

JIAYI WU

Phone: (+1) 401-743-3391 ◊ Email: jiayi_wu4@brown.edu

Google Scholar ◊ GitHub ◊ LinkedIn ◊ Personal Webpage

EDUCATION

Brown University

Sept 2024 - May 2028 (expected)

Sc.B. Mathematics-Computer Science and A.B. Applied Mathematics

Related Coursework:

- Machine Learning / Artificial Intelligence: CSCI 2952N Advanced Topics in Deep Learning, CSCI 2470 Deep Learning (Graduate Level), CSCI 2952W Critical Data and Machine Learning, APMA 1655 Introduction to Probability and Statistics with Theory, MATH 0540 Linear Algebra with Theory, CSCI 1973 Individual Independent Study (on AI and Formal Theorem Proving)
- Theory / Formal Methods: CSCI 1010 Theory of Computation, CSCI 1715 Formal Proof and Verification, CSCI 1710 Logic for Systems, CSCI 1951Y Using an Interactive Proof Assistant to Do Mathematics, MATH 1530 Abstract Algebra, MATH 1540 Topics in Abstract Algebra

RESEARCH INTERESTS

I hope to build the formal and empirical infrastructure for **trustworthy machine learning** across two complementary dimensions:

- On the *construction* side, model architectures informed by various learning paradigms (neurosymbolic programming in particular), and how they support more robust and interpretable reasoning models.
- On the *guarantee* side, evaluation and auditing pipeline design informed by formal verification, and how they scaffold more aligned, accountable automated systems.

PUBLICATIONS

* represents equal contribution.

- [1] E. Poole-Dayan, **Jiayi Wu**, T. Sorensen, J. Pei, and M. A. Bakker, “Benchmarking overton pluralism in LLMs,” in *The Fourteenth International Conference on Learning Representations (ICLR 2026)*, also accepted to *NeurIPS 2025 Workshop on Evaluating the Evolving LLM Lifecycle: Benchmarks, Emergent Abilities, and Scaling*, 2025. [Online]. Available: <https://arxiv.org/abs/2512.01351>.
- [2] C. J. Li*, **Jiayi Wu***, Z. Mo, A. Qu, Y. Tang, K. I. Zhao, Y. Gan, J. Fan, J. Yu, J. Zhao, P. P. Liang, L. A. A. Pastor, and K. Larson, “Simulating society requires simulating thought,” in *The Thirty-Ninth Annual Conference on Neural Information Processing Systems (NeurIPS 2025) Position Paper Track*, 2025. [Online]. Available: <https://arxiv.org/abs/2506.06958>.
- [3] C. J. Li*, Z. Mo*, Y. Tang, A. Qu, **Jiayi Wu**, K. I. Zhao, Y. Gan, J. Fan, J. Yu, H. Jiang, P. P. Liang, J. Zhao, L. A. A. Pastor, and K. Larson, “HugAgent: Evaluating LLMs in simulating individual-level human reasoning on open-ended tasks,” in *NeurIPS 2025 Workshop on Bridging Language, Agent, and World Models for Reasoning and Planning (Spotlight)*, also accepted to *NeurIPS 2025 Workshop on PersonaLLM: Workshop on LLM Persona Modeling (Oral)* and *NeurIPS 2025 Workshop on Socially Responsible and Trustworthy Foundation Models (ResponsibleFM)*, 2025. [Online]. Available: <https://arxiv.org/abs/2510.15144>.

RESEARCH EXPERIENCE

Pluggable Analyses for Modern Real Systems

Jan 2026 - Present

Brown University

Providence, RI

Undergraduate Teaching and Research Awards (UTRA) project advised by Prof. Nikos Vasilakis, in collaboration with Lukas Lazarek *et al.* at ATLAS Group.

| | |
|---|---------------------------------------|
| Pattern Mining and Automated Tactic Discovery in Theorem Proving <i>Brown University</i> | Oct 2025 - Present Providence, RI |
| Research project (CSCI 1973 Individual Independent Study) advised by Prof. Robert Lewis and Prof. Stephen Bach, in collaboration with Gavin Zhao. | |
| Formalizing and Benchmarking Overton Pluralism in Large Language Models [1] <i>Massachusetts Institute of Technology</i> | May 2025 - Dec 2025 Cambridge, MA |
| Visiting summer researcher project advised by Prof. Michiel Bakker, in collaboration with Elinor Poole-Dayan <i>et al.</i> | |
| Motif of Thoughts (MoT): Reusable Abstractions for Neurosymbolic Reasoning [2] [3] <i>Massachusetts Institute of Technology</i> | Mar 2025 - Present Cambridge, MA |
| Research project in collaboration with Chance Jiajie Li <i>et al.</i> | |
| Algorithmic Fairness Primer, Socially Responsible Computing (SRC) Handbook <i>Brown University</i> | Dec 2024 - Present Providence, RI |
| Research project advised by Prof. Suresh Venkatasubramanian and Prof. Julia Netter at Center for Technological Responsibility, Reimagination and Redesign (CNTR). | |
| Multi-Objective Trade-Off Modeling in Platform-Based Algorithmic Management <i>Brown University</i> | Nov 2024 - May 2025 Providence, RI |
| Research project advised by Prof. Harini Suresh at the Data in Society Collective (DISCO Lab). | |

PROJECT EXPERIENCE

| | |
|---|--|
| Probing Structural Signals in Lean 4 Proof Graphs with GNNs <i>Brown University</i> | Sept 2025 - Dec 2025 Providence, RI |
| Class final project at CSCI 2470 Deep Learning (Graduate Level). | |

| | |
|---|--|
| Formalizing Hartshorne's Foundations of Projective Geometry in Isabelle <i>Brown University</i> | Sept 2025 - Dec 2025 Providence, RI |
| Class final project at CSCI 1951Y Using an Interactive Proof Assistant to Do Mathematics. | |

LEADERSHIP & TEACHING

| | |
|---|---------------------|
| Undergraduate Teaching Assistant , CSCI 1952B Responsible Computer Science in Practice | Jan 2026 - Present |
| Co-President , AI Robotics Ethics Society (AIRES) at Brown | Sep 2025 - Present |
| Undergraduate Teaching Assistant , CSCI 0170 Introduction to Computer Science | Aug 2025 - Dec 2025 |
| AI Governance Panel Director , Brown China Summit | Sep 2024 - Mar 2025 |

TECHNICAL SKILLS

| | |
|------------------------------|--|
| Programming Languages | Java, Python, JavaScript, HTML/CSS, SQL, MATLAB, Lean, Isabelle, ReasonML, Racket, C/C++, Processing |
| Programming Tools | PyTorch, TensorFlow, scikit-learn, NumPy, pandas, Git/GitHub, Linux |