ReFoRM Reading Group

Rethinking Foundations for Real-world ML

Amin Saberi & Andrew Ilyas

Welcome to ReFoRM!

Not a slight to other ML!
Just starts with R 😅

What this is: an experimental reading group on foundations of "real-world" ML

What does this mean?

Idealized picture of ML: something like $\theta^* = \arg\min_{\theta \in \Theta} \mathbb{E}_{z_i \sim D}[\ell(z_i; \theta)]$

ML powering systems like Claude, DALL-E, Google Photos:

Welcome to ReFoRM!

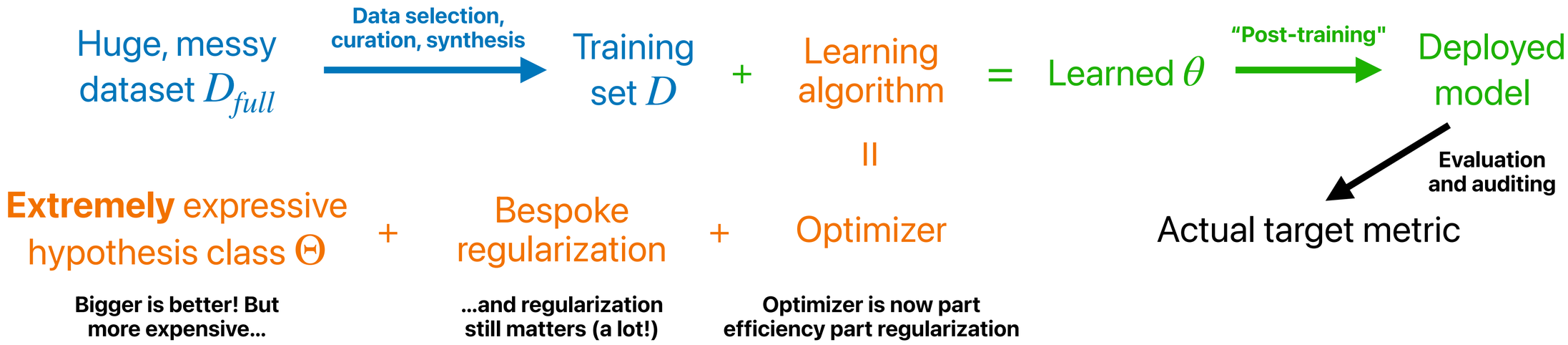
Not a slight to other ML!
Just starts with R 😅

What this is: an experimental reading group on foundations of "real-world" ML

What does this mean?

Idealized picture of ML: something like $\theta^* = \arg\min_{\theta \in \Theta} \mathbb{E}_{z_i \sim D}[\ell(z_i; \theta)]$

ML powering systems like Claude, DALL-E, Google Photos:



Goal of this group

Goal of this group

What do rigorous foundations for this new age of ML look like?

Goal of this group

What do rigorous foundations for this new age of ML look like?

How can tools from statistics, CS theory, and operations inform a better understanding of machine learning algorithms and systems?

Goal of this group

What do rigorous foundations for this new age of ML look like?

How can tools from statistics, CS theory, and operations inform a better understanding of machine learning algorithms and systems?

What are the right questions to ask, and phenomena to explain—at what level of abstraction should we be aiming to explain them?

Goal of this group

What do rigorous foundations for this new age of ML look like?

How can tools from statistics, CS theory, and operations inform a better understanding of machine learning algorithms and systems?

What are the right questions to ask, and phenomena to explain—at what level of abstraction should we be aiming to explain them?

What theoretical models not only explain unexpected phenomena, but also predict new phenomena that we can verify experimentally?

Today’s meeting

Logistics/plan for the quarter

Brief intro to this quarter’s topic: reasoning

Topic for this quarter: LLM Reasoning

Topic for this quarter: LLM Reasoning

Topics by weighted combination of {interest, coverage}:

Data selection, curation, and synthesis

Scaling laws & prediction

Post-training

Fine-tuning

LLM “Reasoning”

All past presentations are online: https://andrewilyas.github.io/REFORM-reading-group/

Topic for this quarter: LLM Reasoning

Topics by weighted combination of {interest, coverage}:

Fall
- Data selection, curation, and synthesis
- Scaling laws & prediction

Post-training

Fine-tuning

LLM “Reasoning”

All past presentations are online: https://andrewilyas.github.io/REFORM-reading-group/

Topic for this quarter: LLM Reasoning

Topics by weighted combination of {interest, coverage}:

Fall
- Data selection, curation, and synthesis
- Scaling laws & prediction

Winter
- Post-training
- Fine-tuning

LLM “Reasoning”

All past presentations are online: https://andrewilyas.github.io/REFORM-reading-group/

Intended format (please sign up!)

Intended format (please sign up!)

Goal: Build intuition, leverage diversity in this group, start collaborations (bringing new perspectives from everyone’s field)

Intended format (please sign up!)

Goal: Build intuition, leverage diversity in this group, start collaborations (bringing new perspectives from everyone's field)

Sign up to be a discussant at

Intended format (please sign up!)

Goal: Build intuition, leverage diversity in this group, start collaborations (bringing new perspectives from everyone's field)

Sign up to be a discussant at

Goal(s) of the discussant:

Intended format (please sign up!)

Goal: Build intuition, leverage diversity in this group, start collaborations (bringing new perspectives from everyone's field)

Sign up to be a discussant at

Goal(s) of the discussant:

1. A single "deep dive" per week about one subject (can be multiple papers) by 1-2 discussants

Intended format (please sign up!)

Goal: Build intuition, leverage diversity in this group, start collaborations (bringing new perspectives from everyone's field)

Sign up to be a discussant at

Goal(s) of the discussant:

1. A single "deep dive" per week about one subject (can be multiple papers) by 1-2 discussants
2. We have suggested several papers for each week, more than one can cover thoroughly in a week. **Pick a small, focused set of papers and read them thoroughly**

Intended format (please sign up!)

Goal: Build intuition, leverage diversity in this group, start collaborations (bringing new perspectives from everyone's field)

Sign up to be a discussant at

Goal(s) of the discussant:

1. A single "deep dive" per week about one subject (can be multiple papers) by 1-2 discussants
2. We have suggested several papers for each week, more than one can cover thoroughly in a week. **Pick a small, focused set of papers and read them thoroughly**
3. Prepare a 20-30 minute presentation, accessible to a second year PhD student, focusing on (a) seeding discussion and (b) identifying gaps and connections, and (c) formulating open problems

Intended format (please sign up!)

Goal: Build intuition, leverage diversity in this group, start collaborations (bringing new perspectives from everyone's field)

Sign up to be a discussant at

Goal(s) of the discussant:

1. A single "deep dive" per week about one subject (can be multiple papers) by 1-2 discussants
2. We have suggested several papers for each week, more than one can cover thoroughly in a week. **Pick a small, focused set of papers and read them thoroughly**
3. Prepare a 20-30 minute presentation, accessible to a second year PhD student, focusing on (a) seeding discussion and (b) identifying gaps and connections, and (c) formulating open problems

Everyone else: Read the paper/watch a podcast/something! Try to come with some familiarity

Intended format (please sign up!)

Intended format (please sign up!)

Goal: Build intuition, leverage diversity in this group, start collaborations (bringing new perspectives from everyone’s field)

Intended format (please sign up!)

Goal: Build intuition, leverage diversity in this group, start collaborations (bringing new perspectives from everyone's field)

Sign up to be a discussant at https://tinyurl.com/reform-S25

Intended format (please sign up!)

Goal: Build intuition, leverage diversity in this group, start collaborations (bringing new perspectives from everyone's field)

Sign up to be a discussant at https://tinyurl.com/reform-S25

New this quarter: Brainstorming sessions

Intended format (please sign up!)

Goal: Build intuition, leverage diversity in this group, start collaborations (bringing new perspectives from everyone's field)

Sign up to be a discussant at https://tinyurl.com/reform-S25

New this quarter: Brainstorming sessions

1. People expressed interest in working on/thinking about open problems together

Intended format (please sign up!)

Goal: Build intuition, leverage diversity in this group, start collaborations (bringing new perspectives from everyone's field)

Sign up to be a discussant at https://tinyurl.com/reform-S25

New this quarter: Brainstorming sessions

1. People expressed interest in working on/thinking about open problems together
2. Our attempt at this: brainstorming docs/sessions

Intended format (please sign up!)

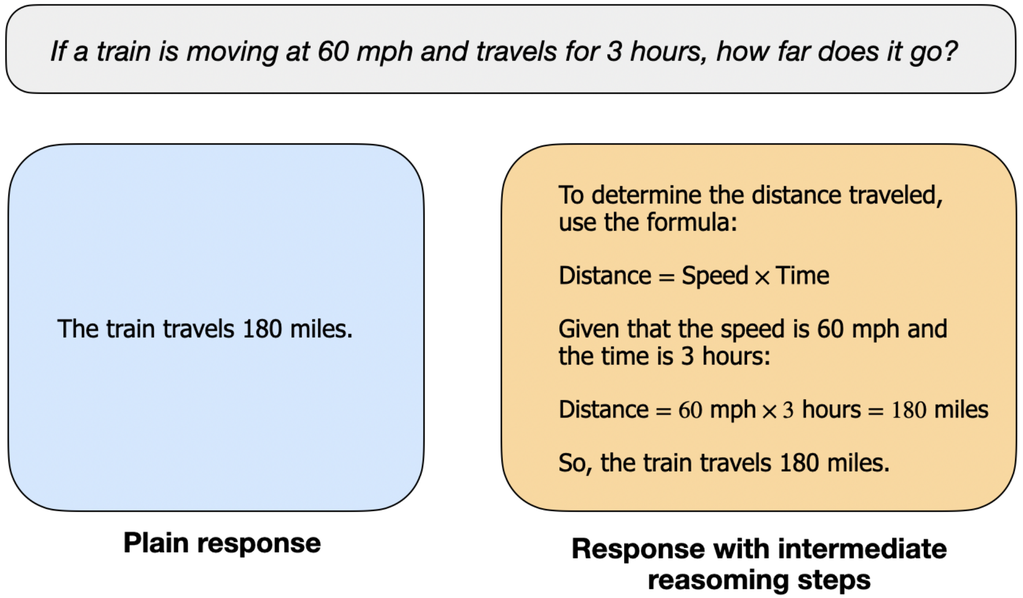
Goal: Build intuition, leverage diversity in this group, start collaborations (bringing new perspectives from everyone's field)

Sign up to be a discussant at https://tinyurl.com/reform-S25

New this quarter: Brainstorming sessions

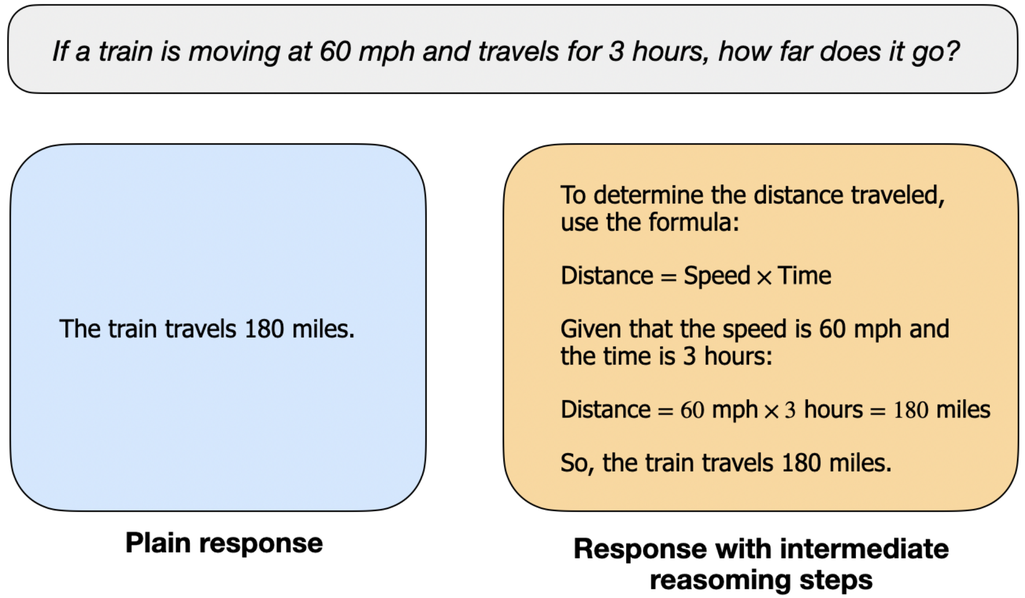
1. People expressed interest in working on/thinking about open problems together
2. Our attempt at this: brainstorming docs/sessions
3. Main idea:
    1. We'll keep a **running document** throughout the quarter where people can dump open problems that they'd be open to thinking about, and other people can express interest
    2. Some meetings will be open problem sessions where people can present on problems they like from the doc

Introduction to reasoning



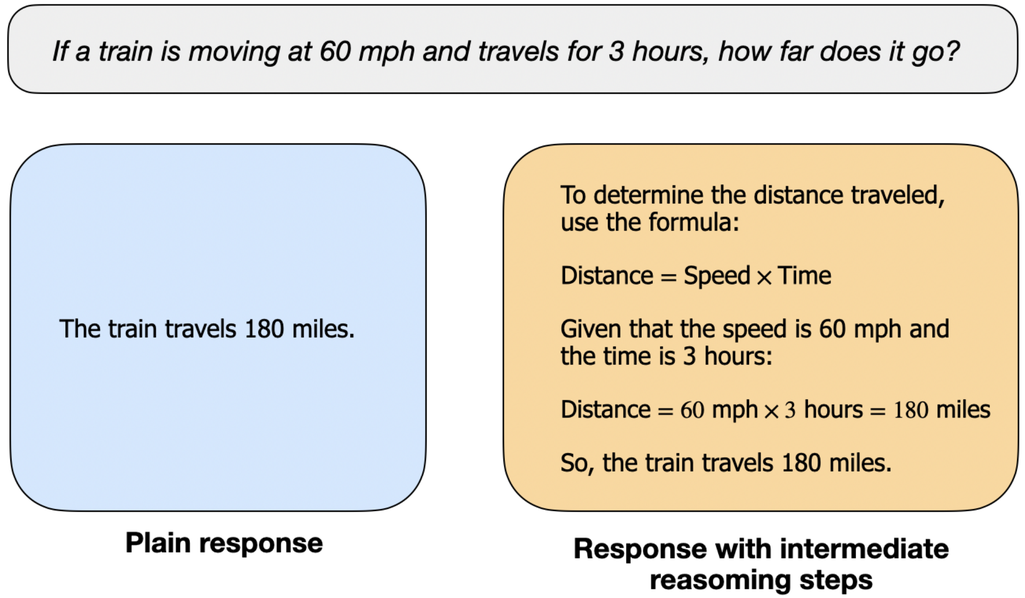
Introduction to reasoning

General goal: Get a language model to solve multi-step questions that require putting together multiple steps



Introduction to reasoning

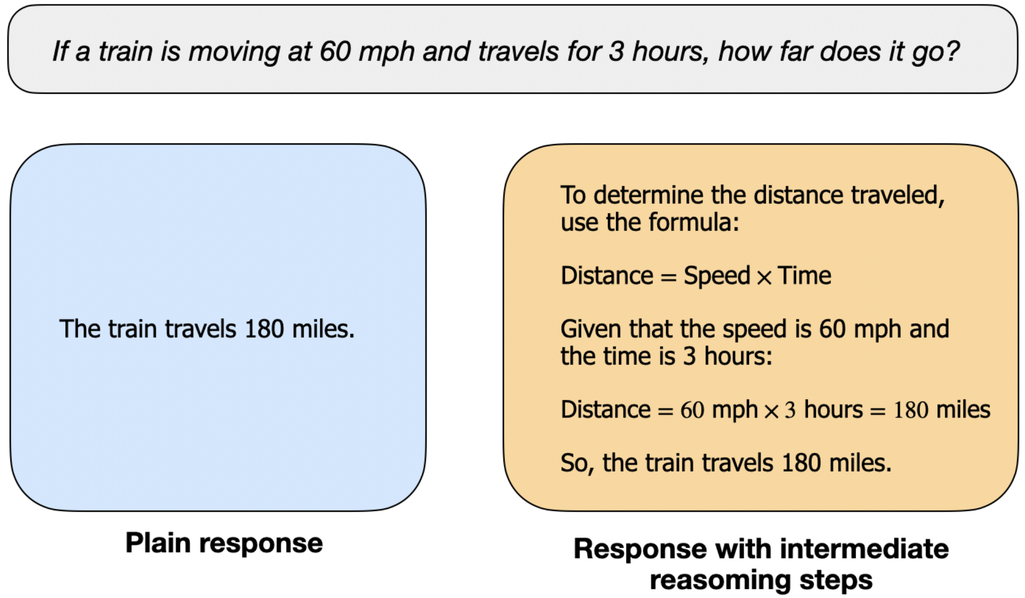
General goal: Get a language model to solve multi-step questions that require putting together multiple steps



Introduction to reasoning

General goal: Get a language model to solve multi-step questions that require putting together multiple steps

Aside: There's lots of philosophical debate about what it means to "reason" (especially how it relates to human reasoning), but we'll mostly ignore this



Evaluating reasoning

Mathematical reasoning: Benchmarks (GSM8k, MATH) use math word problems

Commonsense and logical reasoning: Benchmarks (Commonsense-QA, Strategy-QA, BigBench-Hard) evaluate everyday, multi-hop reasoning, and complex puzzles

Other reasoning tasks: Symbolic/algorithmic (e.g., list sorting, rubik's cube) and Code-based (e.g., programming problems)

Goal: High accuracy on benchmarks + faithful "reasoning traces"

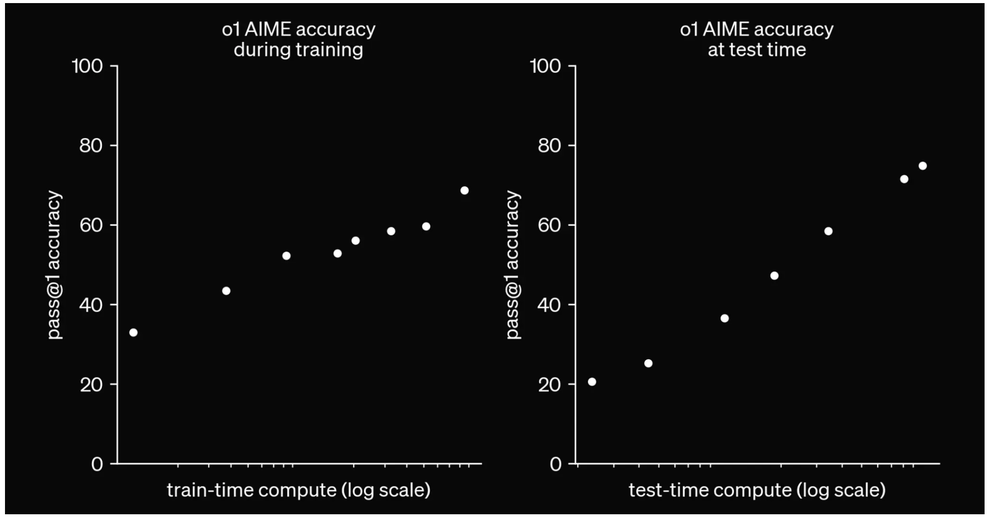
Basic idea: Generate a lot of stuff

Unifying theme of a lot of reasoning stuff is that the model “thinks” for a while before outputting an answer (i.e., test-time computation)

As a result, key questions are usually:

How do we encourage models to generate a lot of stuff?

How do we make sure that “generating a lot of stuff” actually helps reach the right answer?



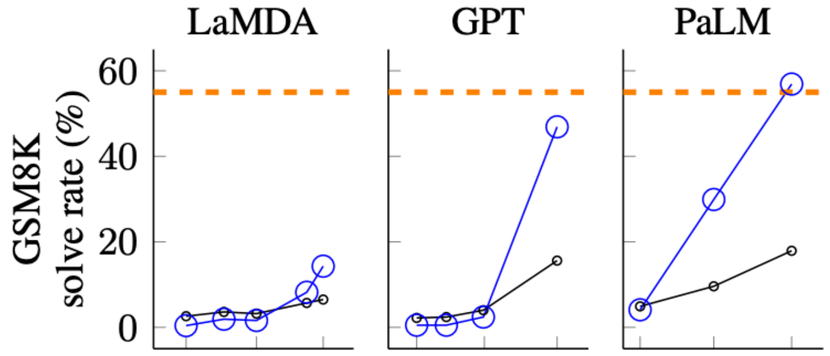
Simplest/Earliest forms of reasoning

Scratchpad/Chain of thought (Nye et al., 2021; Wei et al., 2022; Kojima et al. 2022)

Change nothing about the way the model is trained

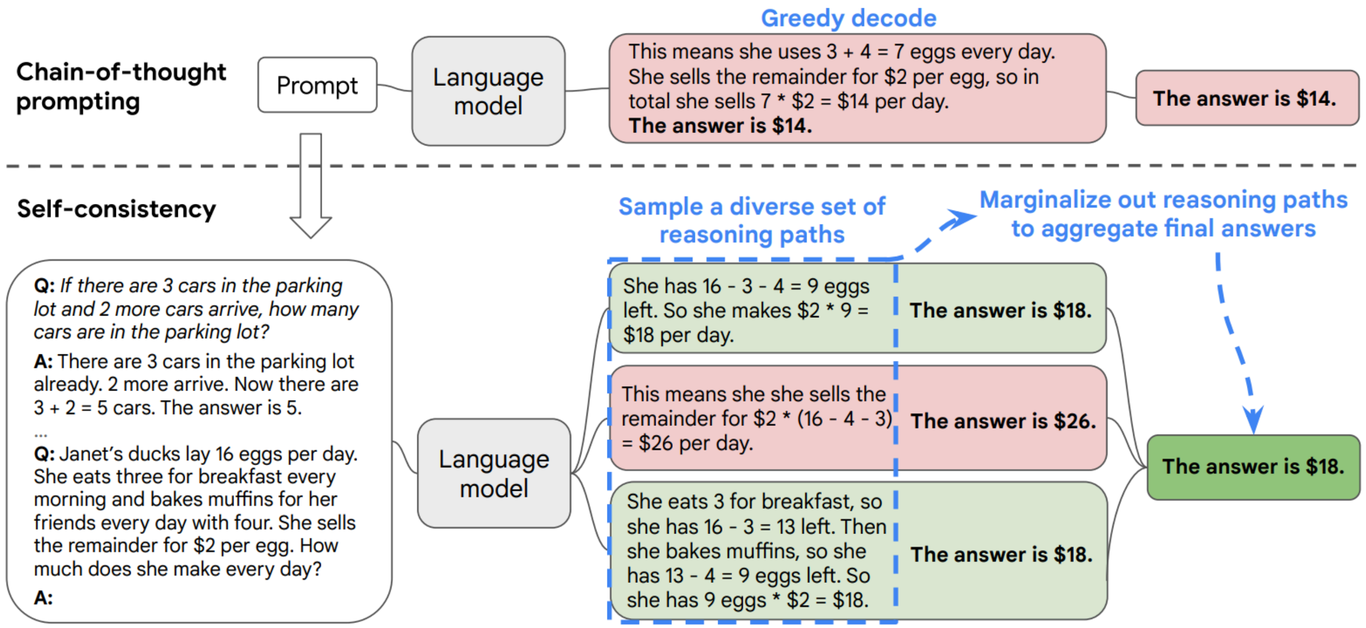
At test time, just ask the model to “show its work,” or “think step by step, or give it “scratch space” to perform big computations

Improved benchmark performance a lot but only worked for really big models



Refinement: Self-consistency

Idea: instead of asking the model to generate a lot of text once, ask it to generate a lot of text many times and take a majority vote (Wang et al., 2023)



PaLM-540B	Zero-shot CoT (Kojima et al., 2022)	43.0
	+ Self-consistency (40 paths)	**69.2**

More refinements

More complex search at test time

e.g., Tree of thought (Yao et al., 2023)

Encouraging models to reason at training time

Incorporating verifiers

e.g., STaR (Zelikman et al. 2022), ReST (Gulcehre et al., 2023), R1 (DeepSeek, 2025)

Rewarding thinking directly

e.g., Process Reward (Uesato et al., 2022; Lightman et al., 2023)

Teaching models to search

e.g., incorporating MCTS during preference learning (Xie et al., 2024)

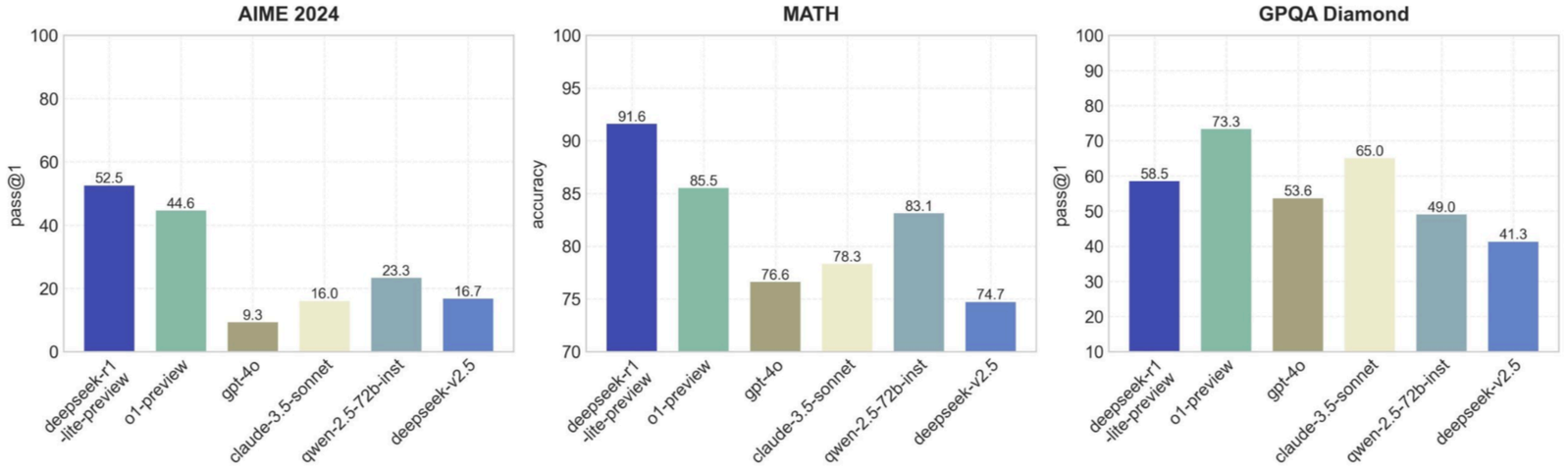
An unexpected success with verifiers

Reminder from Joyee and Wanqiao last quarter: **DeepSeek-R1**

Main idea: Direct RL on correctness

Model automatically learns to output a lot of stuff (even though not explicitly told to!)

Drastically improves performance on reasoning benchmarks



Thank you (and please sign up!)

Sign-up sheet: https://tinyurl.com/reform-S25

Mailing list: reform-ml-list@stanford.edu

Contact: andrewi@stanford.edu, saberi@stanford.edu

Presenter signup

Ideas doc & Reading list