

COMP 3361 Natural Language Processing

Lecture 19: Intro to Advanced Topics

Spring 2025

Tentative schedule

- 1% extra credits.
 - Paper readings
 - Attend the talk in person (attendance will be taken)
 - Ask questions

• Participate in two for the class participation (2% for each) + 1 more for

Retrieval-augmented LLMs

- How do we keep LLMs update-to-date without further training?
- Alleviate problems of hallucinations, lack of attributions, copyright in LLMs
- Smaller size, better performance?



Instruction tuning for LLMs

- How do we make LLMs more controllable by following specific instructions?
- Why following human instructions is important?
- That is the reason why you can prompt a single LLM to perform many tasks !!!



Finetuned Language Models Are Zero-Shot Learners Self-Instruct: Aligning Language Model with Self-Generated Instructions

Inference on unseen task type
Input (Natural Language Inference)
Premise: At my age you will probably have learnt one lesson.
Hypothesis: It's not certain how many lessons you'll learn by your thirties.
Does the premise entail the hypothesis?
OPTIONS: -vesit is not possible to tellno
FLAN Response
It is not possible to tell

LLM alignment

- How do we create LLMs that behaves in accordance with what a human wants?
- How can we ensure powerful AI systems remain aligned with human values and reliable in their reasoning?
- To build trustworthy and aligned Al systems as their capabilities continue to grow.

Zhiqing Sun

孙之清

Hey there, welcome!

I am a final-year Ph.D. candidate at CMU LTI, advised by Prof. Yiming Yang. My research is generously supported by the Google PhD Fellowship in Natural Language Processing (2023) and the OpenAI Superalignment Fast Grants (2024). I received my B.S. in Computer Science from Peking University.

Update (Feb 2025): I've joined OpenAI, where I trained the LLM that powers Deep Research, our latest AI agent that Sam estimates can do "a singledigit percentage of all economically valuable tasks in the world."

Besides Deep Research, my recent research interests include scalable reasoning beyond human supervision, scalable alignment, and scalable training of more capable and trustworthy AI agents.



Research Scientist at OpenAI PhD Candidate at CMU -> SCS -> LTI

ø ...

https://www.cs.cmu.edu/~zhiqings



Sam Altman 🤣 💩 @sama

congrats to the team, especially <u>@isafulf</u> and <u>@EdwardSun0909</u>, for building an incredible product.

my very approximate vibe is that it can do a single-digit percentage of all economically valuable tasks in the world, which is a wild milestone.

9:11 AM · Feb 3, 2025 · 1.1M Views

Zhiqing is the lead of Deep Research agent at OpenAl

LLM explainability & reliability

- How can we build explainable AI systems to better understand the behavior of LLMs?
- What methods improve the calibration and reduce the sensitivity of LLMs to ensure reliable outputs?

Yanda Chen

I am a Member of Technical Staff (Research Scientist) at the Alignment Science team at Anthropic. I work on natural language processing, AI safety, and machine learning.

Previously, I did my PhD in Computer Science at Columbia University, where I was very fortunate to be co-advised by Prof. Kathy McKeown, Prof. He He, and Prof. Zhou Yu. I received my bachelor's degree in Computer Science at Columbia University in April 2021.



Email / CV / Google Scholar / Twitter / Github

https://yandachen.github.io

New Anthropic research: Do reasoning models accurately verbalize their reasoning?

Our new paper shows they don't.

This casts doubt on whether monitoring chains-of-thought (CoT) will be enough to reliably catch safety issues.

Reasoning Models Don't Always Say What They Think

Chen et al.



ANTHROP\C

12:31 AM · Apr 4, 2025 · 996.1K Views



Scaling RL for LLM reasoning

- How and why large-scale reinforcement learning can enhance reasoning capabilities in large language models?
- How do learning and search mechanisms enable language model agents to scale effectively for complex reasoning tasks?



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



Jiayi Pan

潘家怡

Hi 👋

I am a second-year PhD student at Berkeley AI Research. I work with Alane Suhr at Berkeley NLP Group.

I enjoy understanding and making things. I try to learn broadly but bet on a single direction at a time.

Recently, I am most excited about **post-training**, in particular, developing scalable methods to evaluate and improve language model agents and reasoners.

https://www.jiayipan.com





LLMs/VLMs beyond chatting Embodied LLMs/VLMs

(database, web browser, systems, physical world)?



LLMs/VLMs beyond chatting

• Not just chatting with you, can we use LLMs/VLMs as brains of intelligent agents that can interact with and learn from humans and real-world environments

LLM/VLMs as agents

- Definition: An **intelligent agent** perceives its environment via sensors and acts rationally upon that environment with its effectors.
- A discrete agent receives **percepts** one at a time, and maps this percept sequence to a sequence of discrete **actions**.
- Properties sensors -Autonomous percepts -Reactive to the environment -Pro-active (goal-directed) environment actions -Interacts with other agents via the environment effectors





Embodiment	Instructions	>	Actions

LLM/VLMs as agents

- Reasoning by generating tokens is a new type of action (vs. actions in external environments)
- Internal environment, where reasoning takes place in an inner monologue fashion
- Self-reflection is a 'meta' reasoning action (i.e., reasoning over the reasoning process), akin to metacognitive functions
- Reasoning is for better acting, by inferring environmental states, retrospection, dynamic replanning, etc.
- Percept and external action spaces are substantially expanded, thanks to using language for communication and multimodal perception



Reasoning with acting

Question: Aside from the Apple Remote, what other device can control the program Apple Remote was originally designed to interact with?

ReAct (Reason + Act) Thought 1: I need to search Apple Remote and find the program it was originally designed to interact with. Act 1: Search [Apple Remote] **Obs 1:** The Apple Remote is a remote control introduced in October 2005 by Apple ... originally designed to control the Front Row media center program(long text omitted) Thought 2: Apple Remote was originally designed to control the Front Row media center program. I need to search Front Row next and find what other device can control it. Act 2: Search [Front Row] **Obs 2:** Could not find [Front Row]. Similar: ['Front Row Seat to Earth', 'Front Row Motorsports', 'Front Row (software)', ...] **Thought 3:** Front Row is not found. I need to search Front Row (software) . **Act 3:** Search [Front Row (software)] **Obs 3:** Front Row is a discontinued media center software ... (long text omitted) Thought 4: Front Row (software) is controlled by an Apple Remote or the keyboard function keys. So the answer is keyboard function keys. **Act 4:** Finish[keyboard function keys]

LLMs/VLMs as agents

- The LLM/VLM agents need to make decisions for solving complex/abstract problems.
- How LLMs/VLMs as agents? How to use LLMs in an interactive software engineering development environment?



Can Language Models Resolve Real-World GitHub Issues? SWE-agent: Agent-Computer Interfaces Enable Automated Software Engineering



👋 Hello! I'm John.

I'm a 1st year PhD student in Computer Science at Stanford University 🎄

I am currently rotating with Professor Ludwig Schmidt, and I worked with Professor Diyi Yang last fall.

Previously, I completed my M.S.E. in Computer Science at Princeton University 🐯, advised by Professor Karthik Narasimhan. Before that, I finished my B.S. in Computer Science at the University of California, Berkeley 🗱

I'm interested in Language Agents, Language Model Evaluation, and Software Engineering.

If you'd like to get in touch, feel free to reach out through Twitter, LinkedIn, or schedule a meeting through Calendly

Thanks for visiting!

https://john-b-yang.github.io



LLMs/VLMs + robotics/embodied Al

- LLMs/VLMs + robotic agents enables LLMs/VLMs to take actions in real-world environments
- Multimodal LMs





LLM efficient training and tuning methods

Core challenges with scale: efficiency

• The computational and storage costs high, how can we reduct the costs?



• The computational and storage costs of LLM tuning and inference are usually too

Efficient LLM methods

- high, how can we reduct the costs?
- Parameter-efficient LM tuning approaches only fine-tune a small number of thereby greatly decreasing the computational and storage costs.



LoRA: Low-Rank Adaptation of Large Language Models FlashAttention: Fast and Memory-Efficient Exact Attention with IO-Awareness

• The computational and storage costs of LLM tuning and inference are usually too

(extra) model parameters while freezing most parameters of the pretrained LLMs,

LLM parallel pretraining

- Bigger models means more compute to train them. How to conduct data and model parallel training?
- Split the data and distribute data batches among replicas of the model.
 Partition the model across GPUs
- DeepSpeed architecture



Other topics

- LM evaluation, data, and benchmarking
- Bias, toxicity, and privacy in LLMs
- • • •