

Training, Prompt Engineering, and LLM Agents

Neil Gong

Pre-training

Post-training

- Supervised fine-tuning
- Reinforcement learning

Chain-of-thought reasoning

Reasoning Problems

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**

Q: Take the last letters of the words in "Elon Musk" and concatenate them

A: The answer is **nk**.

Q: What home entertainment equipment requires cable?
Answer Choices: (a) radio shack
(b) substation (c) television (d) cabinet

A: The answer is **(c)**.

Arithmetic Reasoning (AR)
 $(+ - \times \div \dots)$

Symbolic Reasoning (SR)

Commonsense Reasoning (CR)

Chain of Thought (CoT)

Definition:

A chain of thought is **a series of intermediate natural language reasoning steps** that lead to the final output.

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. 

Post-training with CoT

- Supervised fine-tuning
- Reinforcement learning

LLM Agent

- Environment
- Actions
 - Tool call
- Reasoning/planning
- Memory