Future Directions

CSC2541H1 Topics in Machine Learning, Winter 2025, UToronto

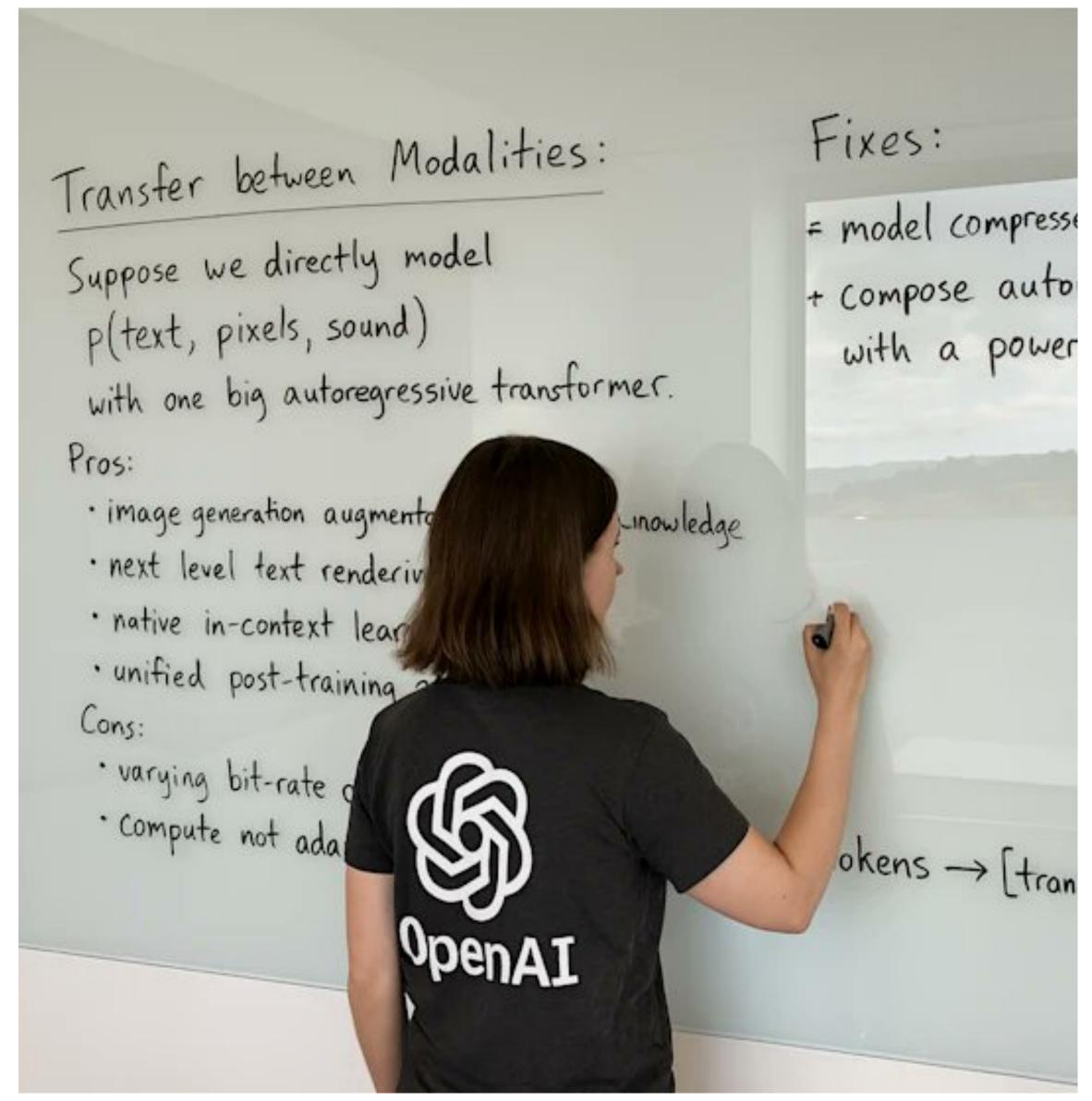
Announcements

Course Evaluations

What are the next frontiers?

Why not model bytes? Back to basics

- All data is stored as bytes.
- Why don't we just model sequences of bytes?
- Multi-modality in one fell swoop.
- Varying bit-rate concerns notwithstanding, the T^2 scaling of Transformers is a key concern.



Synthetic image from GPT 40 native image generation release.

Why not model bytes? Back to basics

- Tokenization is our current solution to long context issues, but it's hacky and brittle.
- State space models (better scaling) are one attempt at a solution.
- Learned 'tokenization' is another.
- We'll hear about these today.



What's in the base model?

The "aha" moment

- DeepSeek-R1-Zero.
 - Finetune a base model with RL against an accuracy reward + format reward.
- The model learns to reason from scratch - no SFT.
- Common theme in the course: what's in the base model?

Question: If a > 1, then the sum of the real solutions of $\sqrt{a - \sqrt{a + x}} = x$ is equal to

Response: <think>

To solve the equation $\sqrt{a - \sqrt{a + x}} = x$, let's start by squaring both \cdots

$$\left(\sqrt{a-\sqrt{a+x}}\right)^2 = x^2 \implies a-\sqrt{a+x} = x^2.$$

Rearrange to isolate the inner square root term:

$$(a-x^2)^2 = a+x \implies a^2-2ax^2+(x^2)^2 = a+x \implies x^4-2ax^2-x+(a^2-a)=0$$

. . .

Wait, wait. Wait. That's an aha moment I can flag here.

Let's reevaluate this step-by-step to identify if the correct sum can be · · ·

We started with the equation:

$$\sqrt{a-\sqrt{a+x}}=x$$

First, let's square both sides:

$$a - \sqrt{a + x} = x^2 \implies \sqrt{a + x} = a - x^2$$

Next, I could square both sides again, treating the equation: · · ·

..

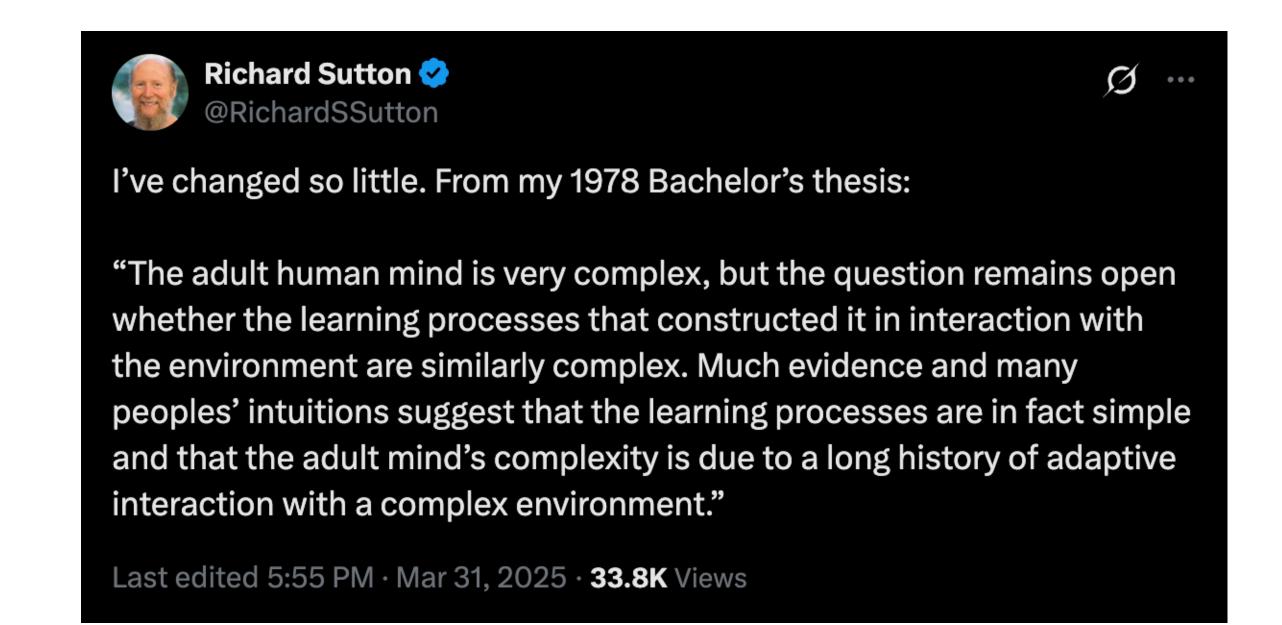
Table 3 | An interesting "aha moment" of an intermediate version of DeepSeek-R1-Zero. The model learns to rethink using an anthropomorphic tone. This is also an aha moment for us, allowing us to witness the power and beauty of reinforcement learning.

DeepSeek-Al. 2025. DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning

What's in the base model?

The "aha" moment

- Common theme in the course: what's in the base model?
- Speculation: this tells us something about ourselves, about our language, more than it tells us something about the methods.
 - We "store" algorithms in language.
 - Simple learning methods can recover them.



What I hope you got out of the course

- An overview of a model pipeline
- A sense that you can still contribute to better model pipelines or our understanding, there's still room for improvement
- Some research and presentation skills
- Some lessons
 - "classify methods by their source of supervision" -> labels? verifier? compiler?
 - "scale is a powerful abstraction" -> some abstractions simplify complex systems

Thank you

for everything

- The amazing presentations and code notebooks
 - You all deserve lots of credit for the course content—I learned a lot
- Your patience with my gibberish
 - It was an important term in my life, returning to work
- Your engagement