

A decorative graphic on the left side of the slide consisting of two overlapping parallelograms. The front one is blue and the back one is a light green color. They are positioned diagonally, with the blue one in front of the green one.

Chain of Thought Prompting

How to use Generative AI to create content

Get the LLM to show it's work

At a high level, Chain of thought prompting (or CoT) prompting breaks apart a complicated concept into smaller pieces by showing its work

This can be triggered via few shot prompting and showing an example of it

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. ✅



Short responses are more inaccurate

Getting a short answer from an LLM is less reliable, but a longer response can get a more accurate answer.

Example:

Prompt: I went to the market and bought 10 apples. I gave 2 apples to the neighbor and 2 to the repairman. I then went and bought 5 more apples and ate 1. How many apples did I remain with?

Answer: 11 Apples ❌ This is wrong



Showing its work for logical reasoning

Context is a combination of the current prompt and the previous responses into building a new response from the LLM.

Example:

- Q: Roger has 11 balls. He buys 2 more cans of tennis balls, each with 3 tennis balls. How many tennis balls does Roger have?
- A: The answer is 17 balls
- Q: The cafeteria has 23 plates. They use 20 for lunch and buy 6 more. How many plates do they have?

The **A: The answer is 17 balls** provides context on how the LLM should respond



How to give an example

If we change the format of the response in the context, it changes the way that the LLM will respond.

Example:

- Q: Roger has 11 balls. He buys 2 more cans of tennis balls, each with 3 tennis balls. How many tennis balls does Roger have?
- A: Roger started with 11 balls, 2 cans of tennis balls with 3 each is 6 balls. $6 + 11 = 17$ balls. The answer is 17 balls
- Q: The cafeteria has 23 plates. They use 20 for lunch and buy 6 more. How many plates do they have?

Now the context window has an example that the LLM can reference to show it's work.

Zero shot and few shot CoT

Can trigger chain of thought prompting with Few shot and Zero Shot.

Few Shot prompting can be triggered via examples

Zero Shot can be triggered with a prompt “Let’s think step by step”

(a) Few-shot

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

A: The answer is 11.

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

A:

(Output) The answer is 8. ✗

(b) Few-shot-CoT

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

A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11.

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

A:

(Output) The juggler can juggle 16 balls. Half of the balls are golf balls. So there are $16 / 2 = 8$ golf balls. Half of the golf balls are blue. So there are $8 / 2 = 4$ blue golf balls. The answer is 4. ✓

(c) Zero-shot

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

A: The answer (arabic numerals) is

(Output) 8 ✗

(d) Zero-shot-CoT (Ours)

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

A: **Let's think step by step.**

(Output) There are 16 balls in total. Half of the balls are golf balls. That means that there are 8 golf balls. Half of the golf balls are blue. That means that there are 4 blue golf balls. ✓



“Let’s think step by step”

Now lets try with a prompt that makes the LLM think

Example:

Prompt: I went to the market and bought 10 apples. I gave 2 apples to the neighbor and 2 to the repairman. I then went and bought 5 more apples and ate 1. How many apples did I remain with?

Let's think step by step.

Answer: First, you started with 10 apples.

You gave away 2 apples to the neighbor and 2 to the repairman, so you had 6 apples left.

Then you bought 5 more apples, so now you had 11 apples.

Finally, you ate 1 apple, so you would remain with 10 apples.

With chain of thought prompting, you get the right answer because the prompt leads to a longer response.