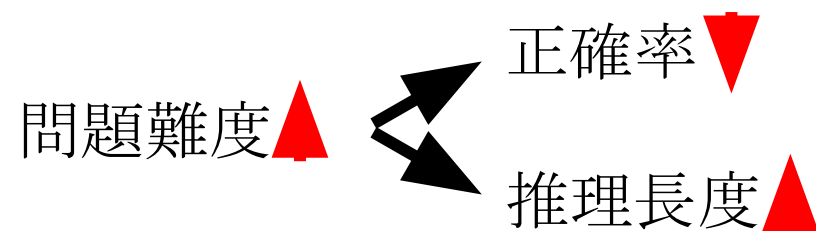
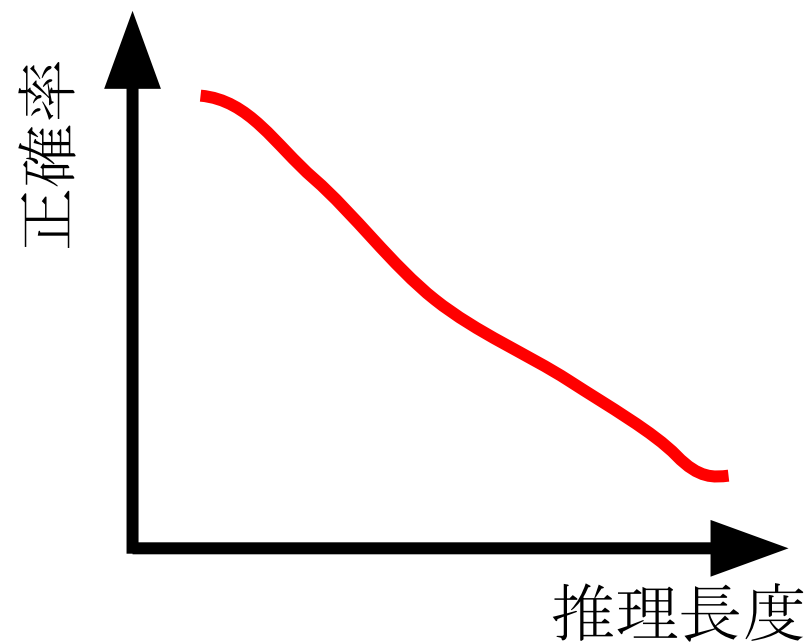
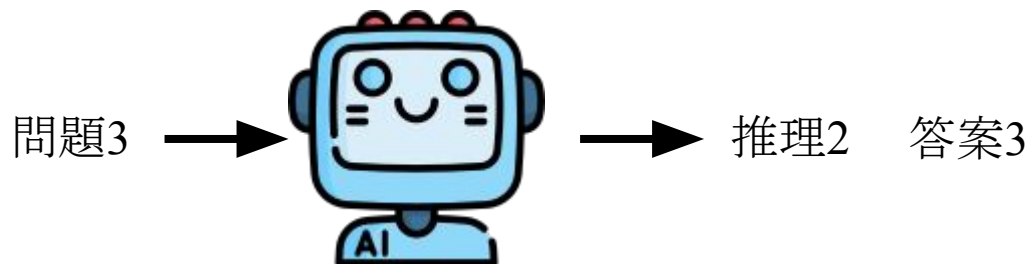
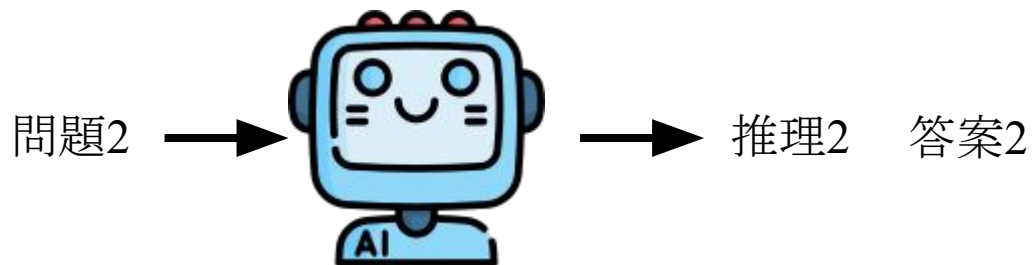
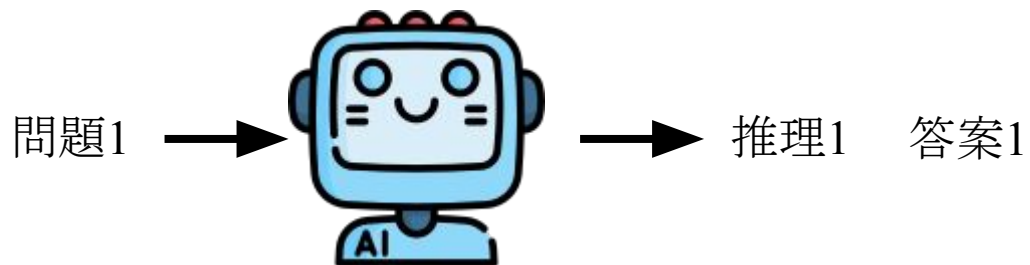


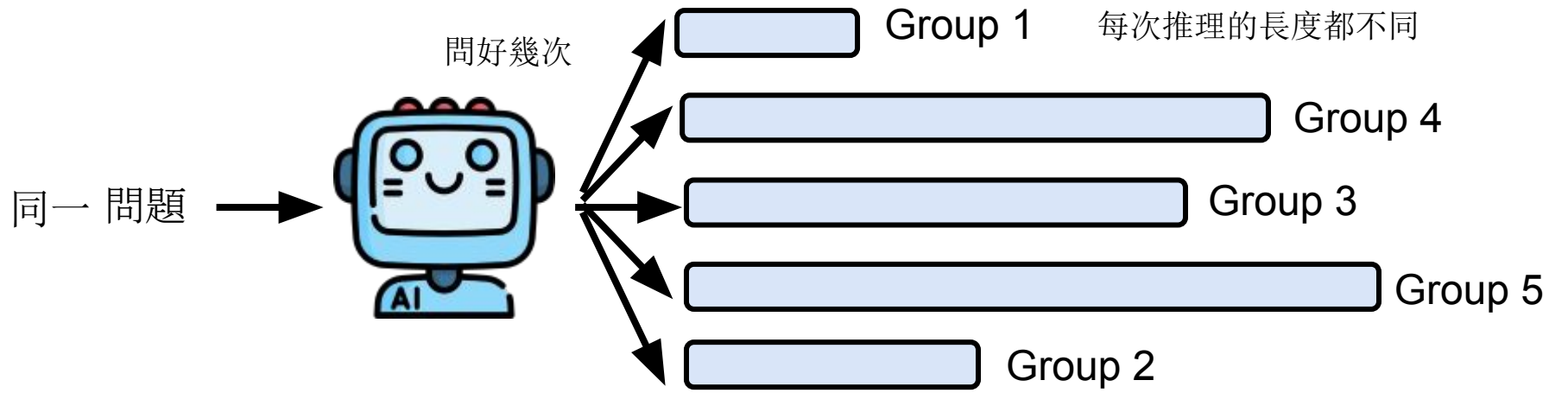
別讓「推理」 大型語言模型想太多

避免他浪費一堆運算資源

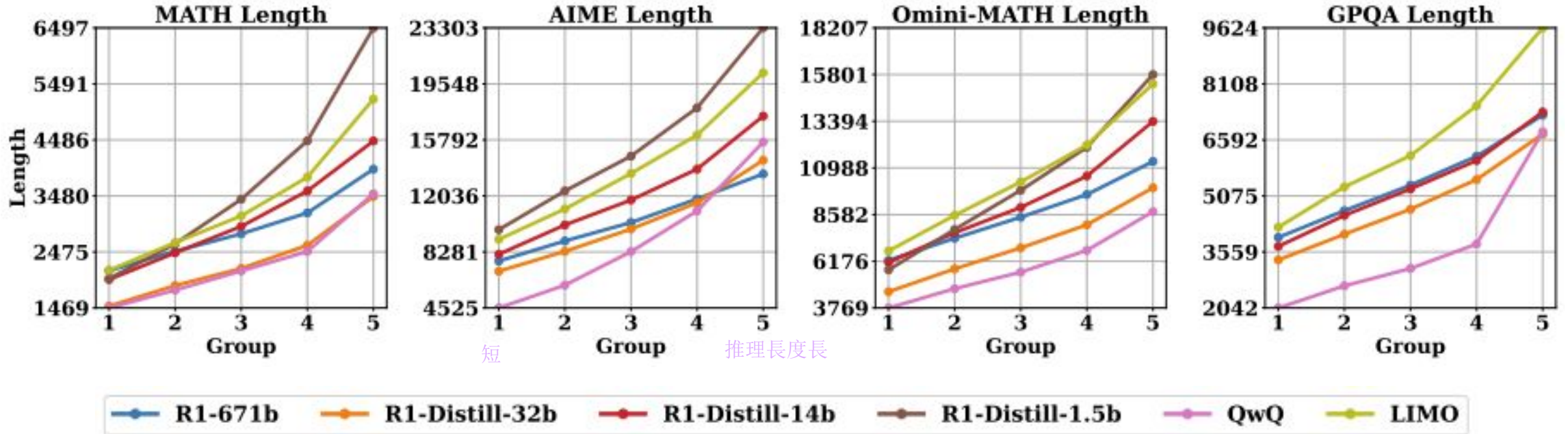
2025/05/02

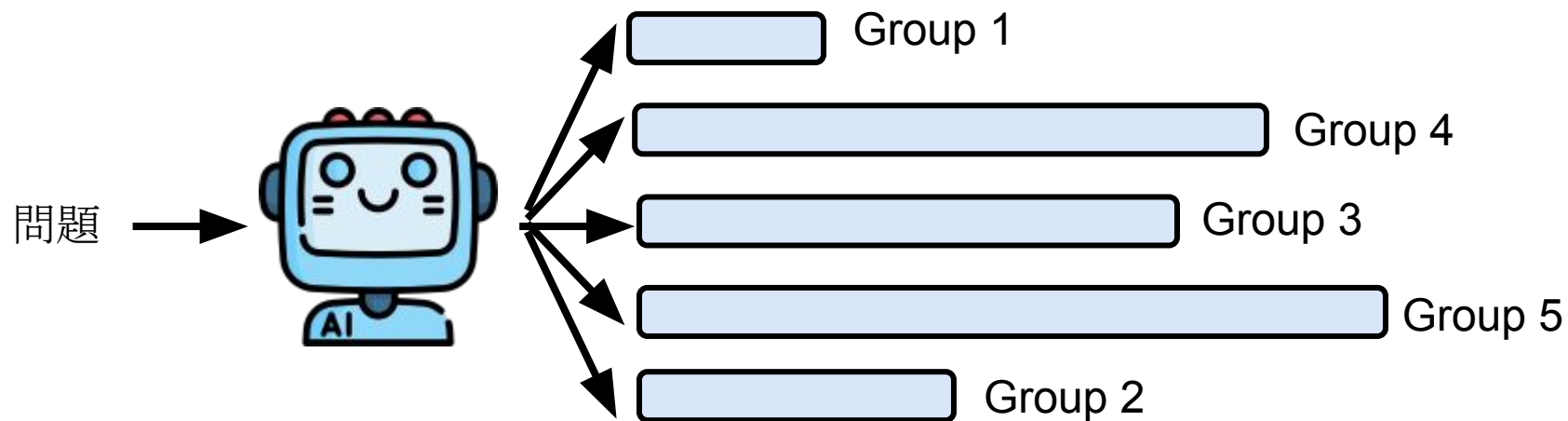
「推理」越長、結果越好？ 不一定



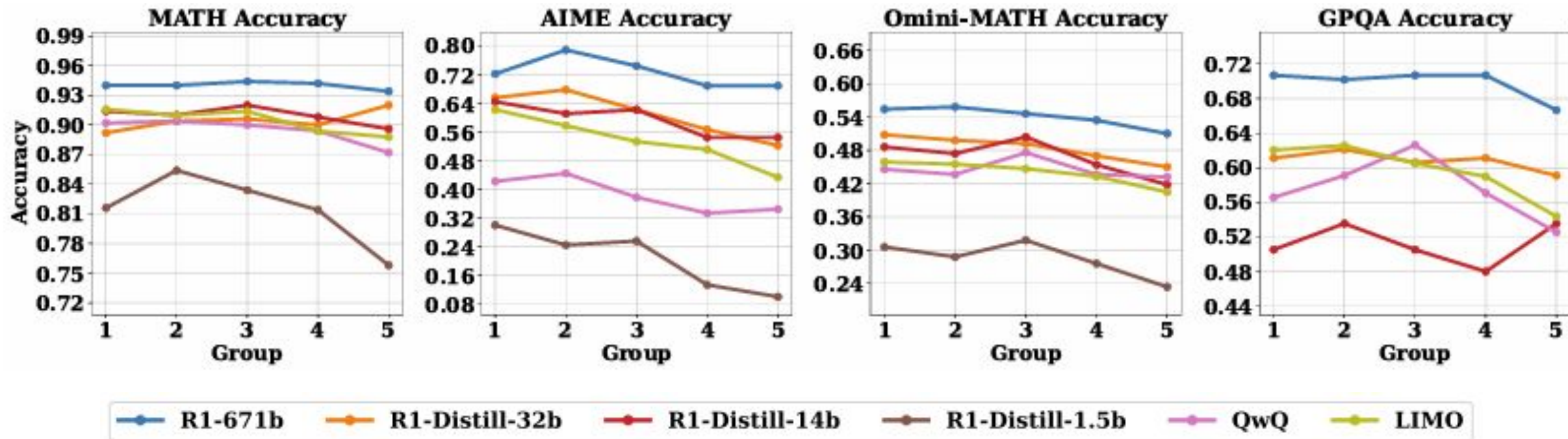


數學競賽





不一定越正確



人工智慧也是一樣

最好的工程師

不是把事情做到完美

而是在有限資源下

把事情做到最好

如何避免「想太多」

更強的思維鏈 (Chain-of-Thought, CoT)

給模型推論工作流程

教模型推理過程 (Imitation Learning)

以結果為導向學習推理 (Reinforcement Learning, RL)

就是Chain of Thought概念，但要用較短的草稿

Chain of Draft

<https://arxiv.org/pdf/2502.18600>

Standard

Answer the question directly. Do not return any preamble, explanation, or reasoning.

Chain-of-Thought

Think step by step to answer the following question. Return the answer at the end of the response after a separator #####.

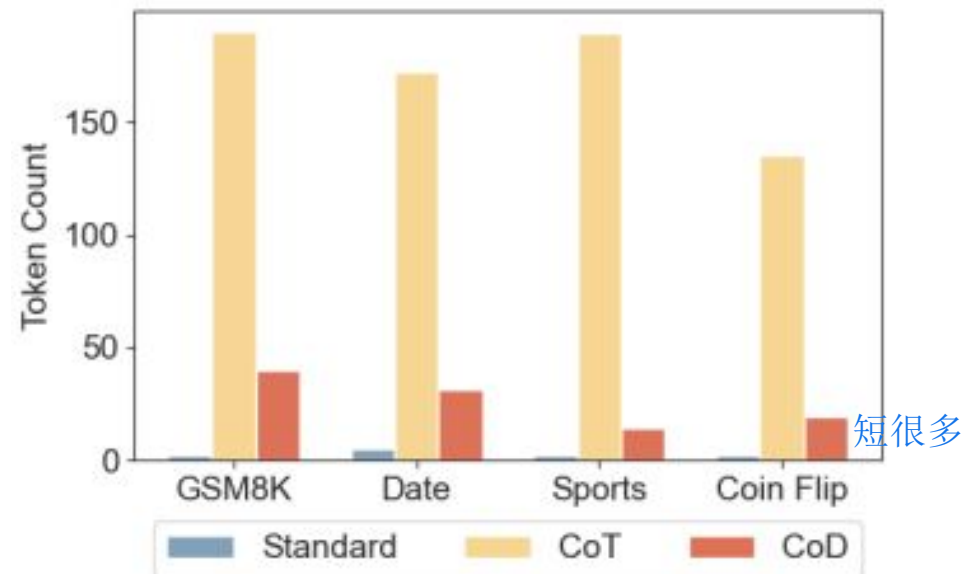
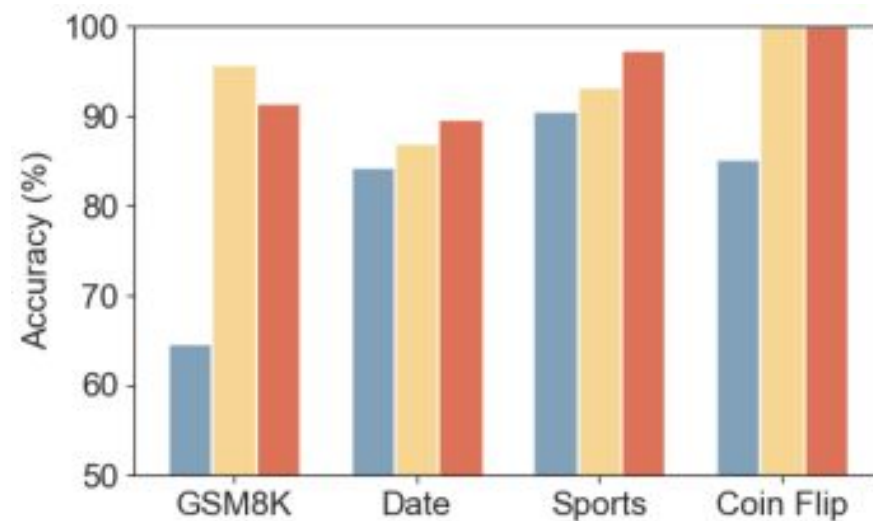
Chain-of-Draft

Think step by step, but only keep a minimum draft for each thinking step, with 5 words at most. Return the answer at the end of the response after a separator #####.

簡單，但結果滿有用的

藍:不做CoT
膚:做CoT
紅:做CoD

Claude 3.5 Sonnet



如何避免「想太多」

更強的思維鏈 (Chain-of-Thought, CoT)

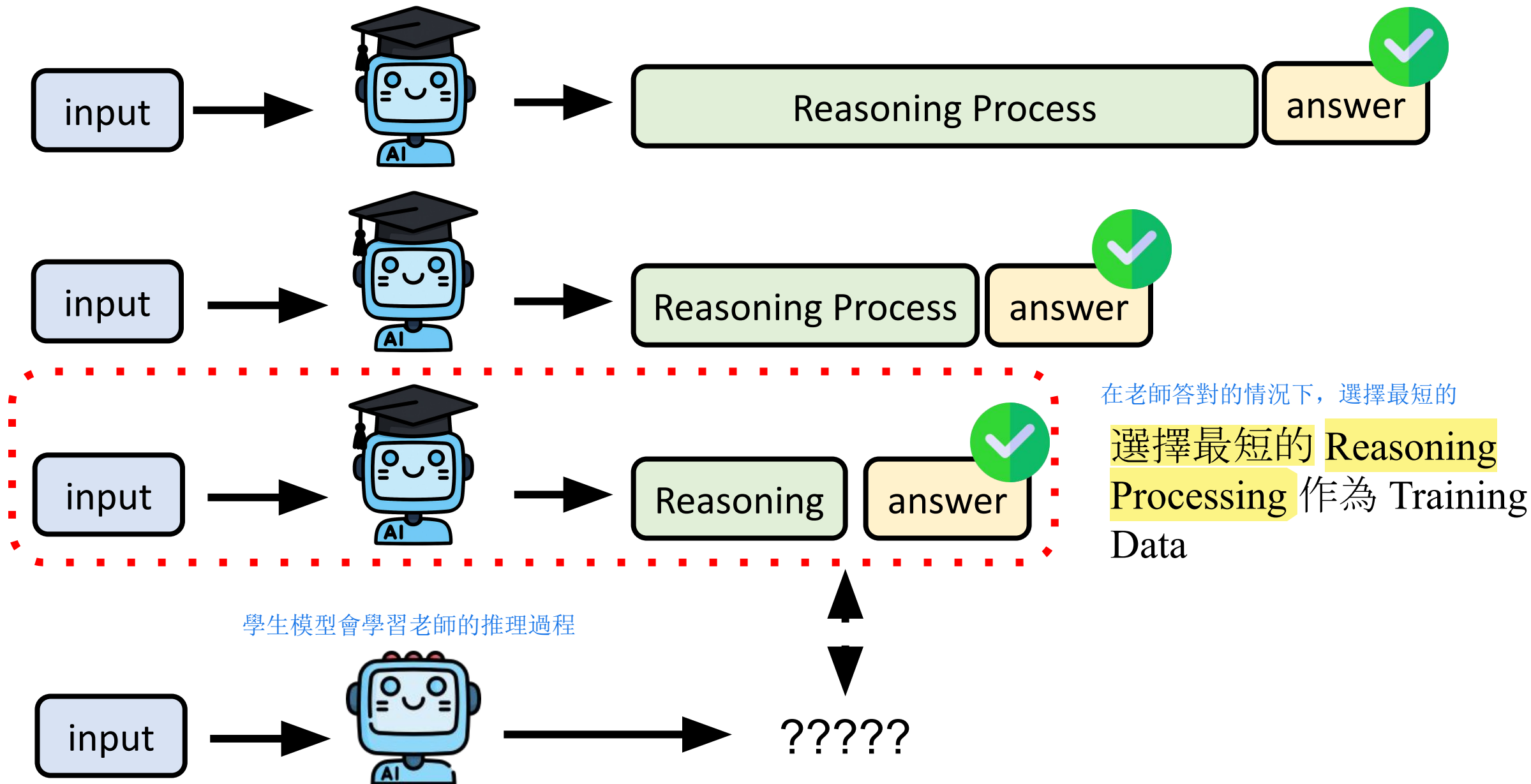
給模型推論工作流程

由人設定工作流程
ex. sampling少一點, beam search時beam少一點等

教模型推理過程 (Imitation Learning)

現在已有ai教ai怎麼做推理

以結果為導向學習推理 (Reinforcement Learning, RL)



對於像是乘法這種簡單的數學有用

From Explicit CoT to Implicit CoT

若直接省去CoT就產生答案，效果不佳

把reasoning過程的第一個token拿掉

		Input					
Explicit CoT	Stage 0:	2	1	×	4	3	=
	Stage 1:	2	1	×	4	3	=
	Stage 2:	2	1	×	4	3	=
	Stage 3:	2	1	×	4	3	=
	Stage 4:	2	1	×	4	3	=
	Stage 5:	2	1	×	4	3	=
Implicit CoT	Stage 6:	2	1	×	4	3	=

		CoT					Output		
		8	4	+	0	6	3	=	8 0 4
		4	+	0	6	3		=	8 0 4
			+	0	6	3		=	8 0 4
				0	6	3		=	8 0 4
					6	3		=	8 0 4
						3		=	8 0 4
								=	8 0 4

讓模型產生少了一個token的reason過程，再產生答案

用漸進式學法
減少模型推理的過程展現

<https://arxiv.org/abs/2405.14838>

如何避免「想太多」

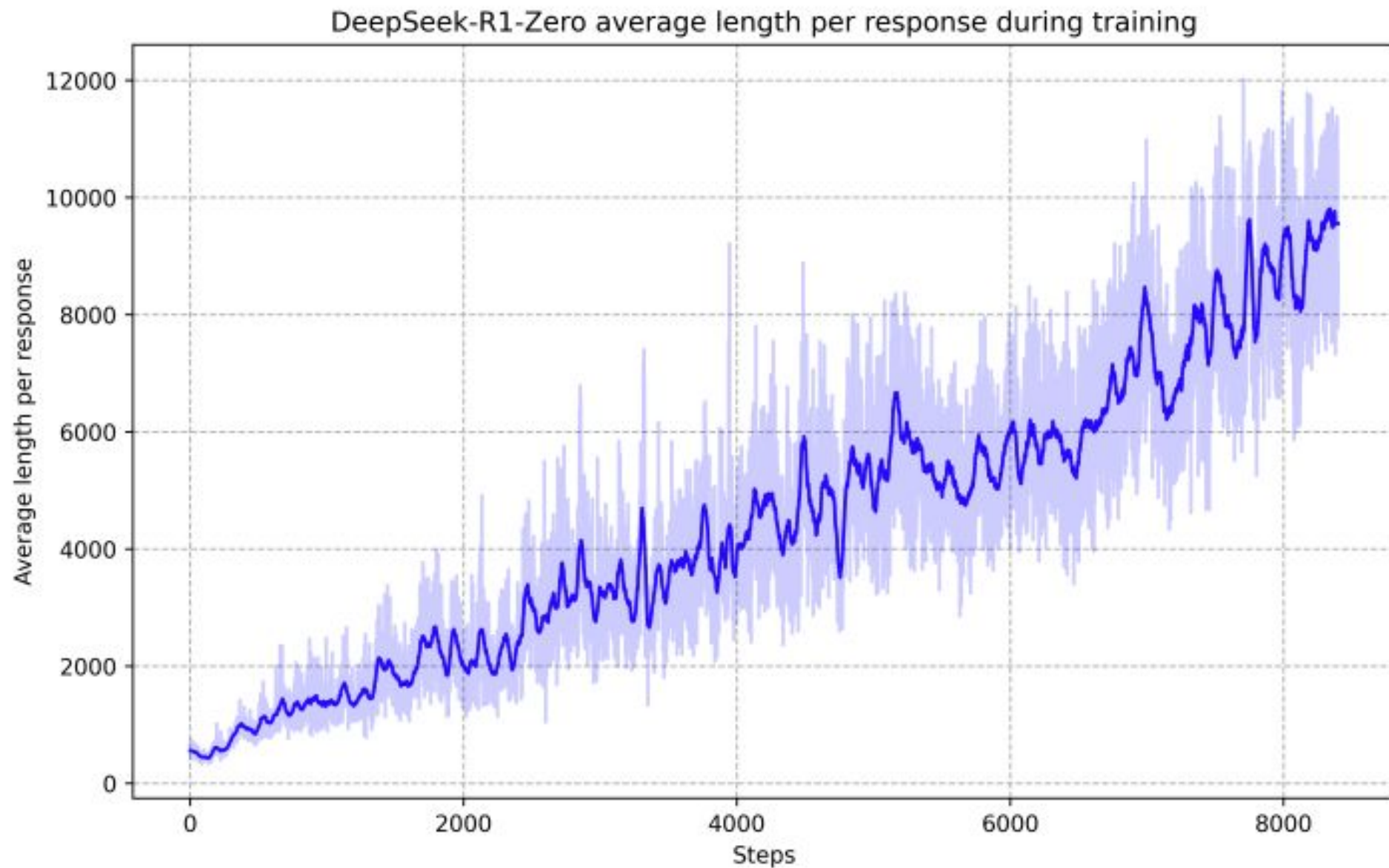
更強的思維鏈 (Chain-of-Thought, CoT)

給模型推論工作流程

教模型推理過程 (Imitation Learning)

以結果為導向學習推理 (Reinforcement Learning, RL)

HW. 7會做RL的作業



DeepSeek-R1

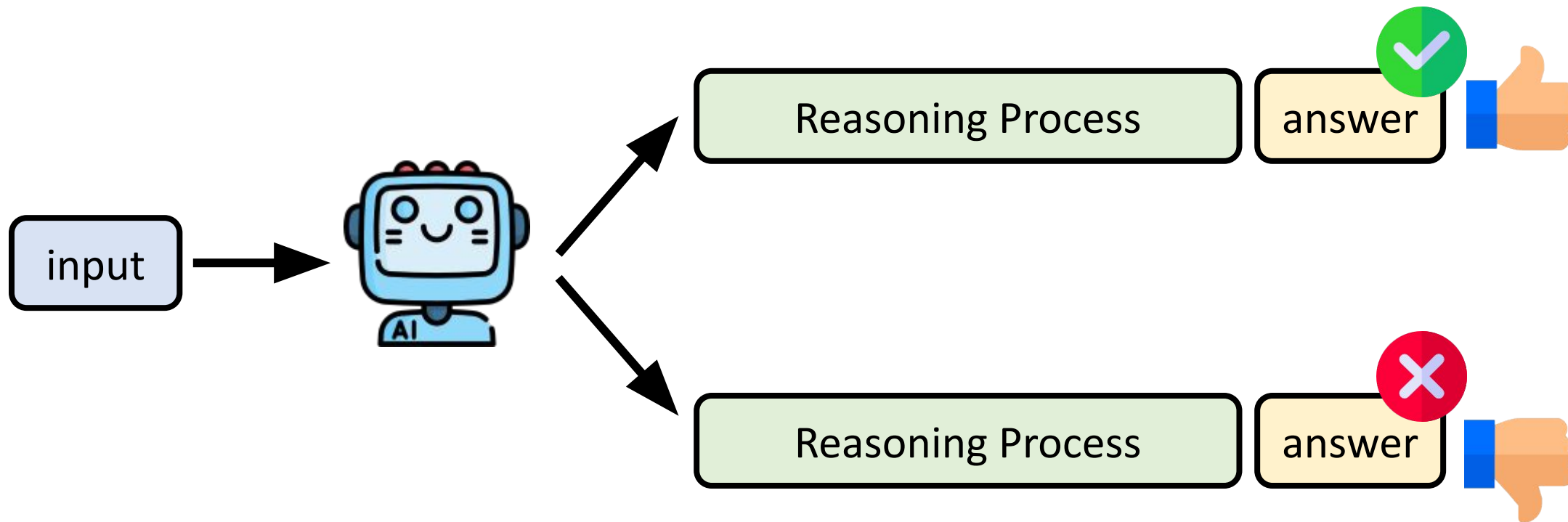
<https://arxiv.org/abs/2501.12948>

deepseek論文顯示隨著r1訓練進行，模型輸出會越長

因為在RL訓練過程，沒有告訴模型長度不要太長

RL 的方法產生了超長的「推理」過程

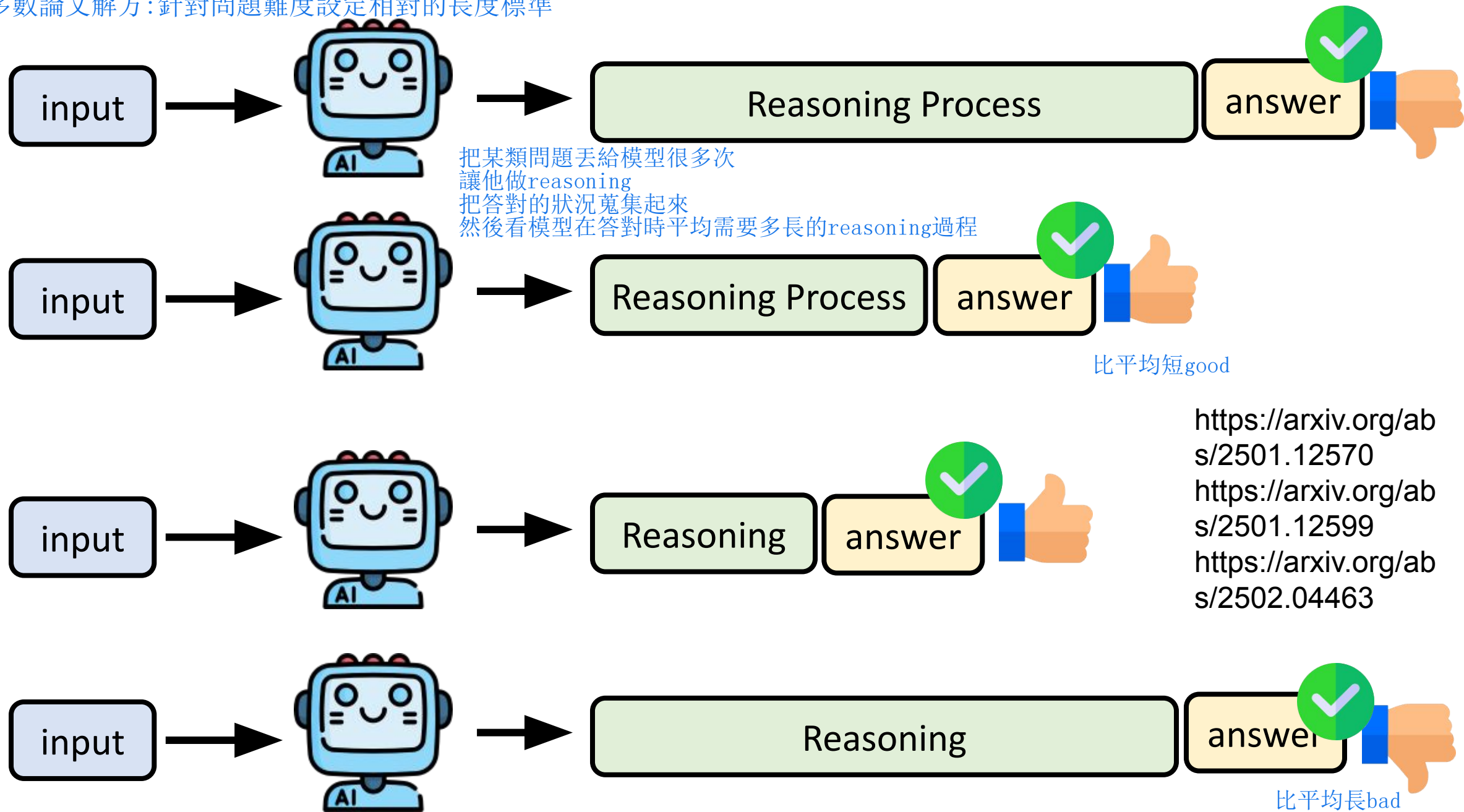
通常只會對答案給予反饋
reward: 正確率



沒人教 AI 要在意 Reasoning Process 的長度！

超過一定長度就算是不好的？ 這樣太僵化

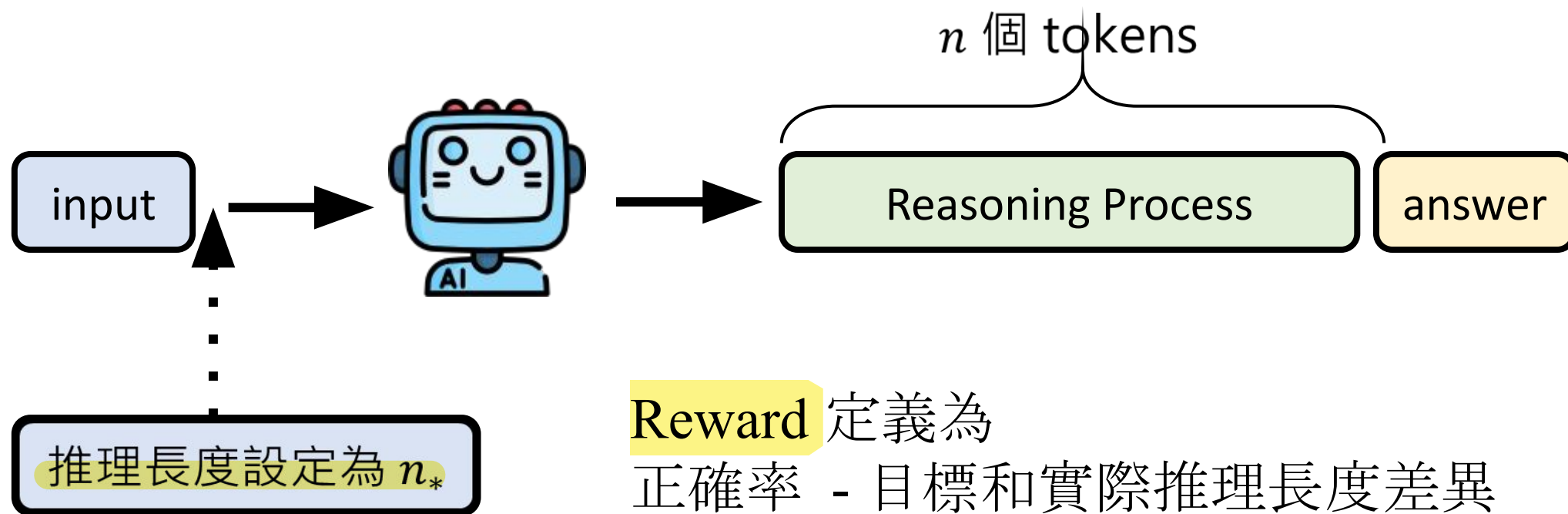
多數論文解方: 針對問題難度設定相對的長度標準



也可以教模型

控制「推理」的長度

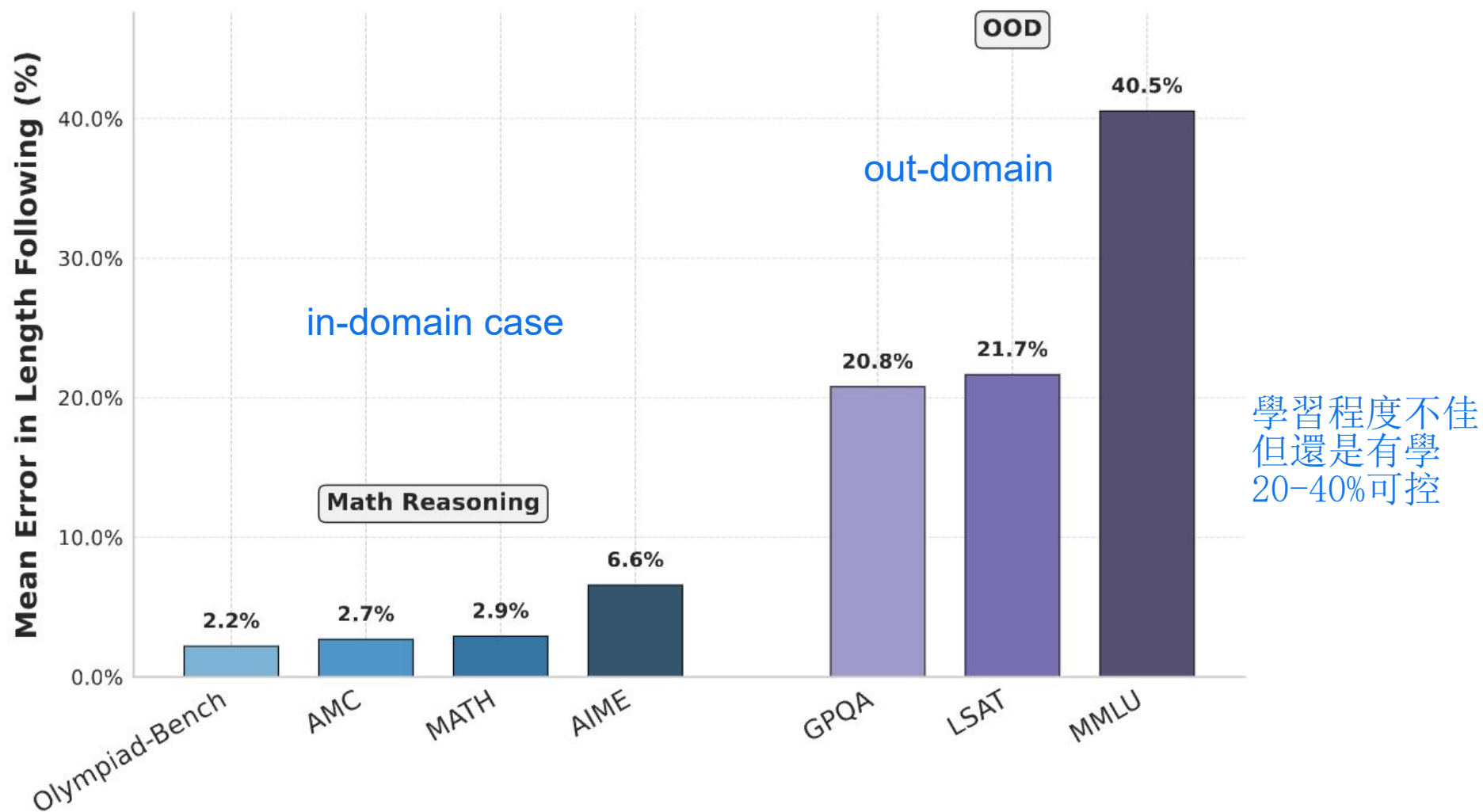
<https://arxiv.org/abs/2503.04697>



控制「推理」的長度

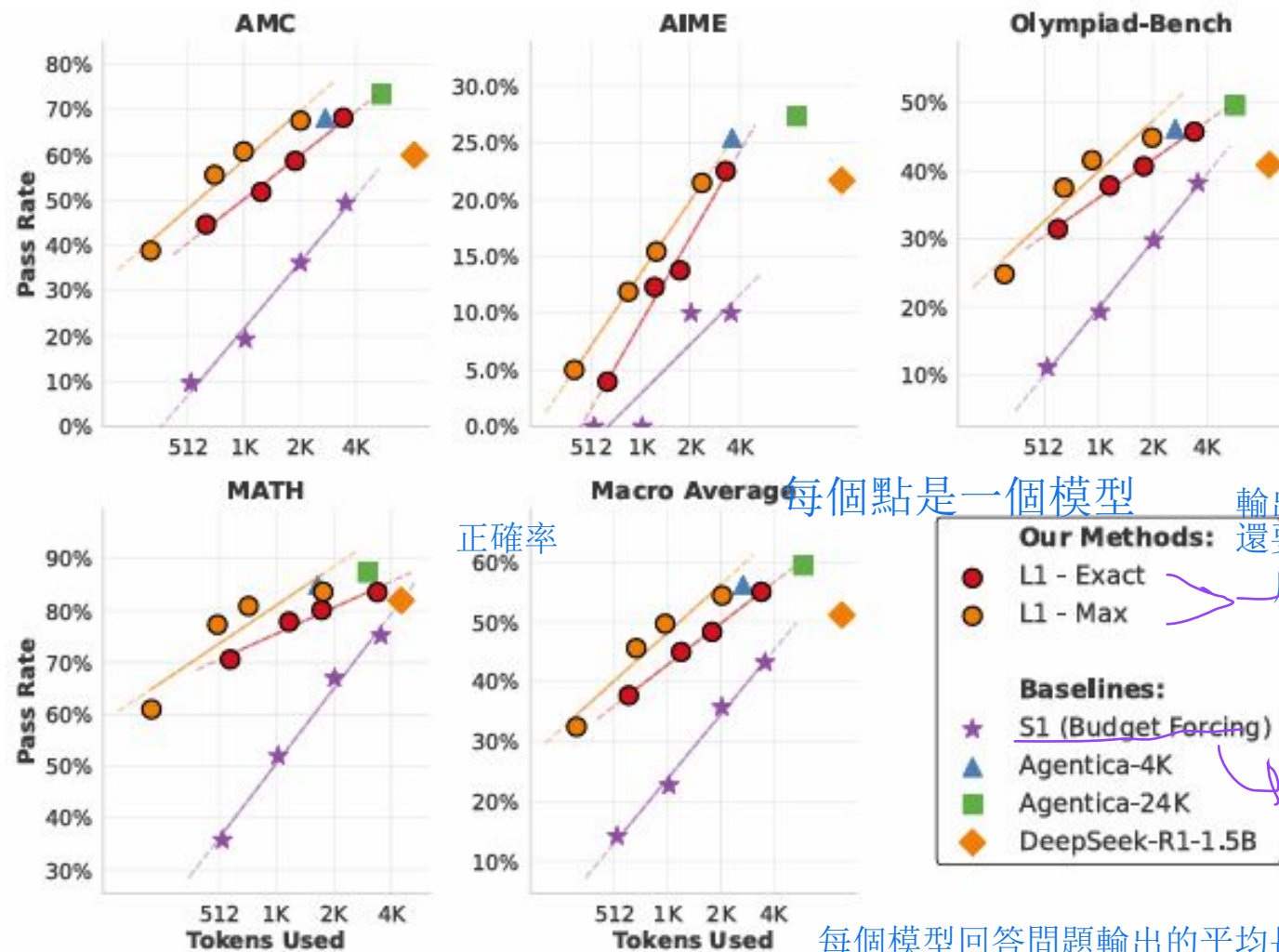
看起來可行

<https://arxiv.org/abs/2503.04697>



控制「推理」的長度

<https://arxiv.org/abs/2503.04697>



exact: 剛好xx字
max: 最多xx字

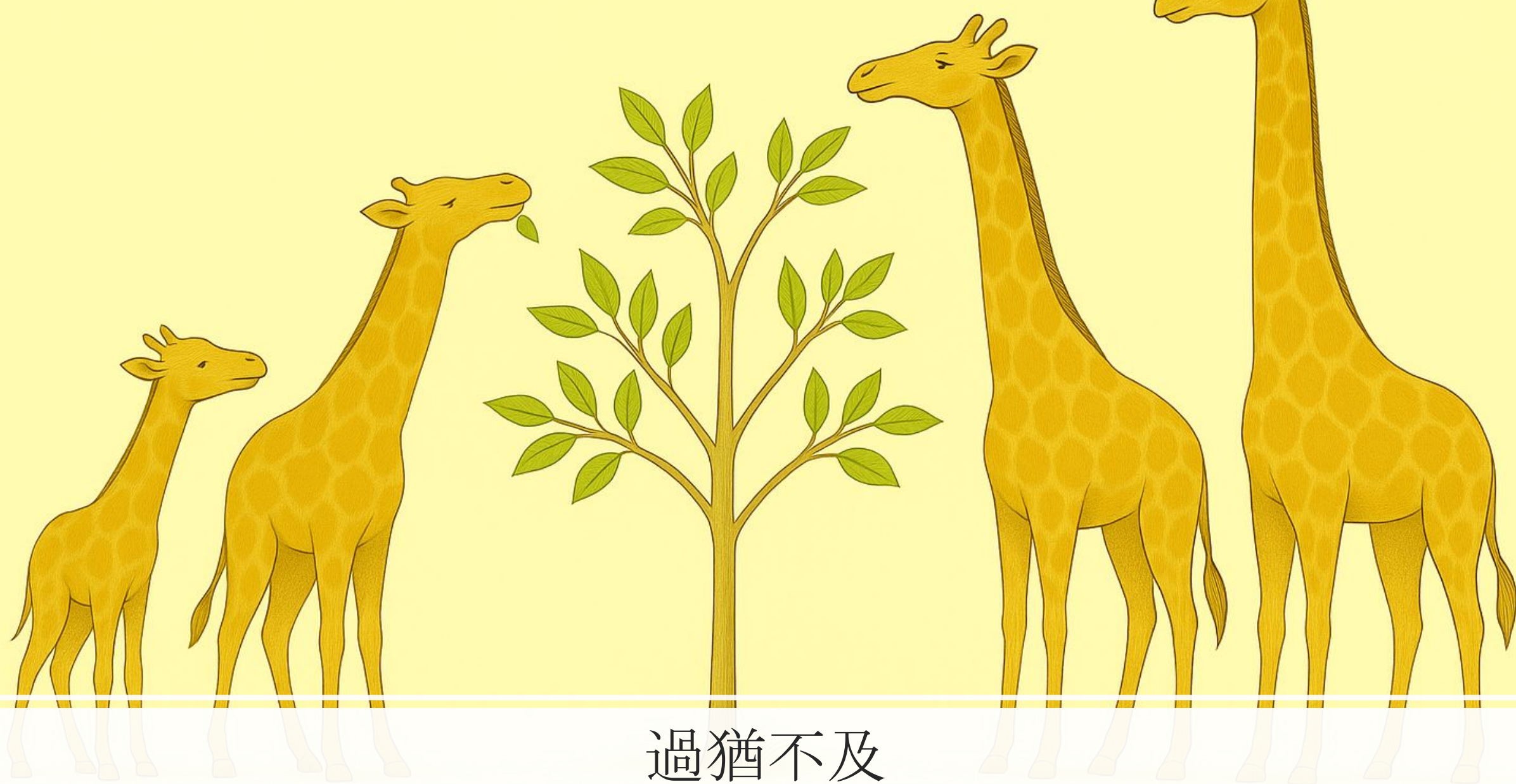
每個點是一個模型

輸出短時，結果比s1強制結束的方法還要好

用r1控制模型輸出長度

對gemini做knowledge distillation的模型n得到，在本篇paper中有控制輸出長度
粗暴方法：模型講太長就cut

每個模型回答問題輸出的平均長度



過猶不及

reasoning 長度不是越長越好

(本圖由 gpt-4o 生成, 與實際演化過程無關)