



Natural Language Processing

Note of Denny Zhou's (Google DeepMind) Talk in Standford, April 2025



LLM Reasoning

Denny Zhou

Google DeepMind



Stanford, CS25: Transformers United V5, April 29, 2025

Stanford

<https://www.youtube.com/watch?v=ebnX5Ur1hBk>

Common brief:

*Pretrained LLMs cannot reason
without further prompting
engineering or finetuning ?*

Pretrained LLMs are ready to reason

All we need is decoding

Reasoning by decoding

I have 3 apples. My dad has 2 more apples than me. How many apples do we have in total?

5 apples. (Greedy Decoding)

I have 3 apples, my dad has 2 more apples than me, so he has 5 apples. $3+5=8$.

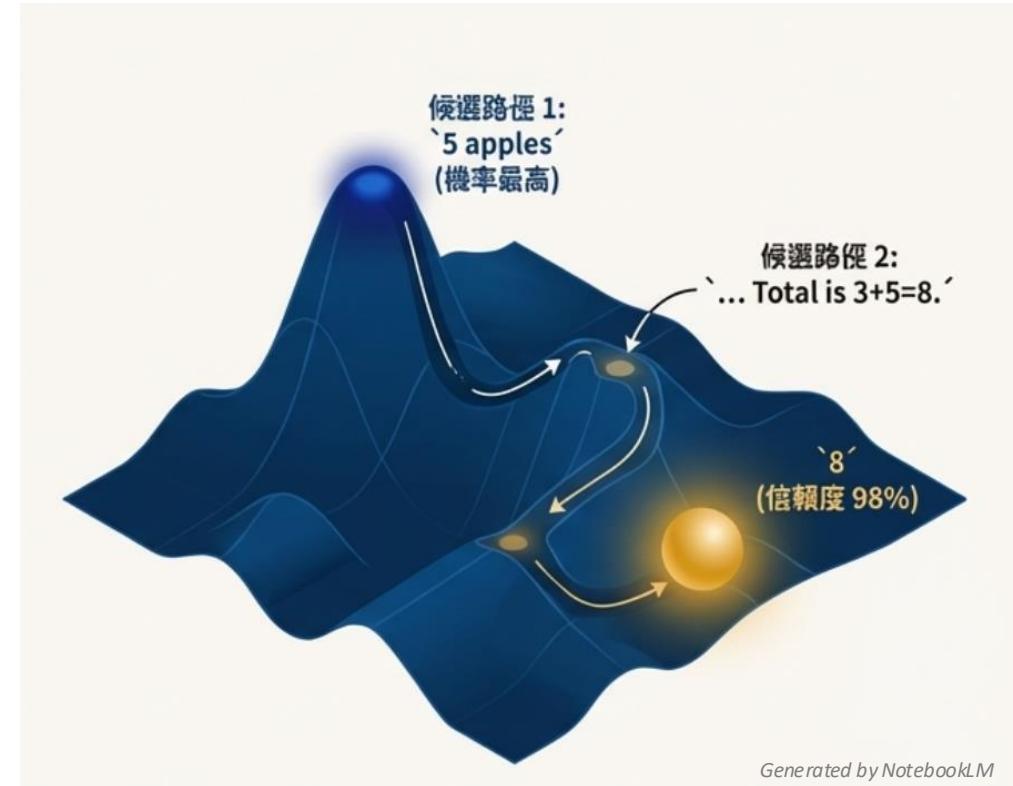
We have 8 apples in total.

You have 3 apples, your dad has 2 more apples than you, so he has 5 apples. $3+5=8$.

The answer is 5.

Chain-of-Thought Decoding

1. Go beyond greedy coding by checking more generation candidates.
2. Choose candidates which have the highest confidence on the final answer.



Prompting change output distribution

Chain-of-Thought Prompting

Q: Elsa has 3 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 $2 + 3 = 5$ apples. So Elsa and Anna have $3 + 5 = 8$ apples together.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Need task-specific examples

Let's Think Step by Step

The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Let's think step by step.

Generic, but perform worse than few-shot

LLM reasoning v.s. Human reasoning

[Question] Using the numbers from 1 to 10 to make 2025. Using each number once and primary operations + and *

[Answer] $(10 * 4 + 5) * (9 * 3 + 8 + 7 + 2 + 1)$

Let's break it down:

- First part: $10 * 4 + 5 = 40 + 5 = 45$
- Second part: $9 * 3 + 8 + 7 + 2 + 1 = 27 + 8 + 7 + 2 + 1 = 45$
- Final Calculation: $45 * 45 = 2025$

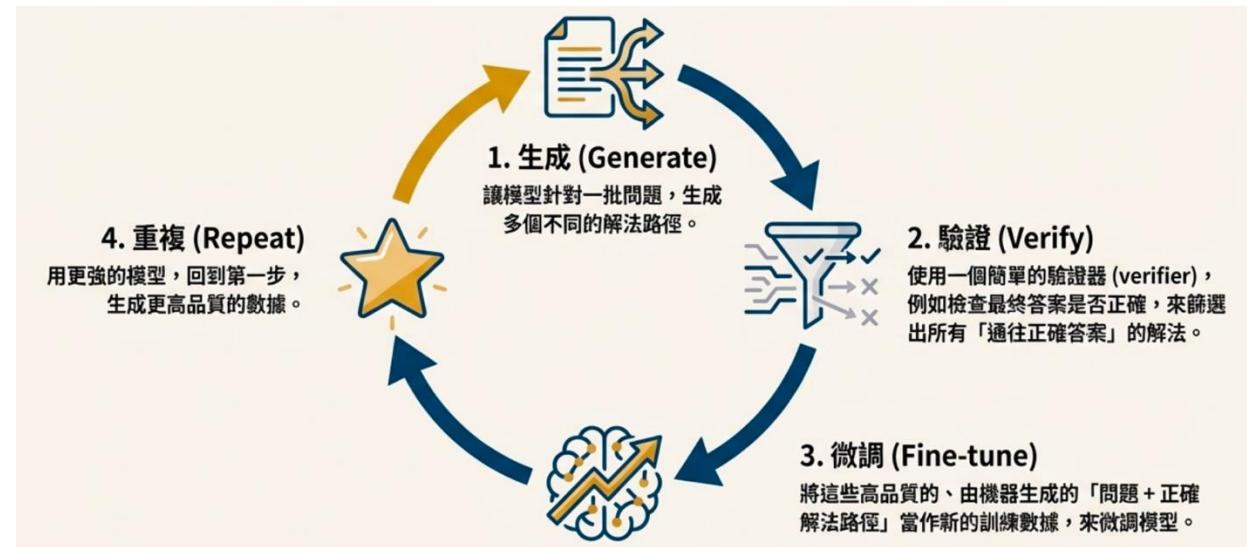
Therefore, the expression $(10 * 4 + 5) * (9 * 3 + 8 + 7 + 2 + 1)$ uses the numbers 1 through 10 exactly once and only the primary operations of addition (+) and multiplication (*) to equal 2025.

Gemini 2.0 thinking mode (Dec, 2024)

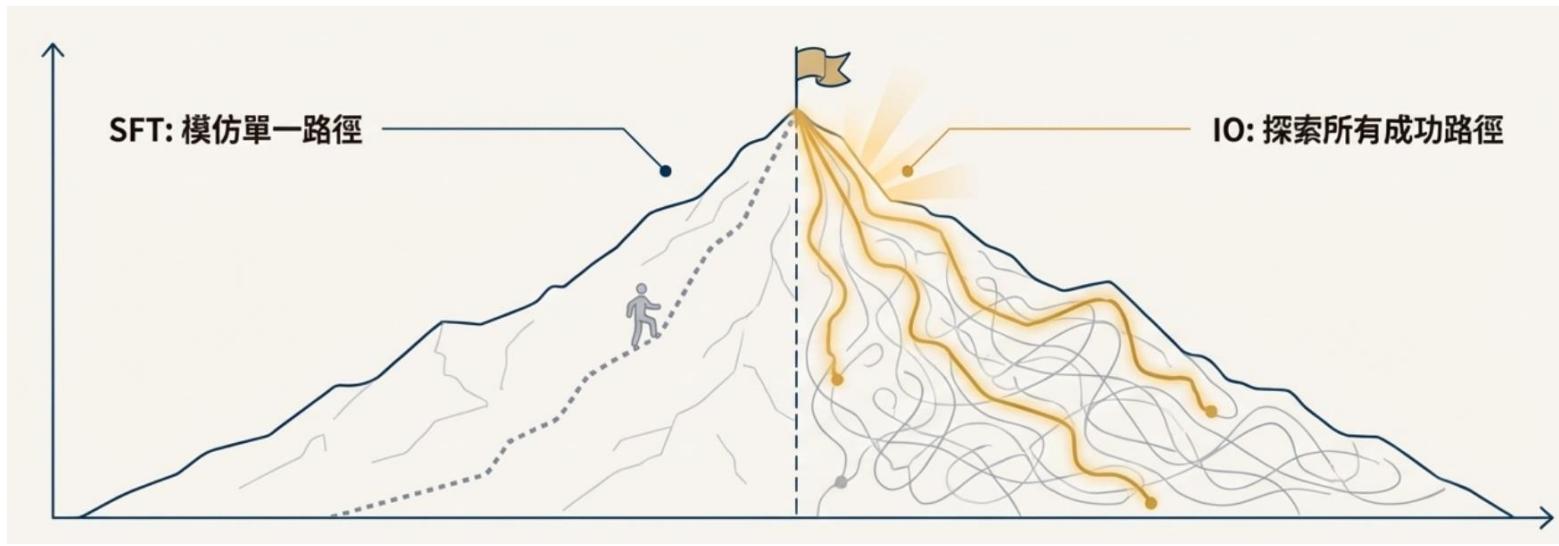
The target number is 2025. This is a relatively large number, suggesting multiplication will be heavily involved. It's also worth noting that 2025 is 45 squared ($45 * 45$). This might or might not be directly helpful, but it gives us a sense of scale. **Prioritize Multiplication:** Since the target is large, start thinking about how to get large intermediate products using multiplication. The largest numbers (10, 9, 8, 7) are good candidates for multiplication early on ... This confirms multiplication is key, but simply multiplying everything together is too much. We need to combine additions. ... Let's aim for products that get us closer to the square root of 2025 (which is 45) ...

Stanford

SFT (human knowledge) v.s. IO fine-tuning



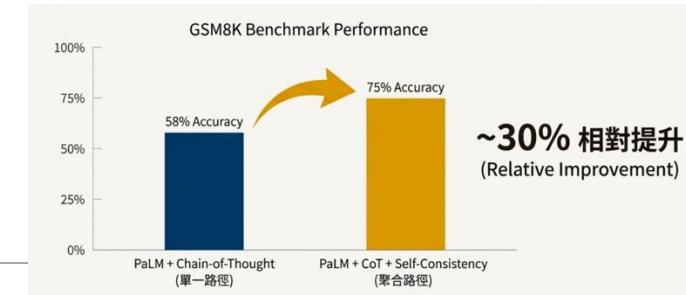
Generated by NotebookLM



Generated by NotebookLM

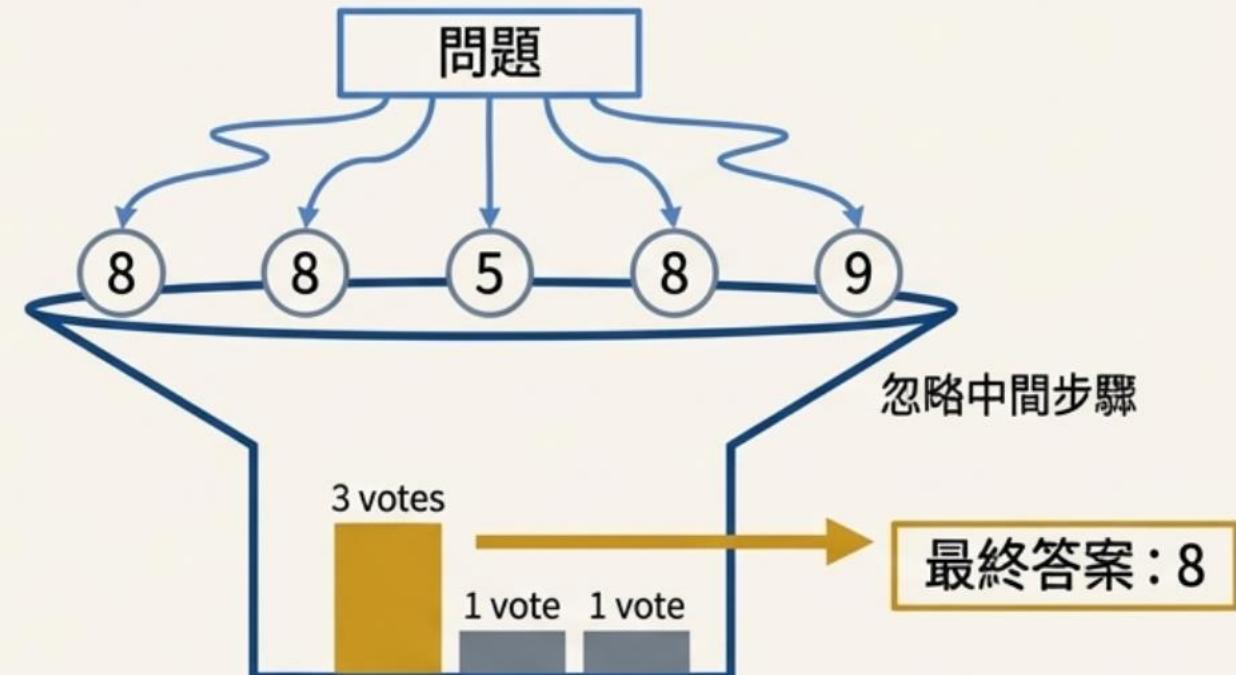
Always keep in mind that LLMs are probabilistic models trained to predict next tokens. **They are not humans!**

Self-consistency Aggregation



The Solution (Self-Consistency)

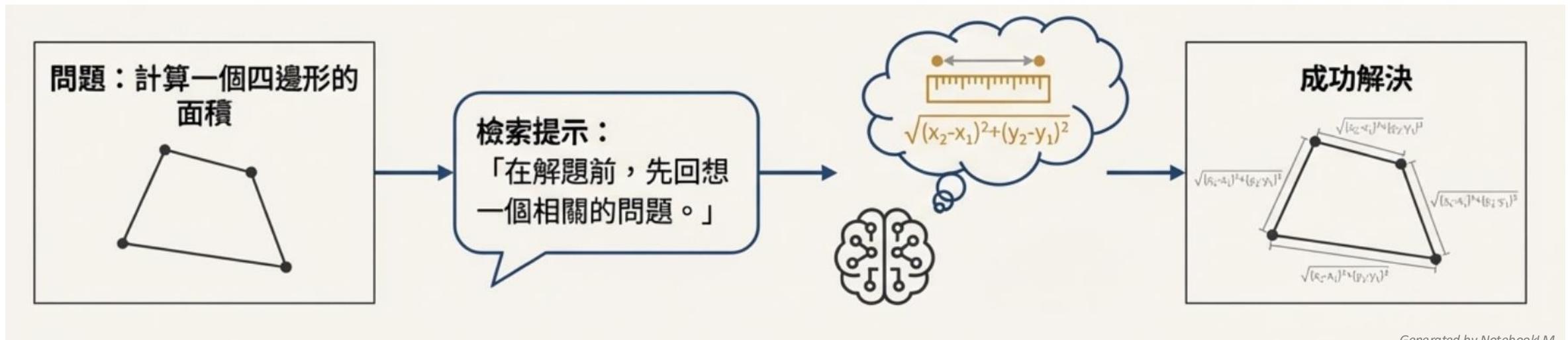
1. 透過隨機採樣 (random sampling)，讓模型生成多個（例如 40 條）不同的推理路徑。
2. 完全忽略中間的推理步驟。
3. 只看最終答案，並選擇出現最多次的那個答案作為最終結果 (majority vote)。



Generated by NotebookLM

Retrieval or reasoning?

Retrieval + Reasoning



Generated by NotebookLM

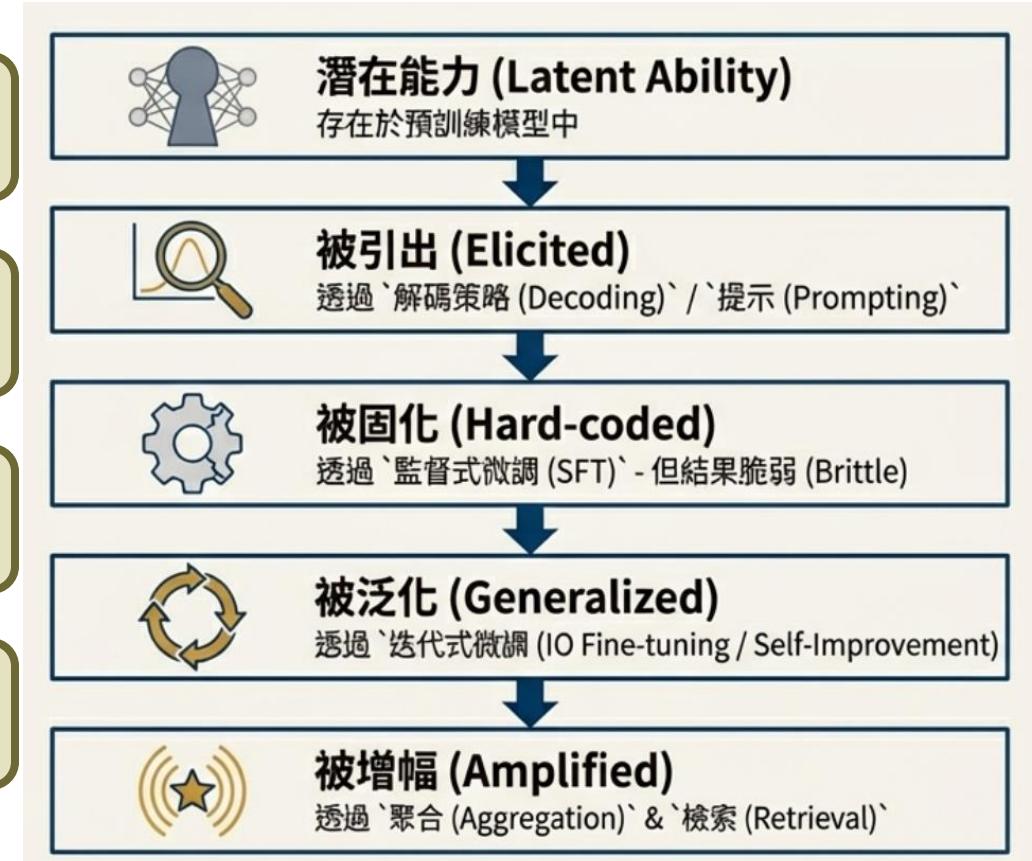
Summary

Reasoning > no Reasoning

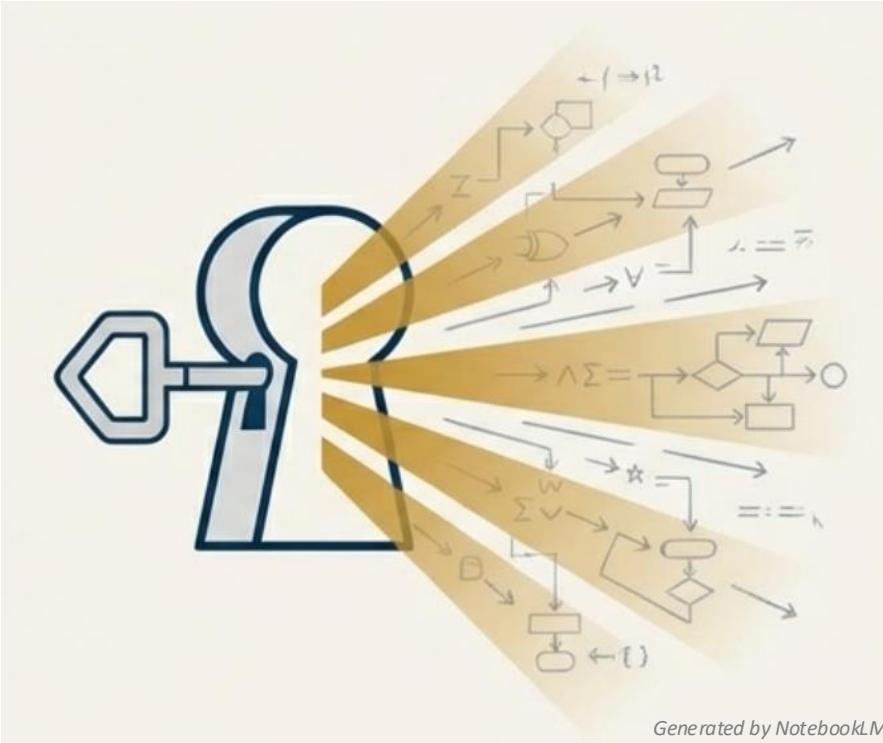
RL fine-tuning > SFT

Aggregating multiple answers > one answer

Retrieval + Reasoning > reasoning only



Generated by NotebookLM



Next-token prediction

→ emerging intelligence

The truth always turns out to be simpler than you thought.

- Richard Feynman