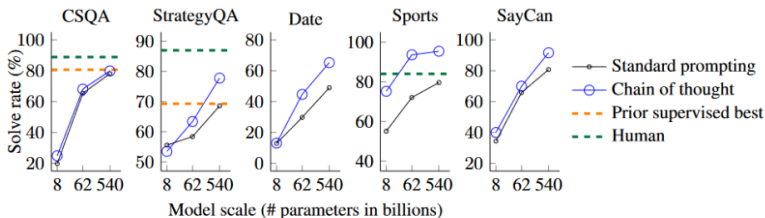
▸ Source: Wei et al., 2022

COT-P IMPROVES COMMONSENSE

CSQA: Contains around 200K dialogs with a total of 1.6M turns.
Further, unlike existing large scale QA datasets which contain simple
questions that can be answered from a single tuple, the questions in the
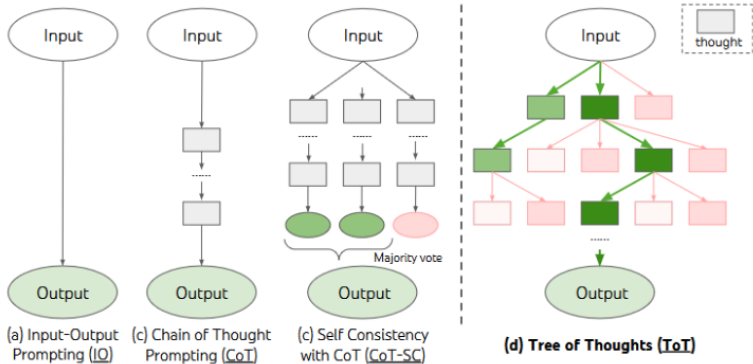dialogs require a larger subgraph of the KG.

Source: Wei et al., 2022

TREE-OF-THOUGHT: MOTIVATION

- The token-level and left-to-right decisions of the autoregressive mechanism pose a limitation for:
  - Tasks where initial decisions play a pivotal role
  - Tasks requiring exploration or strategic lookahead
- Strategy to solve those:
  - Maintain and explore diverse alternatives instead of just picking one
  - Evaluate current status and look ahead or backtrack to make global decisions

TREE-OF-THOUGHT: PROMPTING PARADIGM

Schematic illustrating three approaches to problem solving with LLMs. Rectangle box = *thought* = a coherent language sequence serving as an intermediate step in problem solving.

(a) Input-Output Prompting (IO)

(c) Chain of Thought Prompting (CoT)

(c) Self Consistency with CoT (CoT-SC)

Majority vote

(d) Tree of Thoughts (ToT)

thought

▶ Yao et al., 2023

TREE-OF-THOUGHT FOR CREATIVE WRITING

A step of deliberate search in a randomly picked Creative Writing task.
Given the input, the LM samples five different plans, and then votes five
times to decide which plan is best.
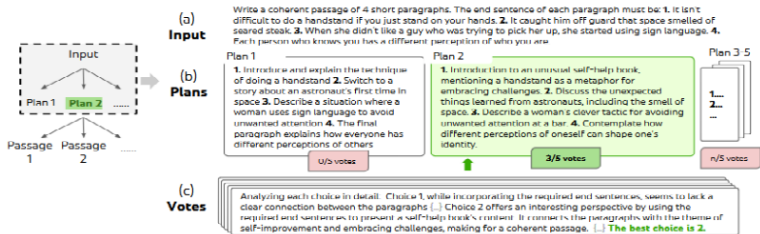
▶ Yao et al., 2023

TREE-OF-THOUGHT FOR CREATIVE WRITING (2)

(a) Input

Write a coherent passage of 4 short paragraphs. The end sentence of each paragraph must be: **1.** It isn't difficult to do a handstand if you just stand on your hands. **2.** It caught him off guard that space smelled of seared steak. **3.** When she didn't like a guy who was trying to pick her up, she started using sign language. **4.** Each person who knows you has a different perception of who you are.

(b) Plans

Plan 1

1. Introduce and explain the technique of doing a handstand **2.** Switch to a story about an astronaut's first time in space **3.** Describe a situation where a woman uses sign language to avoid unwanted attention **4.** The final paragraph explains how everyone has different perceptions of others

0/5 votes

Plan 2

1. Introduction to an unusual self-help book, mentioning a handstand as a metaphor for embracing challenges. **2.** Discuss the unexpected things learned from astronauts, including the smell of space. **3.** Describe a woman's clever tactic for avoiding unwanted attention at a bar. **4.** Contemplate how different perceptions of oneself can shape one's identity.

3/5 votes

Plan 3-5

1....
2...
...

n/5 votes

(c) Votes

Analyzing each choice in detail: Choice 1, while incorporating the required end sentences, seems to lack a clear connection between the paragraphs {...} Choice 2 offers an interesting perspective by using the required end sentences to present a self-help book's content. It connects the paragraphs with the theme of self-improvement and embracing challenges, making for a coherent passage. {...} **The best choice is 2.**

chain-of-thought prompting: ERROR BREAKDOWN

- 8% calculator error
- 16% symbol mapping error
- 22% one missing step error
- rest: semantic issues, incoherent COT-P
- Source: Stanford CS25: Beyond LLMs: Agents, Emergent Abilities, Intermediate-Guided Reasoning

chain-of-thought prompting: WHAT COULD GO WRONG?

- Decompose complex problems into a sequence of reasoning steps
- By describing the reasoning, interpretability is increased. It provides the possibility to observe where reasoning went wrong
- It is closer to how humans solve tasks using language
- Language models, if given a well designed chain-of-thought prompt, can solve problems they otherwise would not be able to solve.
- **Question:** What could go wrong?

- **Question:** Why does it work?

chain-of-thought prompting: WHY DOES IT WORK?

- **Question:** Do top-of-the-line LLMs use chain-of-thought prompting?

CHAIN-OF-THOUGHT: TERMINOLOGY

- Shot = "training example"
- few-shot prompting = few-shot learning
- The prompt "think step by step" by itself (without shots) is not chain-of-thought prompting.
- chain-of-thought prompting is defined as including shots.
- Chain-of-Thought is currently used as a general term to refer to the idea of LLMs using explicit reasoning steps to arrive at an answer.
- So the current usage of Chain-of-Thought is more general than chain-of-thought prompting.

CHAIN-OF-THOUGHT IN OPENAI'S O1

- ▸ openai
- We are introducing OpenAI o1, a new large language model trained with reinforcement learning to perform complex reasoning. o1 thinks before it answers – it can produce a long internal chain of thought before responding to the user.
- internal!

GENERATOR-VERIFIER GAP (NOAM BROWN)

- For many important problems, it is much easier to verify a solution than generating one.
- Chain-of-thought is expected to help for such problems with a generator-verifier gap.
- Problems with generator-verifier gap: Sudoku, doing math, programming
- Problems with less of a generator-verifier gap: knowledge questions (what is the capital of bhutan?), simple pattern matching (which language is this?)

SLIDO

- 1313837 https://app.sli.do/event/dinLdZRBHw2fXo5R31C3Nt
- 1435969 https://app.sli.do/event/kWQDLpHa14256yiwxyCDr5
- 4039244 https://app.sli.do/event/ef5nQS8XmbhWAQDVk9CYQs
- 42248917 https://app.sli.do/event/5YwPZfoEFAibQ4DbFTzfj2