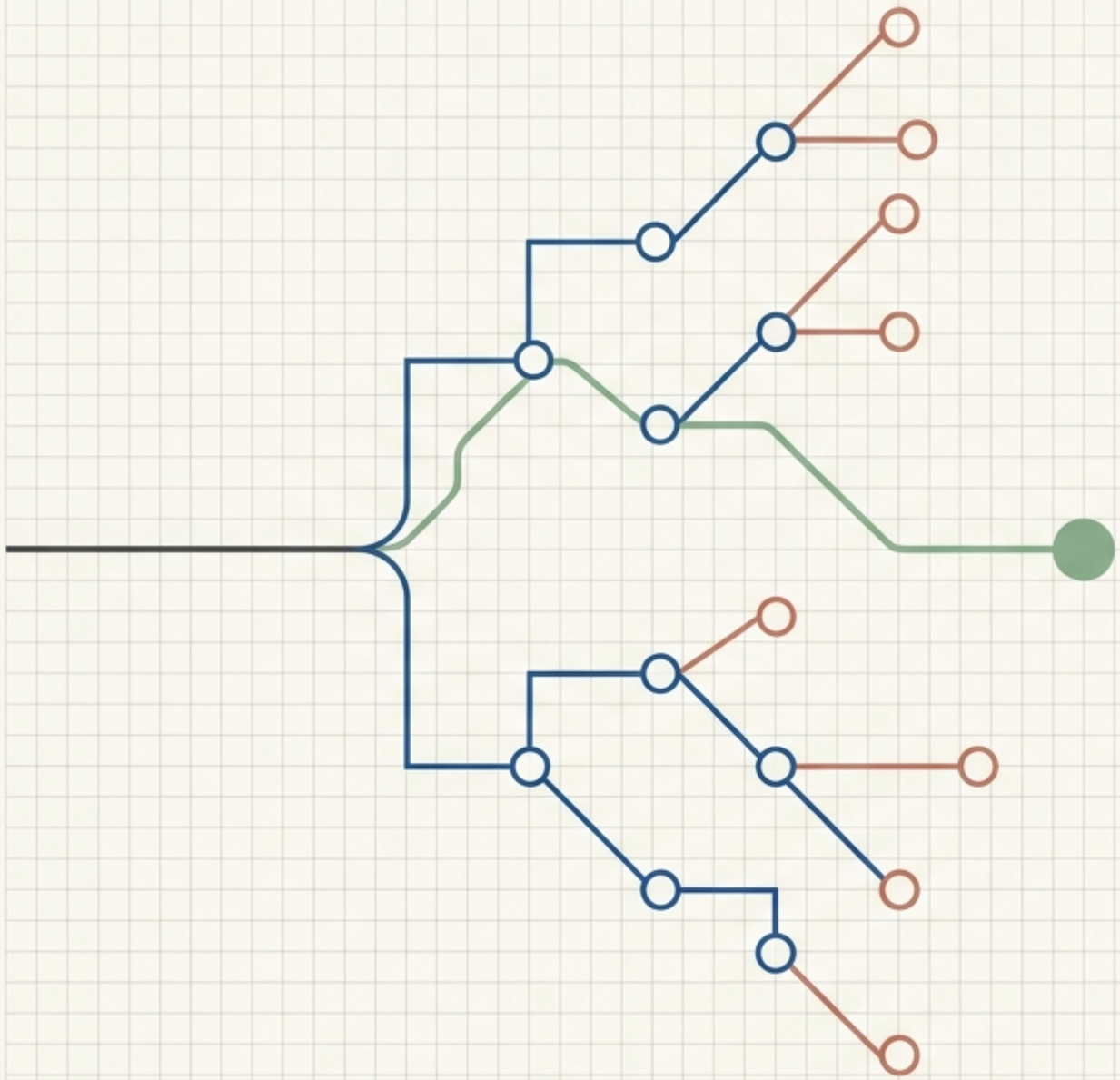


สู่ยุคแห่งการคิดวิเคราะห์เชิงลึกของ AI

การสอนให้ AI เรียนรู้วิธีคิดผ่าน Meta Chain-of-Thought (Meta-CoT) และ In-Context Search

สรุปงานวิจัยฉบับสมบูรณ์: ทำไมโมเดลระดับแนวหน้า (o1, DeepSeek-R1) ถึงแก้โจทย์ที่ซับซ้อนได้?

NotebookLM

ทำไมโมเดลเก่งๆ ถึงตกม้าตายในโจทย์ยาก?

The Linear Illusion

$$\frac{A \cdot B}{C} = D + E$$

Where $A = \int (x^2 + 1)\, dx$

$$B = \sum_{n=1}^{\infty} \frac{1}{n^2}$$

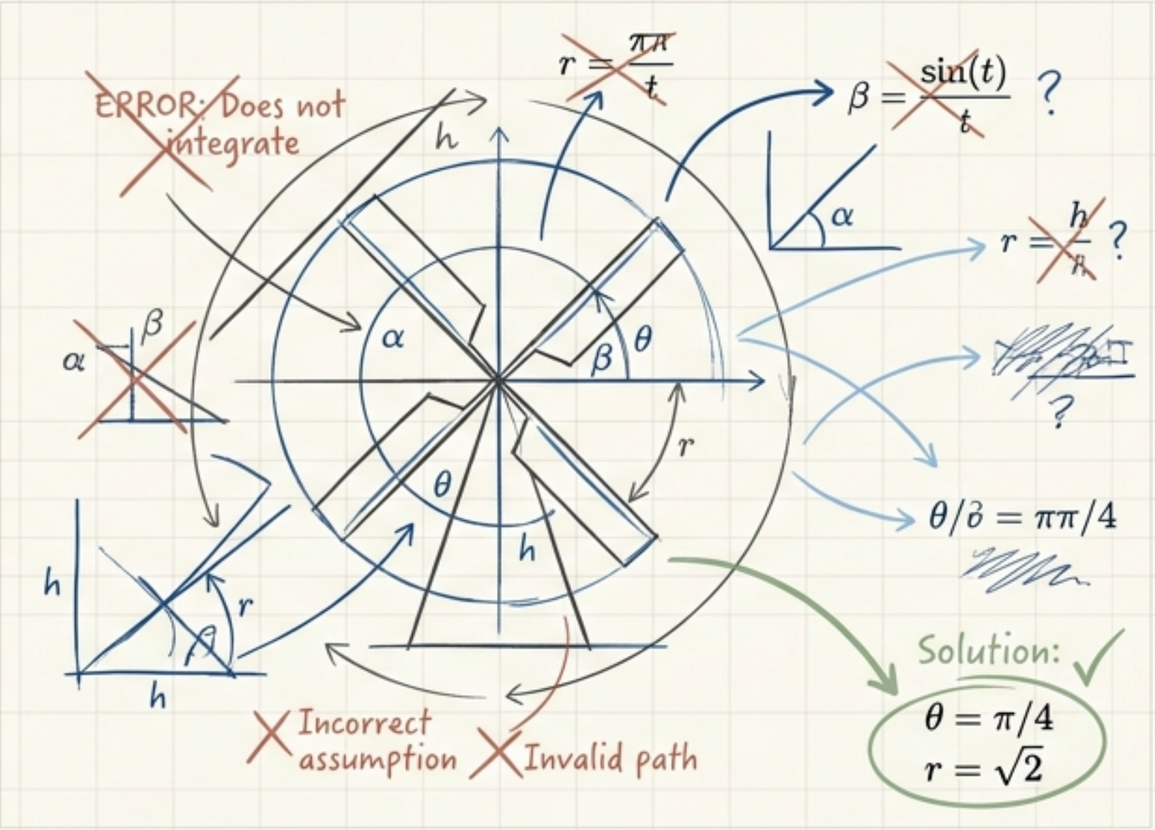
$$C = \sqrt{|-4| + 5}$$

$$D = e^{\pi i}$$

$$E = \lim_{t \to 0} \frac{\sin(t)}{t}$$

Thus, $\frac{A \cdot B}{C} = D + E$ holds true.

Latent Human Reasoning



The Illusion of Pre-training

AI ถูกฝึกด้วยคำตอบที่สมบูรณ์แบบจากตำรา แต่กระบวนการคิดจริงๆ ของมนุษย์ (Latent Reasoning) ไม่ได้เป็นเส้นตรงตั้งแต่ต้นจนจบ

The Failure of Standard CoT

การสั่งให้ AI 'Think step-by-step' ได้ผลแค่กับโจทย์ง่าย เพราะมันบังคับให้ AI สร้างคำตอบแบบจากซ้ายไปขวาโดยไม่เปิดโอกาสให้ลองผิดลองถูก

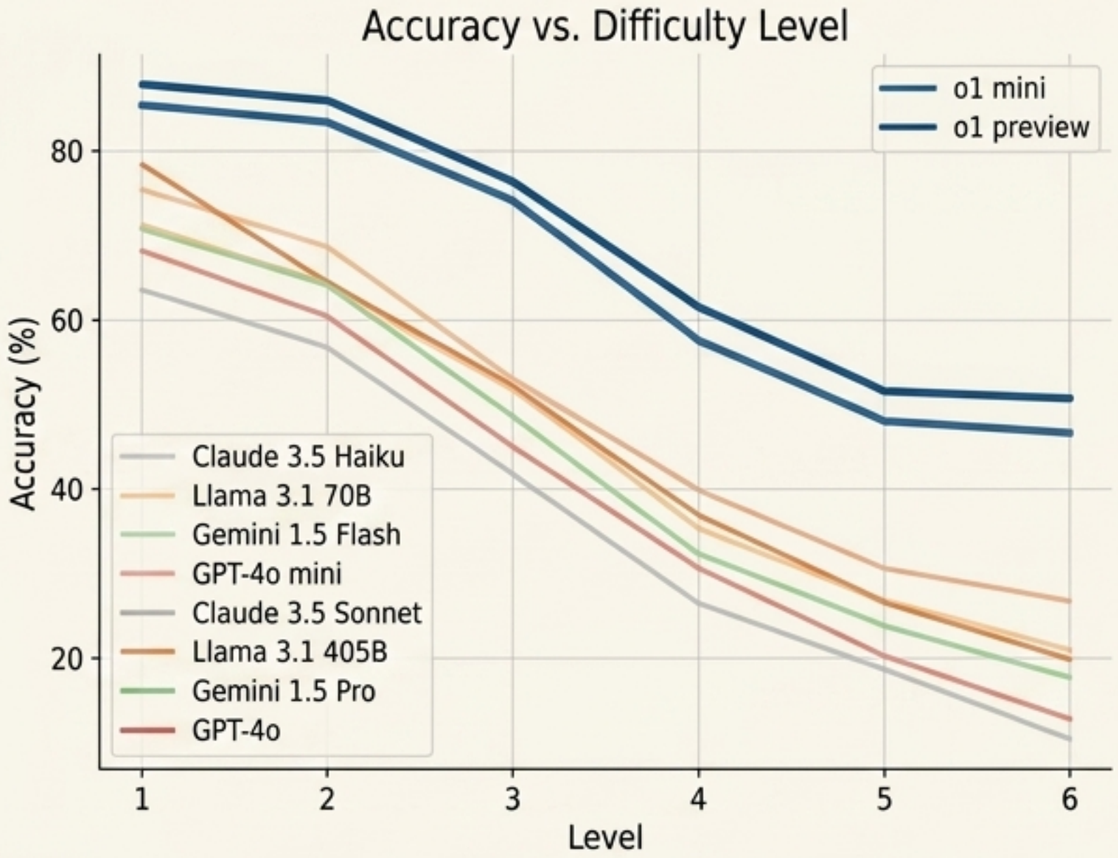
NotebookLM

การเปลี่ยนผ่านกระบวนทัศน์: System 1 สู่ System 2

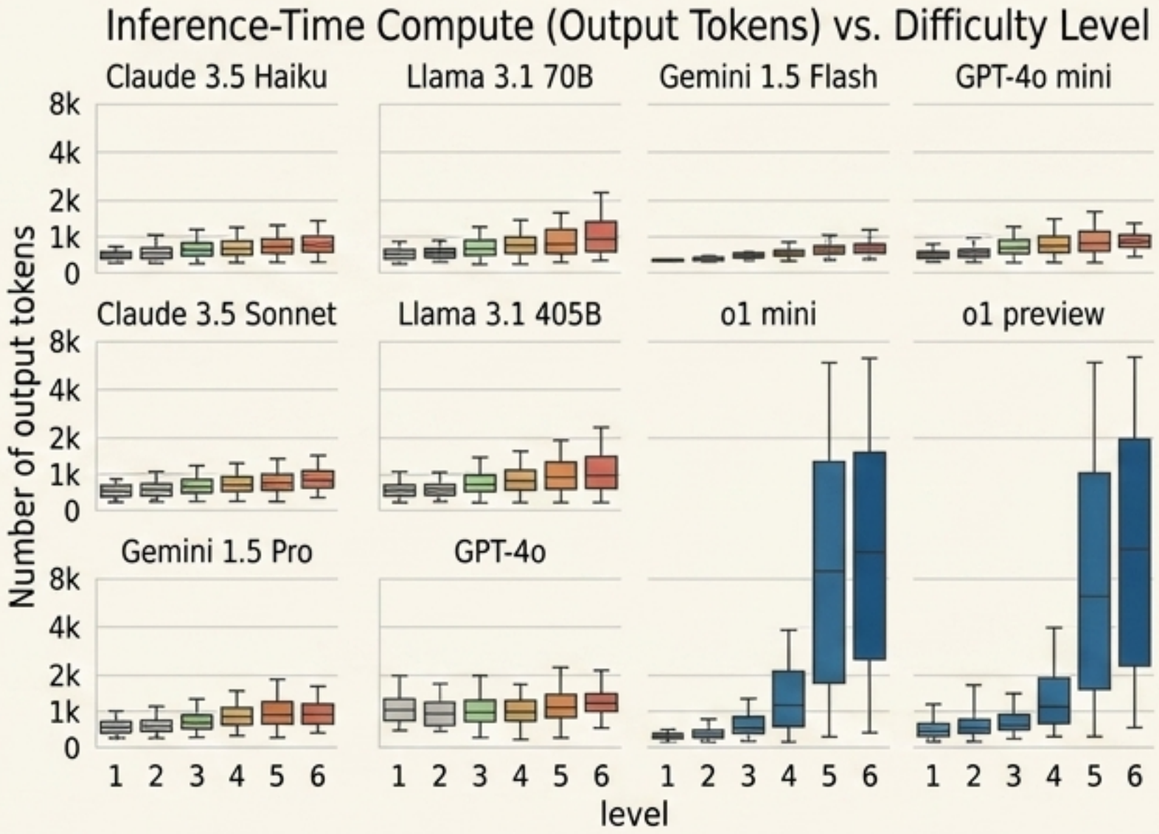
	System 1 (Standard CoT)	System 2 (Meta-CoT)
รูปแบบเส้นทาง (Trajectory)	เส้นตรงเส้นเดียว (Linear)	แตกกิ่งก้านสาขาและย้อนกลับ (Tree/Maze Search)
การใช้พลังงานประมวลผล (Compute)	ใช้เท่าเดิมเสมอไม่ว่าโจทย์จะยากแค่ไหน	ยิ่งโจทย์ยาก ยิ่งใช้เวลาและ **Compute** เยอะขึ้น (Inference-Time Compute)
พฤติกรรมที่เทียบเคียง (Cognitive Equivalent)	การท่องจำและการตอบสนองทันที	การทดลอง ตั้งสมมติฐาน และตรวจสอบ (Deliberate Reasoning)
กลไกหลัก (Core Mechanism)	Next-token prediction	Meta Chain-of-Thought (Meta-CoT)

NotebookLM

ความลับของ Frontier Models: ยิ่งยาก ยิ่งต้องคิดนาน



โมเดลทั่วไปความแม่นยำดิ่งลงเมื่อเจอโจทย์ซับซ้อน แต่ตระกูล o1 ยังคงรักษาความแม่นยำไว้ได้



ข้อมูลสะท้อนชัดเจน: โมเดลตระกูล o1 ใช้เวลาสร้าง Token เพื่อคิดวิเคราะห์ (Inference-Time Compute) มากเป็นเงตามตัวตามระดับความยาก

NotebookLM

กุญแจสำคัญ: ช่องว่างระหว่างการสร้างและการตรวจ

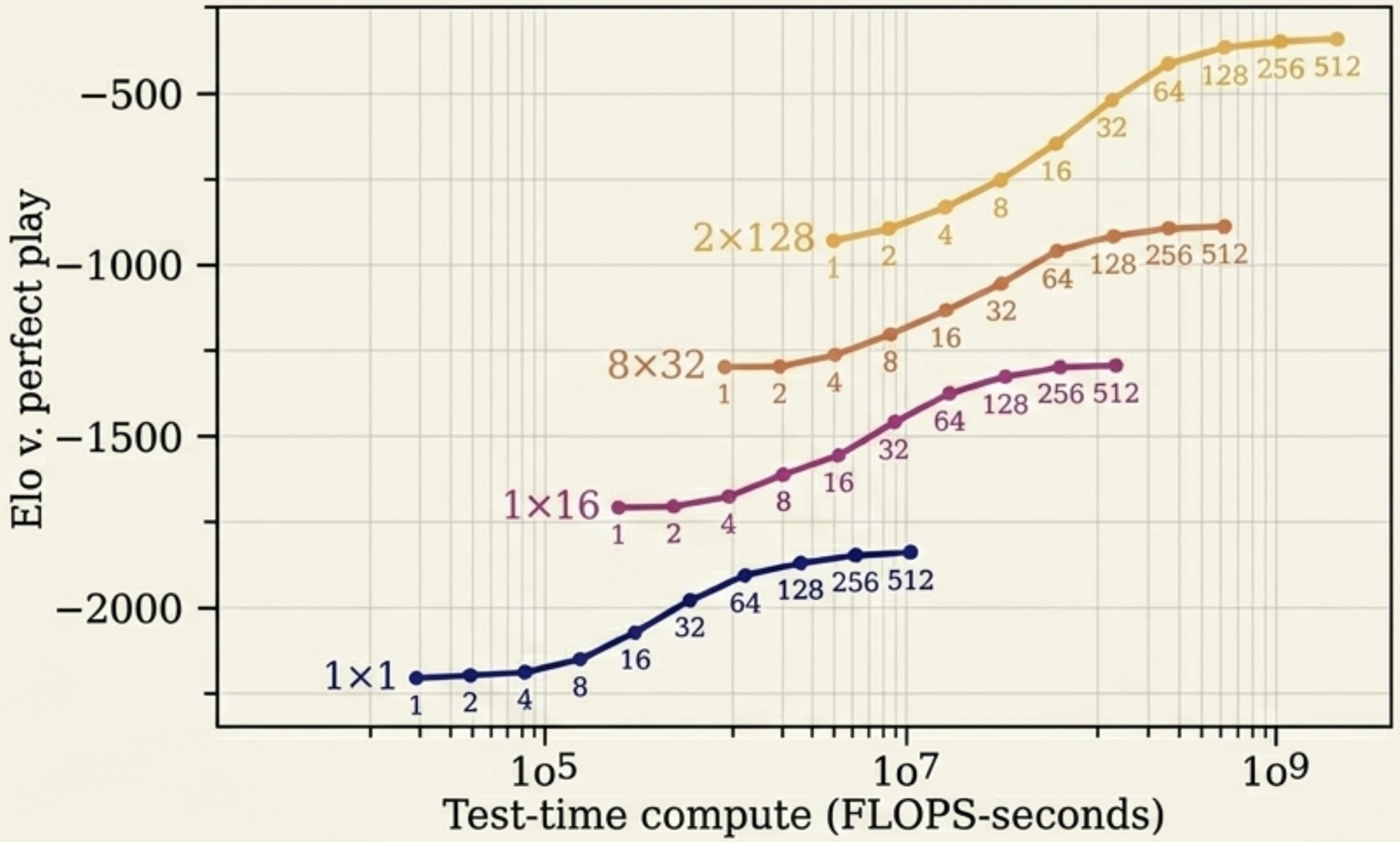
Generation (การสร้างคำตอบ)
หาคำตอบที่ถูกทำได้ยากมาก
มีทางเลือกนับไม่ถ้วนเหมือนงมเข็ม
ในมหาสมุทร (High Entropy)

Verification (การตรวจสอบ)
การประเมินว่าตรรกะนั้นถูกต้องหรือไม่
ทำได้ง่ายกว่าและมีความแน่นอนสูง
(Low Entropy)

นี่คือเหตุผลที่โมเดล System 2 ต้องถูกสอนให้
‘ค้นหา (Search)’ และ ‘ตรวจสอบตัวเอง (Verify)’ สลับกันไป

NotebookLM

Scaling Law ยุคใหม่: พลังของ Test-Time Compute



Key Insight

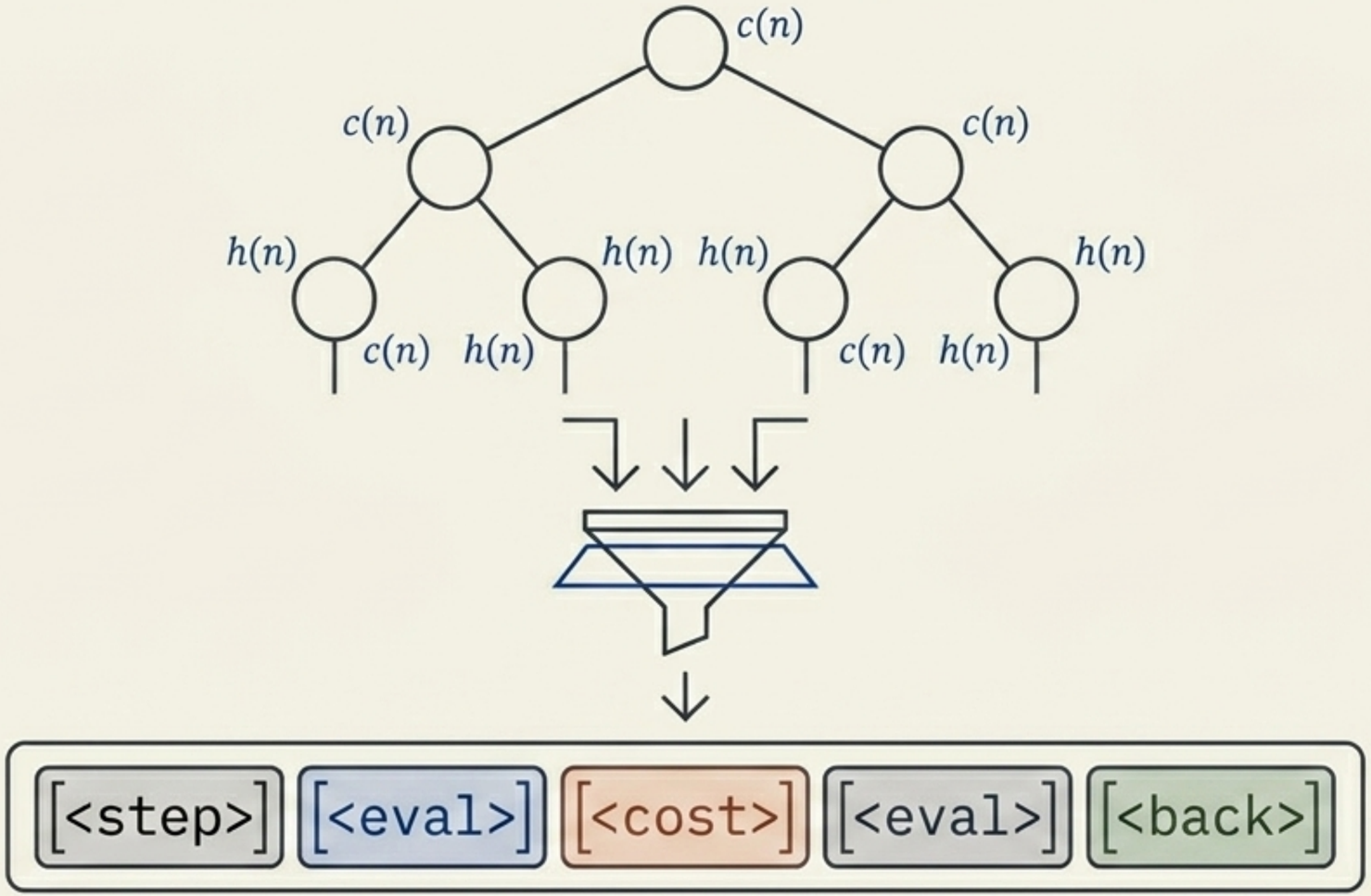
กฎการขยายสเกล (Scaling Law) ไม่ได้หยุดอยู่แค่จำนวนพารามิเตอร์ หรือข้อมูลเทรนอีกต่อไป

The Search Budget

งานวิจัยชี้ให้เห็นว่า หากเราให้เวลา **และโควต้า** Token แก่ AI ในการ**สุ่มและค้นหา**มากขึ้นในขั้นตอนการใช้งานจริง (Test-time) ความแม่นยำจะพุ่งสูงขึ้นอย่างชี้ชัด

NotebookLM

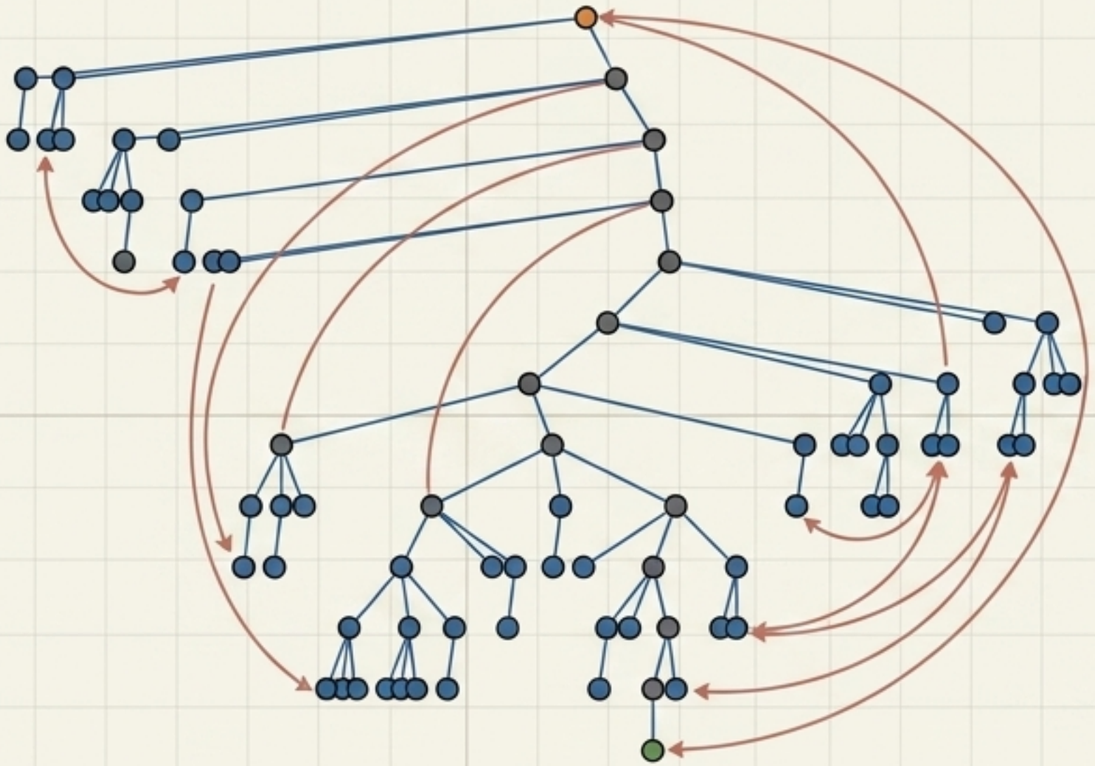
Meta-CoT: จำลอง Tree Search ให้อยู่ในรูปแบบข้อความ



- แทนที่จะใช้ระบบค้นหาภายนอกมาควบคุม AI...
- **Meta-CoT** คือการ**นำกระบวนการค้นหาทั้งหมด** ทั้งการคาดเดา การประเมิน และการย้อนกลับ มาเรียงร้อยเป็นข้อความเส้นตรง (Linearized Search Trace)
- สอนให้ AI อ่านโครงสร้างต้นไม้ให้ออกเป็นตัวอักษร เพื่อให้มันทำทุกกระบวนการจบได้ด้วยตัวเอง (In-context)

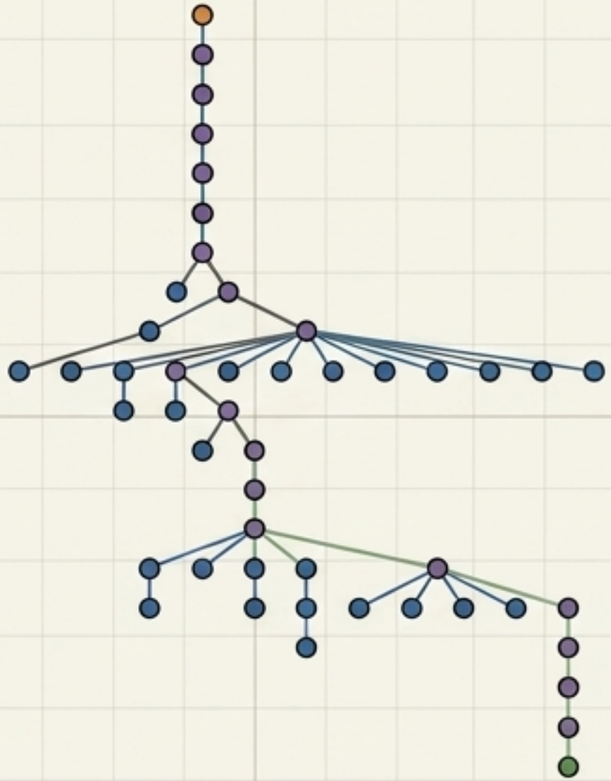
NotebookLM

MCTS vs. A* ในสมองของ AI



Monte Carlo Tree Search (เน้นสำรวจ)

เน้นความสมดุลระหว่างการหาเส้นทางใหม่และการต่อยอดเส้นทางเดิม
โมเดลจะ Backtrack บ่อยครั้งเพื่อกระจายความเสี่ยง

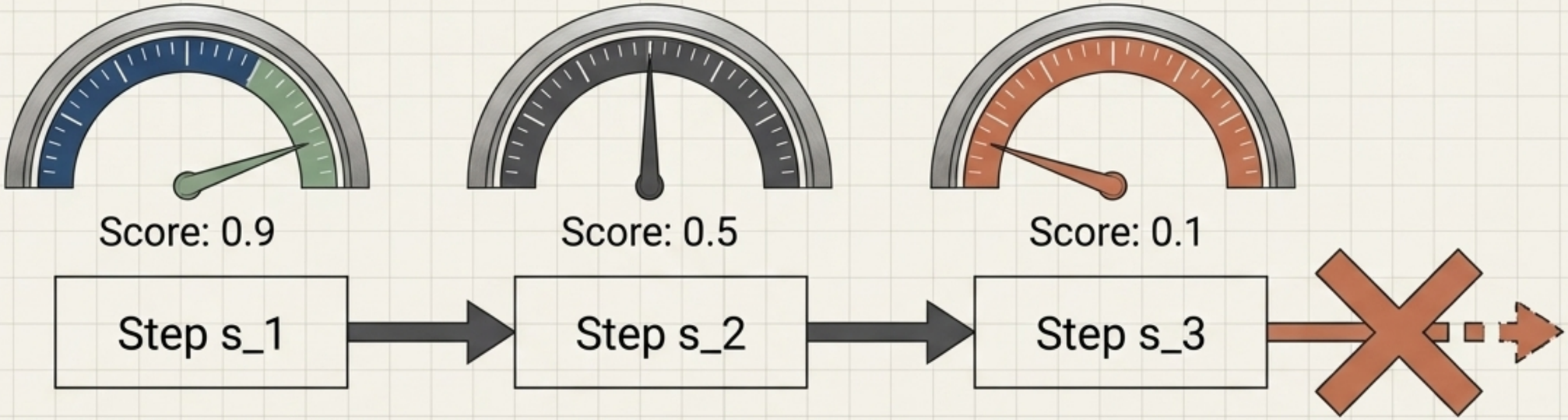


Best-First Search / A* (เน้นพุ่งเป้า)

มุ่งตรงไปยังเส้นทางที่มีคะแนนประเมิน (Heuristic) สูงสุดอย่างรวดเร็ว
โฟกัสเฉพาะจุด ลดการคำนวณซ้ำซ้อน

NotebookLM

Process Reward Models (PRMs): เข็มทิศนำทางความคิด



- AI จะรู้ได้อย่างไรว่าเดินมาถูกทางเมื่อโจทย์ยังไม่จบ?
- PRM คือโมเดลที่ฝึกมาเพื่อประเมินความน่าจะเป็นของแต่ละก้าว (Step-by-step evaluation) หากคะแนนต่ำเกินไป AI จะรู้ตัวและหยุดเส้นทางนั้นทันที ช่วยลดการสูญเสีย Token

NotebookLM

ศิลปะแห่งการถอยหลังตั้งหลัก (Learning to Backtrack)

...so the roots must be x=1 and x=-1.

[EVALUATE] The heuristic score is low.

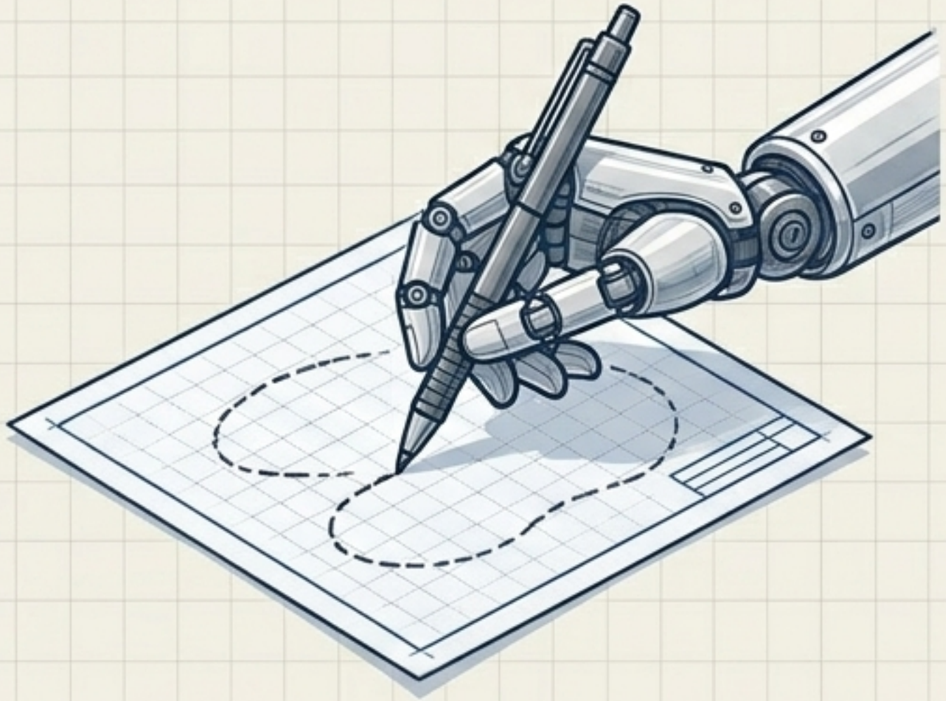
[BACK] Wait, this contradicts the earlier assumption that x must be even. Let's try factoring by 2 instead...

จุดเด่นที่สุดของ Meta-CoT คือการสอนให้ AI รู้จักคำว่า เดี๋ยวก่อน...

เมื่อประเมินด้วย PRM แล้วพบทางตัน โมเดลจะสร้าง Token พิเศษเพื่อย้อนกลับ (Semantic Backtracking) ไปยังสถานะที่ปลอดภัยก่อนหน้า ซึ่งพบได้ชัดเจนใน DeepSeek-R1 และ o1

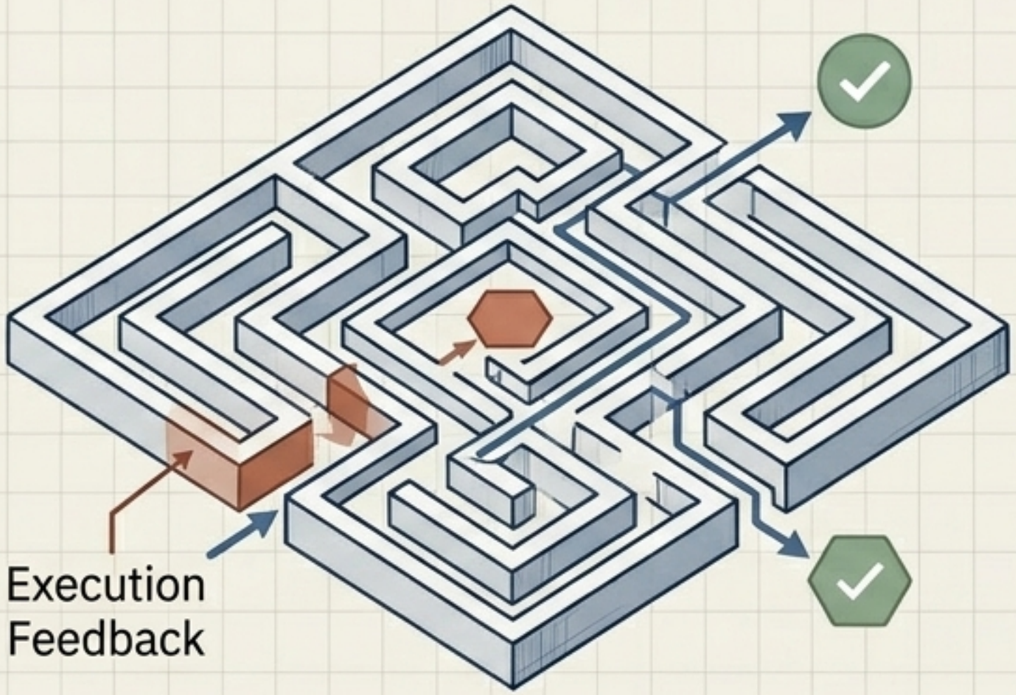
NotebookLM

Meta-RL: เมื่อการเลียนแบบไม่พอ ต้องสอนให้คิดเอง



SFT (แค่ท่องจำรูปแบบ)

ทำให้ AI เลียนแบบสไตล์การเขียน CoT ได้
แต่ไม่ได้เข้าใจหลักการแก้ปัญหาลึกซึ้ง



Meta-RL (ค้นพบอัลกอริทึม)

ปล่อยให้ AI ทดลองแก้โจทย์ผ่านสภาพแวดล้อมจำลอง
เพื่อให้มันปรับแต่งอัลกอริทึมการคิด
(Policy Optimization) ของตัวเองขึ้นมาใหม่

NotebookLM

หลักฐานเชิงประจักษ์: ส่องสมองของโมเดลปัจจุบัน

OpenAI o1

[THINKING]...So the total number of distinct permutations must be 5! which is 120. But wait, I need to account for the identical elements. Let's recalculate. The formula is n!/p!q!r!... So the sum is 5! / 2!1!1!1! which is 120/2 = 60. Let's double check the reasoning. So the sum is...

Repetition

DeepSeek-R1

[THOUGHTS]...Therefore, the area of the shaded region is the integral from 0 to 2. wait, actually... this contradicts the earlier assumption that the boundary is a straight line. Let me re-read the problem statement... Okay, it's a curve. Let's try factoring by...

Backtracking

Gemini Thinking

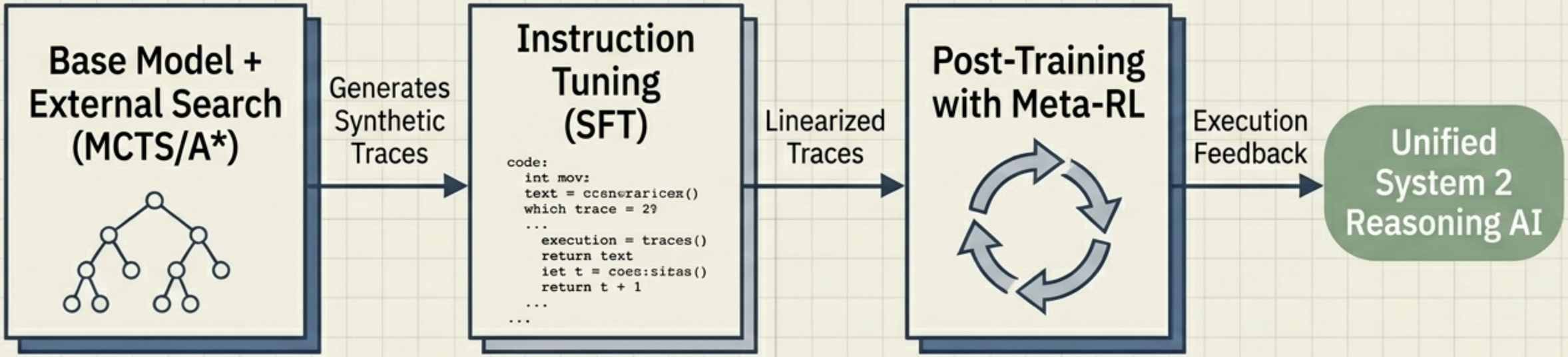
[THINKING STEP]...Following the logic, the solution to the differential equation is y = Ce^x. Now, let's apply the boundary conditions y(0) = 1. This gives C = 1. Let's double check the reasoning. Is this the only solution? Let me verify with the original equation...

Verification

ร่องรอยกระบวนการคิดของโมเดลระดับท็อป แสดงพฤติกรรมที่ตรงกับทฤษฎี Meta-CoT อย่างชัดเจน: มีการตรวจสอบตัวเองตัวเอง (Self-criticism), มีการวกกลับเมื่อเจอทางตัน, และมีการพิมพ์ทวนซ้ำเพื่อขยาย Context

NotebookLM

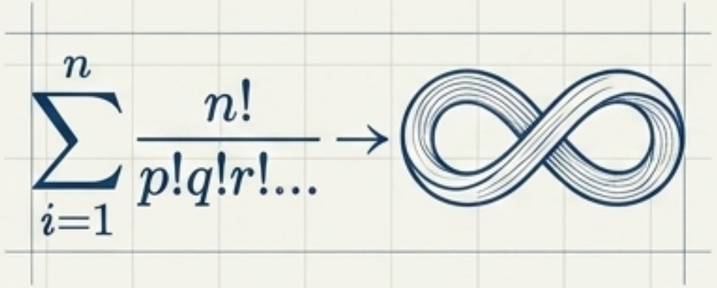
บทสรุปสถาปัตยกรรม System 2 Pipeline



วิธีสร้าง AI ที่คิดลึกซึ้งภายในโมเดลเดียว: ใช้เครื่องมือ Search ภายนอกเพื่อสร้างร่องรอยการคิดจำลอง -> นำร่องรอยนั้นไปสอน AI ผ่าน SFT ให้คุ้นเคย -> ขัดเกลาตรรกะขั้นสุดท้ายให้แหลมคมด้วย Reinforcement Learning

NotebookLM

The Open Frontier: ความท้าทายที่รอการค้นพบ

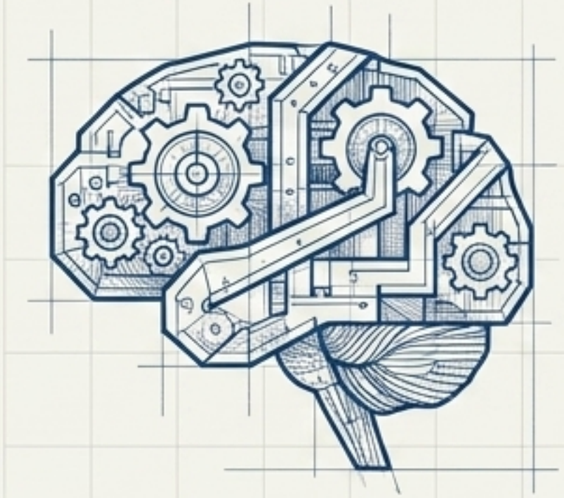


Open-ended Verification

PRM จะทำงานอย่างไรกับงานปลายเปิด อย่างการเขียนเชิงสร้างสรรค์ที่ไม่มีคำตอบตายตัวเหมือนคณิตศาสตร์?

The Verifier Gap

ความเสี่ยงที่ AI จะเรียนรู้วิธีหาช่องโหว่และหลอกโมเดลตรวจสอบ (Reward Hacking)

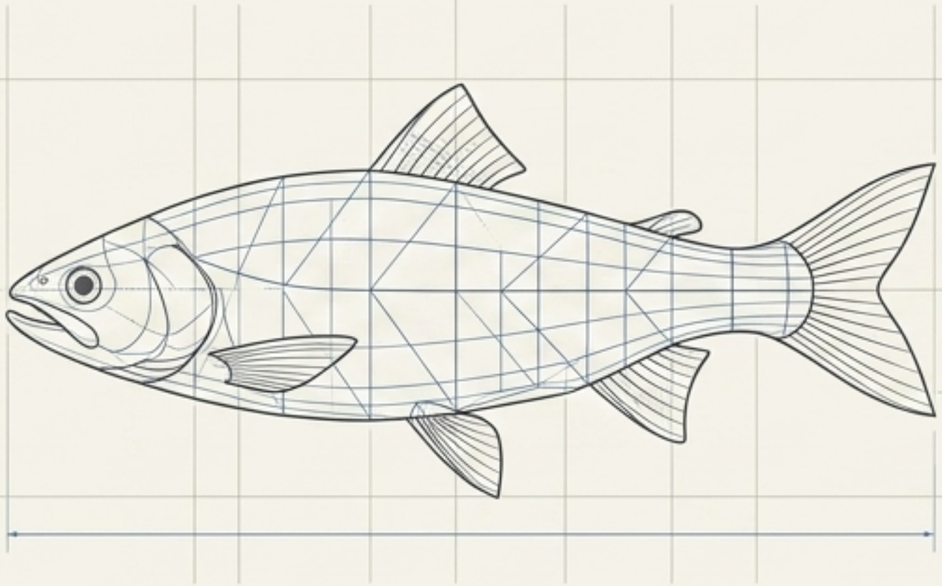


Super-Intelligence

ด้วยพลังของ Meta-RL AI จะค้นพบวิธีคิดคำนวณแบบใหม่ที่มนุษย์ไม่เคยนึกถึงมาก่อนได้หรือไม่?

NotebookLM

บทสรุป



Standard CoT

Give an AI a CoT, and it solves one problem.

Meta-CoT

Teach an AI to search, and it learns how to think for a lifetime.

ยุคต่อไปของ AI ไม่ใช่แค่การป้อนข้อมูลมหาศาล แต่คือการสร้าง ‘สถาปัตยกรรมแห่งความคิด’ ที่มอบเวลาและพื้นที่ให้ AI ได้สืบค้น ประเมิน และทบทวนตัวเองอย่างลึกซึ้งที่สุด

NotebookLM