MUXIN LIU

maxliu@g.hmc.edu

EDUCATION

Harvey Mudd College *Claremont, CA*
B.S. in Computer Science and Mathematics *Expected May 2025*
Second Major: Cognitive Science (at Pomona College)
Study Abroad at Aquincum Institute of Technology, Budapest, Hungary *Spring 2024*
GPA: 3.96/4.0, Dean's List All Semesters
Nominated for the CRA Outstanding Undergraduate Researcher Award *Sept 2024*

Relevant Coursework
Advanced Linear Algebra, Intermediate Probability, Geometry and Topology, Programming Languages, Artificial Learning, Neural Networks, Combinatorial Optimization, Complex Networks, Computer Vision, Computer System, Reinforcement Learning, Quantitative Linguistics, AI Ethics, Psycholinguistics, Brain and Behavior
Audit: Causality, Natural Language Processing, Convex Optimization, Statistical Inference

PUBLICATIONS

LILO: Learning interpretable program libraries via compression and documentation .
Gabriel Grand, Lionel Wong, Matthew Bowers, Theo X. Olausson, **Muxin Liu**, Joshua B. Tenenbaum, Jacob Andreas. *ICLR2024*.

Stream of Search (SoS): Learning to Search in Language .
Kanishk Gandhi, Denise Lee, Gabriel Grand, **Muxin Liu**, Winson Cheng, Archit Sharma, Noah D. Goodman. *COLM2024 Oral*.

PREPRINT

Internal Mechanisms for Tracking Belief State in Language Models.
Muxin Liu, Jacob Andreas, and Belinda Li. *Under Review for ToM4AI Workshop at AAAI 2025. To be Submitted to ICML 2025*

ATOMS: Aligning Task Vectors via Orthogonalization and Merging Sequentially .
Muxin Liu. *In Progress*.

PRESENTATIONS

Probing Language Models for Belief Distribution
The Leadership Alliance National Symposium *Jul 2024*
MIT Summer Research Program Poster Symposium *Aug 2024*

Large Language Models as Few-Shot Library Learners
MIT Summer Research Program Poster Symposium *Aug 2023*

Interactive Proof Assistant for Group Theory
Southern California Applied Mathematics Symposium *Apr 2023*
The Southern California - Nevada Section of the Mathematical Association of America *Apr 2023*

RESEARCH EXPERIENCE

Massachusetts Institute of Technology, CSAIL

Intern, MIT Summer Research Program
Supervisor: Jacob Andreas

Internal Mechanisms for Tracking Belief State in Language Models *Jun 2024 - Present*

- Introduced a new dataset of Theory of Mind tasks by Markov Decision Process and modeled each belief as a distribution for belief tracking through a narrative.
- Trained probes and produced visualization on representations to study where and how facts and beliefs are encoded in the hidden states.
- Designed customized metrics for activation patching to understand the mechanism of belief tracking.
- Working directly under PI and having a first-author paper in progress for ICML 2025.

Stream of Search (SoS): Learning to Search in Language *Jan 2024 - Aug 2024*

- Designed a pipeline to generate a synthetic dataset of search trajectories solving linear equations, produced via BFS with a neural heuristic, and included a parser to optimize large-scale generation.
- Fine-tuned models on trajectories across varying task difficulties and search step granularities, conducted fine-grained evaluations of models' search and backtracking capabilities, and compared results with the countdown domain to assess the generality of LLM-learned backtracking.

Large Language Models as Few-Shot Library Learners *Jun 2023 - Nov 2023*

- Ran and evaluated all experiments using natural language task descriptions in place of synthetic ones.
- Developed pipelines for LLMs to generate library abstractions for DSLs and designed evaluation metrics to systematically compare them with abstractions produced by traditional symbolic compression algorithms.
- Manually documented library abstractions and conducted experiments to compare program synthesis performance against LLM-documented abstractions from the original pipeline.

Harvey Mudd College Department of Computer Science

Interactive Proof Assistant for Group Theory *Sept 2022 - Dec 2023*
Undergraduate Researcher, ALPAQA Lab
Supervisor: Lucas Bang

- Engineered a Python-based Domain-Specific Language with custom definitions and tactics from abstract algebra to construct and verify group theory proofs with error checking.
- Developed a web interface for an interactive proof assistant that provides step-by-step, real-time feedback on proof validations, designed for pedagogical purposes to teach formal proofs.
- Researched LLMs with formal proof systems such as Coq and Lean for proof hint generations using auto-formalization, neural searching and first-order logic.

Cultural Diverse Fashion Planner *May 2022 - Sept 2022*
Undergraduate Researcher, Summer Startup
Supervisor: Zachary Dodds

- Led full-stack development of a Flask-based web application for virtual closets and outfit planning, integrating user authentication, interactive calendar, outfit recommendation and social media features.
- Fine-tuned fashion-CLIP with African and Indian fashion data to build culturally diverse outfit recommendation algorithm based on customized descriptions.

Harvey Mudd College Department of Mathematics *Dec 2022 - May 2023*

Undergraduate Researcher
Supervisor: Francis Su

- Conducted an independent topological combinatorics study on multilabeled versions of Sperner's Lemmas and developed an algorithm for multi-agent fair division using constructive proof techniques.

INDUSTRIAL EXPERIENCE

Laserfiche, Long Beach, CA (Remote) *Sept 2024 - Present*
Part of Harvey Mudd College Clinic Program

- Developed LLM prompting pipelines to automate code generation (in Python and D3) for data extraction, cleaning, insight generation, and visualization on semi-structured data.
- Implemented retrieval-augmented generation (RAG) for code reuse and self error correction.
- First-author paper in progress for ACL 2025 on multimodal LLMs for data visualization.

Harman International, Northridge, CA (Remote) *Oct 2023 - Jan 2024*
Audio Research Intern

- Designed PCA-based dimensionality reduction methods and distance metrics for HRTF samples, and developed an algorithm to construct customized HRTFs for different groups based on unique ear features.

Claudius Legal Intelligence, Princeton, New Jersey (Remote) *Sep 2023 - Jan 2024*
NLP Intern

- Applied prompt tuning on LLMs to provide feedback on legal studies research papers, aligning with reviewer comments, and tuned a semantic similarity metric to assign field labels to these papers.

TEACHING AND COMMUNITY ENGAGEMENT

Academic Excellence Program, Harvey Mudd College *Sep 2023 - Present*
Math Facilitator

- Held weekly mentor sessions for 50+ students on homework problems in core math classes, including Multivariable Calculus, Linear Algebra, Discrete Mathematics, and Differential Equations
- Hosted and attended weekly teaching seminars, made shared study guides, organized social activities

Teaching Assistant, Harvey Mudd College *Sep 2022 - Present*
Held TA sessions and graded homework for Discrete Mathematics, Multivariable Calculus, and Algorithms.

Applied Math Club, Claremont Colleges *May 2023 - May 2024*
Vice President

- Connected with external speakers and organized talks, seminars, and social activities to introduce applied math research, facilitate discussion, and connect students with professors and industry professionals.

Women of the Association of Computing Machinery, Claremont Colleges *May 2023 - Dec 2023*
Member

- Facilitated panels and social activities to connect students with professors and share resources on internships, graduate school applications and interviews for underrepresented groups.

Chinese Students and Scholars Association, Claremont Colleges *Sep 2022 - Dec 2023*
Treasurer

- Organized cultural events for 100+ people in the Chinese student community, coordinated panels and professional trips for industry opportunities, and secured funding by partnering with firms and NGOs.