Yuxiao Qu

cohenqu.github.io
Google Scholar: https://scholar.google.com/citations?user=X5SFuyYAAAAJ&hl=en

Education

Carnegie Mellon University — **Pittsburgh, PA**
PhD in Machine Learning, 4.22/4.33 — *May 2024 - Present*
Advisor: Aviral Kumar & Ruslan Salakhutdinov

Carnegie Mellon University — **Pittsburgh, PA**
Master of Science in Machine Learning , 4.28/4.33 — *Aug 2023 - Present*
Advisor: Aviral Kumar

University of Wisconsin - Madison — **Madison, WI**
Bachelor of Science in Computer Science and Mathematics, 4.00/4.00 — *Sep 2020 – Dec 2022*

The Chinese University of Hong Kong — **Sha Tin, Hong Kong**
Major: Computer Science — *Sep 2017 – May 2020*

Working Experiences

Morgridge Institute for Research — **Madison, WI**
Research Software Engineer — *Apr 2023 – Aug 2023*

NVIDIA — **Remote**
Research Intern — *Nov 2025 – Present*

Awards and Fellowships

- 2025 Amazon PhD Fellowship In AI — 2025
- 2025 Qualcomm Innovation Fellowship (*Finalist*) — 2025
- 2021 ICPC North Central NA Regional Contest (*Rank 19th*) — 2021
- Professor Charles K. Kao Research Exchange Scholarship — 2021
- Best Project Award in Undergraduate Summer Research Internship — 2019

Publications

(* indicates Equal Contribution)

1. **Learning to Reason on Hard Problems with Privileged On-Policy Exploration**
   Yuxiao Qu*, Amrith Setlur*, Virginia Smith, Ruslan Salakhutdinov, Aviral Kumar
   Oral at 5th MATH-AI Workshop, NeurIPS 2025
2. **CaRT: Teaching LLM Agents to Know When They Know Enough**
   Grace Liu*, **Yuxiao Qu***, Jeff Schneider, Aarti Singh, Aviral Kumar
3. **Learning to Discover Abstractions for LLM Reasoning**
   Yuxiao Qu*, Anikait Singh*, Yoonho Lee*, Amrith Setlur, Ruslan Salakhutdinov, Chelsea Finn, Aviral Kumar

Spotlight at Workshop on Efficient Systems for Foundation Models, ICML 2025
Oral at RAM 2: Reasoning, Attention & Memory – 10 Years On, COLM 2025

4. **Optimizing Test-Time Compute via Meta Reinforcement Fine-Tuning**
Yuxiao Qu*, Matthew Y.R. Yang*, Amrith Setlur, Lewis Tunstall, Edward Emanuel Beeching, Ruslan Salakhutdinov, Aviral Kumar
Oral at Workshop on Foundation Models in the Wild, ICLR 2025
International Conference on Machine Learning (ICML), 2025
5. **Harnessing Webpage UIs for Text-Rich Visual Understanding**
Junpeng Liu, Tianyue Ou*, Yifan Song*, **Yuxiao Qu***, Wai Lam, Chenyan Xiong, Wenhu Chen, Graham Neubig, Xiang Yue
International Conference on Learning Representations(ICLR), 2025
6. **Recursive Introspection: Teaching Language Model Agents How to Self-Improve**
Yuxiao Qu, Tianjun Zhang, Naman Garg, Aviral Kumar
Oral at Workshop on Large Language Models and Cognition, ICML 2024
Advances in Neural Information Processing Systems (NeurIPS), 2024
7. **Guided Data Augmentation for Offline Reinforcement Learning and Imitation Learning**
Nicholas E. Corrado, **Yuxiao Qu**, John U. Balis, Adam Labiosa, Josiah P. Hanna
Reinforcement Learning Conference (RLC), 2024
8. **FLEE-GNN: A Federated Learning System for Edge-Enhanced Graph Neural Network in Analyzing Geospatial Resilience of Multicommodity Food Flows**
Yuxiao Qu, Jinmeng Rao, Song Gao, Qianheng Zhang, Wei-Lun Chao, Yu Su, Michelle Miller, Alfonso Morales, Patrick Huber
ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery, 2023
9. **Simulation-Acquired Latent Action Spaces for Dynamics Generalization**
Nicholas E. Corrado, **Yuxiao Qu**, Josiah P. Hanna
Proceedings of Machine Learning Research (PMLR), 2022
10. **Polite or Direct? Conversation Design of a Smart Display for Older Adults Based on Politeness Theory**
Yaxin Hu, **Yuxiao Qu**, Adam Maus, Bilge Mutlu
ACM CHI Conference on Human Factors in Computing Systems (CHI), 2022
11. **TreeNet: Deep Point Cloud Embedding for Routing Tree Construction**
Wei Li, **Yuxiao Qu**, Gengjie Chen, Yuzhe Ma, Bei Yu
Best Paper Award at ACM Asian and South Pacific Design Automation Conference (ASPDAC), 2021

Research Blog Posts

1. **IsoCompute Playbook: Optimally Scaling Sampling Compute for RL Training of LLMs** 2026
Zhoujun Cheng*, Yutao Xie*, **Yuxiao Qu***, Amrith Setlur*, Shibo Hao, Varad Pimpalkhute, Tongtong Liang, Feng Yao, Hector Liu, Eric Xing, Virginia Smith, Ruslan Salakhutdinov, Zhiting Hu, Taylor Killian, Aviral Kumar
Blog Post
2. **QED-Nano: Teaching a Tiny Model to Prove Hard Theorems** 2026
LM-Provers

Blog Post

3. **How to Explore to Scale RL Training of LLMs on Hard Problems?** November 2025
   Yuxiao Qu, Amrith Setlur, Virginia Smith, Ruslan Salakhutdinov, Aviral Kumar
   ML@CMU
4. **Optimizing LLM Test-Time Compute Involves Solving a Meta-RL Problem** January 2025
   Amrith Setlur, **Yuxiao Qu**, Matthew Yang, Lunjun Zhang, Virginia Smith, Aviral Kumar
   ML@CMU

Teaching Experiences

- COMP SCI 577: Introduction to Algorithms (*Honor Session*) Fall 2022
- COMP SCI 537: Introduction to Operating Systems Fall 2022

Professional Service

- Reviewer for ICML (2024-2025), NeurIPS (2024-2025), AISTATS (2025), ICLR (2025-2026)