

OpenAI o1 / o3

Jan 12 2025
Roger Lam

Chain-of-Thought improved performance

One method ppl used to get better performance from LLMs is to literally tell it to “think step by step”.

The model would output like someone explaining their way of thinking **but it didn't actually spend more time thinking.**

(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. ✓*

<https://www.promptingguide.ai/techniques/cot>

o-series models “think” before they respond

o-series on the other hand
**spend more time planning its
approach to get better
performance.**

They call this “test-time
compute”.

This is in addition to more
reinforcement learning
(train-time compute).

The image shows a side-by-side comparison of two chat interfaces: GPT-4o on the left and OpenAI o1-preview on the right. Both interfaces show a user input and a model response. The user input is a ciphertext: "oyfjdnsdr rtqwainr acxz mynzbhbx -> Think step by step". The model response is a decoded message: "oyekaijzdf aaptcg suaokybhai ouow aqht mynznvaatzacdfoulxxz". The GPT-4o response is a single block of text. The OpenAI o1-preview response is a single block of text, but it is highlighted with a yellow box. The highlighted text is: "Thought for 5 seconds v To decode the given ciphertext using the example provided, we'll use the pattern identified in the example." Below this, the text "Ciphertext Example:" is visible. The GPT-4o response also includes a section titled "1. Example given:" with input and output examples, and a section titled "By examining the words:".

GPT-4o

User

oyfjdnsdr rtqwainr acxz mynzbhbx -> Think step by step

Use the example above to decode:

oyekaijzdf aaptcg suaokybhai ouow aqht mynznvaatzacdfoulxxz

Let's break this down step by step based on the example:

1. Example given:

- Input: oyfjdnsdr rtqwainr acxz mynzbhbx
- Output: **Think step by step**

By examining the words:

OpenAI o1-preview

User

oyfjdnsdr rtqwainr acxz mynzbhbx -> Think step by step

Use the example above to decode:

oyekaijzdf aaptcg suaokybhai ouow aqht mynznvaatzacdfoulxxz

Thought for 5 seconds v

To decode the given ciphertext using the example provided, we'll use the pattern identified in the example.

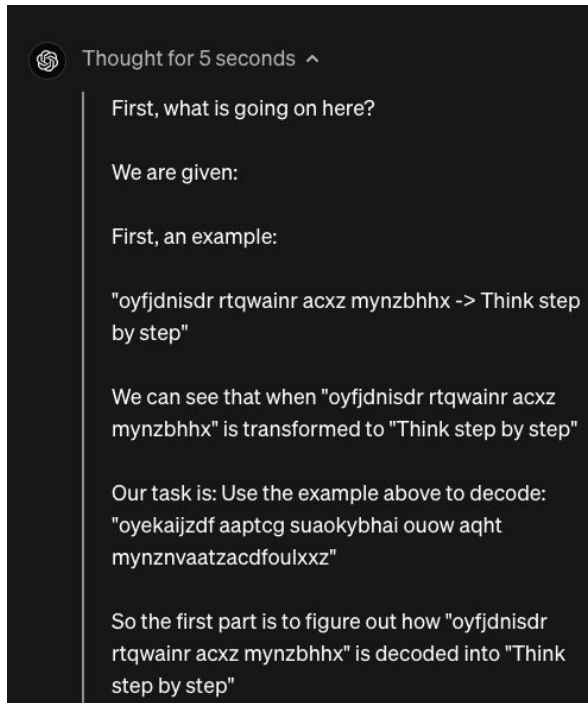
Ciphertext Example:

Our large-scale reinforcement learning algorithm teaches the model how to think productively using its chain of thought in a highly data-efficient training process. We have found that the performance of o1 consistently improves with more reinforcement learning (train-time compute) and with more time spent thinking (test-time compute). The constraints on scaling this approach differ substantially from those of LLM pretraining, and we are continuing to investigate them.

What does the thinking look like?

We're not sure exactly. OpenAI provides a summarized view of what happened but not the exact thought process.

(Probably because they don't want people to steal the output and distill their own models.)



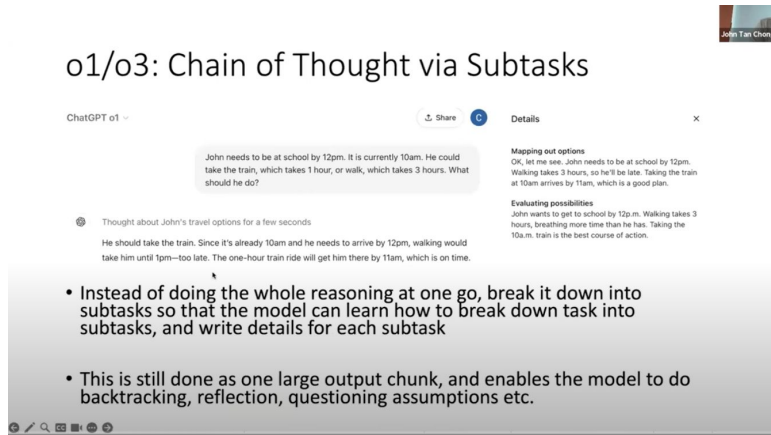
What do we think it looks like?

1. Trained on Chain-of-Thought reasoning
2. Break down requests into sub-tasks
3. Perform advanced (not-so) secret sauce methods on the subtasks

From John Tan Chong Min's speculation on o-models

<https://www.youtube.com/watch?v=f5obaHiOog4>

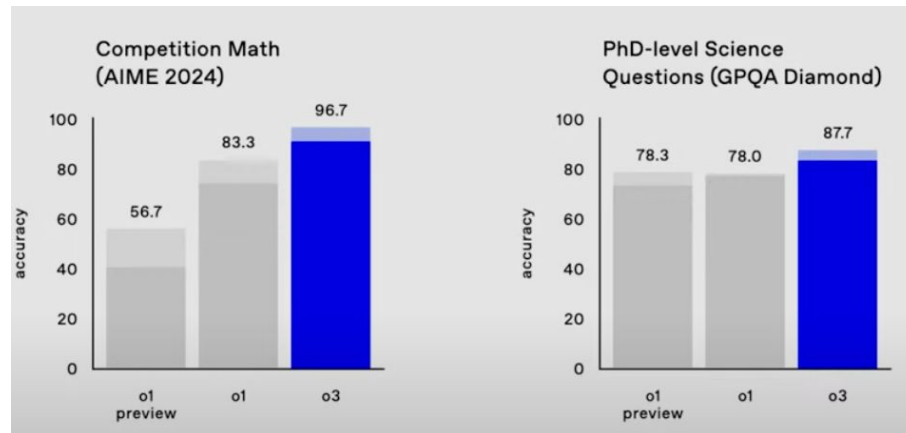
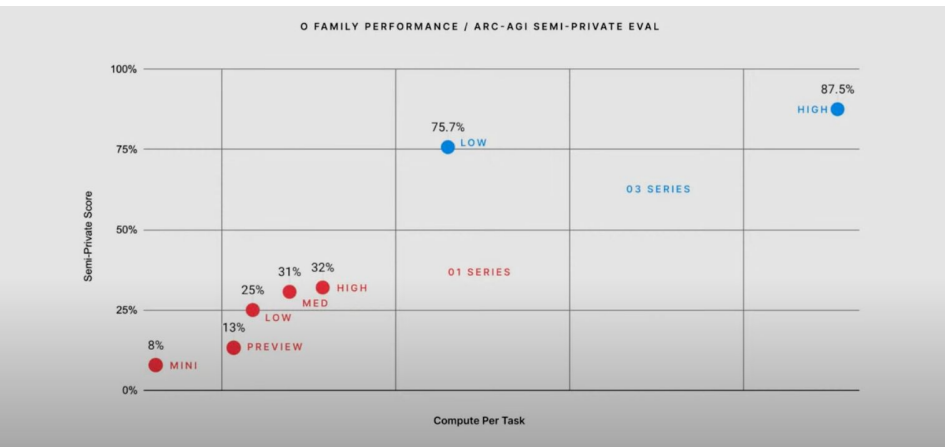
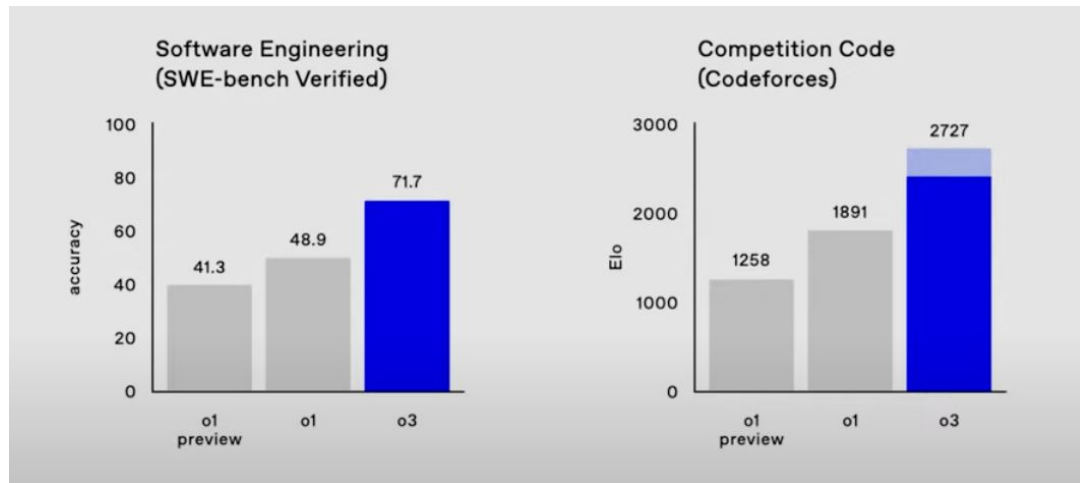
o1/o3: Chain of Thought via Subtasks



How good is o3?

A lot better on coding, math and science benchmarks.

Also on ARC-AGI benchmark.



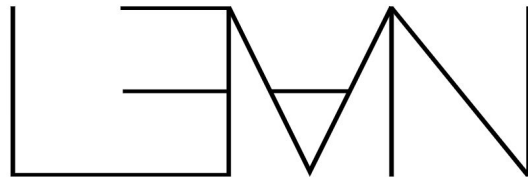
Math research is promising

OpenAI and AI systems in general has captured the attention of Terence Tao

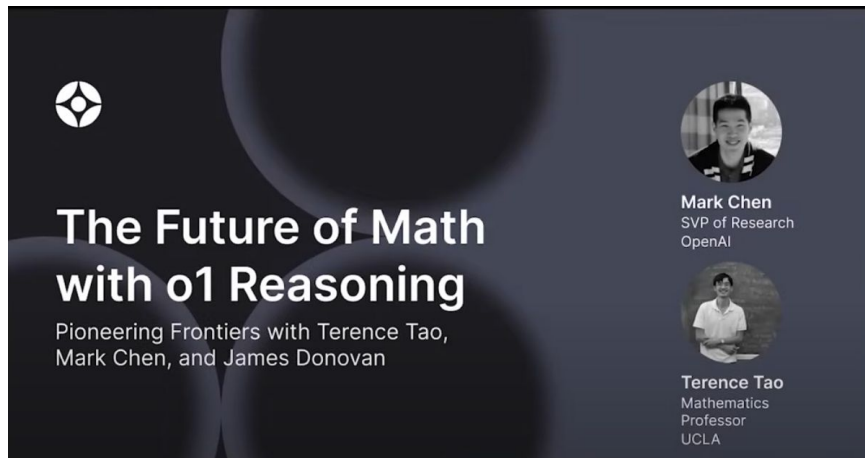
o-models can keep running, thinking and learn with mathematical solvers to (hopefully) solve conjectures.

Math has a ground truth in a way that more open-ended problems don't.

https://www.youtube.com/watch?v=kO192K7_FaQ



Programming Language and Theorem Prover



You should use o-models differently from chat models

Shoutout to Latent Space blog (not just a pod!) for publishing this guide on how to best use o-models.

1. Don't write prompts; write briefs
2. Focus on Goals
3. Know what it does well and what it doesn't

<https://www.latent.space/p/o1-skill-issue>

The Anatomy of an o1 Prompt

The diagram illustrates the structure of an o1 prompt, divided into four main sections on the right side, each with a corresponding colored vertical bar: Goal (green), Return Format (blue), Warnings (orange), and Context Dump (grey). The prompt text is contained within a dark grey box on the left.

Goal

I want a list of the best medium-length hikes within two hours of San Francisco.

Return Format

Each hike should provide a cool and unique adventure, and be lesser known.

For each hike, return the name of the hike as I'd find it on AllTrails, then provide the starting address of the hike, the ending address of the hike, distance, drive time, hike duration, and what makes it a cool and unique adventure.

Warnings

Return the top 3.

Be careful to make sure that the name of trail is correct, that it actually exists, and that the time is correct.

--

Context Dump

For context: my girlfriend and i hike a ton! we've done pretty much all of the local SF hikes, whether that's presidio or golden gate park. we definitely want to get out of town -- we did mount tam pretty recently, the whole thing from the beginning of the stairs to stinson - it was really long and we are definitely in the mood for something different this weekend! ocean views would still be nice. we love delicious food. one thing i loved about the mt tam hike is that it ends with a celebration (Arriving in town to breakfast!) The old missile silos and stuff near Discovery point is cool but I've just done that hike probably 20x at this point. We won't be seeing each other for a few weeks (she has to stay in LA for work) so the uniqueness here really counts.

6x more expensive but other reasoning models are here

o1 is 6x more expensive in both input and output tokens and I imagine o1 uses a lot more output tokens too.

[QwQ](#) is also an open reasoning model from Alibaba and [DeepSeek-R1](#) is offering their own reasoning model.

And as trends go for LLMs, prices will prob go down.

Model	Pricing
gpt-4o	\$2.50 / 1M input tokens
	\$1.25 / 1M cached** input tokens
	\$10.00 / 1M output tokens

Model	Pricing
o1	\$15.00 / 1M input tokens
	\$7.50 / 1M cached* input tokens
	\$60.00 / 1M output** tokens

Hope this helped!

Follow for more AI / ML / Software Engineering dives.

[linkedin.com/in/lam-roger](https://www.linkedin.com/in/lam-roger)

lamroger.com