Peiyang Song

1200 E California Blvd, Pasadena, CA
✉ psong@caltech.edu
peiyang-song.github.io

Education

6/2026 (expected) **California Institute of Technology** Pasadena, CA
B.S. in Computer Science & Minor in Robotics
Advisors: Prof. Steven Low & Prof. Günter Niemeyer. GPA: **4.2/4.0**

Research Interests

My research centers on **LLM reasoning**, **agentic AI**, and **neuro-symbolic AI**. This agenda unfolds along one central axis and extends to two broader directions:

- **Central axis:** Integrating LLM reasoning, agentic AI, and neuro-symbolic methods by combining *neural* models (LLMs) with formal *symbolic* systems (e.g., Lean) to advance LLM-based agents for formal reasoning in mathematics and verifiable code.
- **Broader LLM reasoning:** Extending LLM reasoning beyond formal domains, exploring how LLMs can tackle informal reasoning in natural language, inspired by cognitive science principles and studies of human-like reasoning.
- **Broader neuro-symbolic AI:** Building general neuro-symbolic approaches towards fundamental AI systems beyond formal reasoning, such as developing energy-efficient inference and machine translation across idioms and languages.

Work Experience

6/2025 – Present **University of California, Berkeley** Berkeley, CA
Researcher @ Berkeley Artificial Intelligence Research (BAIR) Lab
Advisors: Prof. Dawn Song (UCB), Dr. Jingxuan He (UCB)

6/2024 – Present **Stanford University** Palo Alto, CA
Researcher @ Stanford AI Lab (SAIL) and Computation & Cognition Lab
Advisors: Prof. Noah Goodman (Stanford), Dr. Gabriel Poesia (Harvard)

2/2023 – 2/2025 **California Institute of Technology** Pasadena, CA
Research Fellow @ Anima AI+Science Lab
Advisors: Prof. Anima Anandkumar (Caltech), Dr. Kaiyu Yang (Meta)

11/2022 – 6/2024 **University of California, Santa Barbara** Santa Barbara, CA

Researcher @ Computer Architecture Lab (ArchLab)

Advisors: Prof. Timothy Sherwood (UCSB), Dr. Jeremy Lau (Google)

Selected Publications

Preprint **Energy-Aware Temporal Function Approximation**

Peiyang Song, Rhys Gretsch, Jeremy Lau, and Timothy Sherwood

In Submission, Manuscript Available upon Request

Preprint **A Survey on Large Language Model Reasoning Failures**

Peiyang Song*, Pengrui Han*, Noah Goodman (* Equal Contribution)

ICML AI for Math (AI4MATH) Workshop, 2025; Under Journal Review

Preprint **The Personality Illusion: Revealing Dissociation Between Self-Reports & Behavior in LLMs**

Pengrui Han*, Rafal D. Kocielnik*, Peiyang Song, Ramit Debnath, Dean Mobbs, Anima Anandkumar, R. Michael Alvarez

NeurIPS LAW Workshop: Bridging Language, Agent, and World Models, 2025; Under Conference Review

Preprint **LeanProgress: Guiding Search for Neural Theorem Proving via Proof Progress Prediction**

Suozhi Huang, Peiyang Song, Robert Joseph George, Anima Anandkumar

ICLR VerifAI: AI Verification in the Wild Workshop, 2025; Under Journal Review

IEEE Micro 2025 **Delay Space Arithmetic and Architecture**

Rhys Gretsch, Peiyang Song, Advait Madhavan, Jeremy Lau, Timothy Sherwood

IEEE Micro, 2025, **Top Picks**

ICLR 2025 **LeanAgent: Lifelong Learning for Formal Theorem Proving**

Adarsh Kumarappan*, Mo Tiwari*, Peiyang Song, Robert Joseph George, Chaowei Xiao, Anima Anandkumar

International Conference on Learning Representations (ICLR), 2025

NeuS 2025 **Lean Copilot: Large Language Models as Copilots for Theorem Proving in Lean**

Peiyang Song, Kaiyu Yang, Anima Anandkumar

International Conference on Neuro-Symbolic Systems (NeuS), 2025

1.2k+ stars on Github, ranking 2nd after Mathlib4 among all Lean projects

EMNLP 2024 **Creative and Context-Aware Translation of East Asian Idioms with GPT-4**

Kenan Tang*, Peiyang Song*, Yao Qin, Xifeng Yan (* Equal Contribution)

Findings of Empirical Methods in Natural Language Processing (EMNLP), 2024

EMNLP 2024 **In-Context Learning May Not Elicit Trustworthy Reasoning: A-Not-B Errors in Pretrained Language Models**

Pengrui Han*, Peiyang Song*, Haofei Yu, Jiaxuan You (* Equal Contribution)

Findings of Empirical Methods in Natural Language Processing (EMNLP), 2024

ASPLOS 2024 **Energy Efficient Convolution with Temporal Arithmetic**

Rhys Gretsch, Peiyang Song, Advait Madhavan, Jeremy Lau, Timothy Sherwood

ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS), 2024

NeurIPS 2023 **LeanDojo: Theorem Proving with Retrieval-Augmented Language Models**

Kaiyu Yang, Aidan Swope, Alex Gu, Rahul Chalamala, Peiyang Song, Shixing Yu, Saad Godil, Ryan Prenger, Anima Anandkumar

Neural Information Processing Systems (NeurIPS), 2023, **Oral Presentation**

Selected Awards

5/2025 **ICLR Notable Reviewer Award**

4/2025 **George W. Housner Student Discovery Fund**

2/2025 **IEEE Micro Top Pick Award**

8/2023 **Early Research Scholarship**

4/2023 **Caltech SURF Award**

Selected Media

2024 **Mathematicians' Newest Assistants Are Artificially Intelligent**

Scientific American

2024 **LeanAgent: The First Life-Long Learning Agent for Formal Theorem Proving in Lean**

MarkTechPost

2024 **Lean Copilot: An AI Tool that Allows Large Language Models (LLMs) to be used in Lean for Proof Automation**
MarkTechPost

2023 **Can LLMs Generate Mathematical Proofs that can be Rigorously Checked?**
MarkTechPost

Invited Talks & Tutorials

LLM Reasoning for Formal Math and Verifiable Code

10/2025 Carnegie Mellon University L3 Lab

Tutorial: Neuro-Symbolic Theorem Proving with Lean

9/2024 3rd Neuro-Symbolic AI Summer School (NSSS)

Towards An AI Mathematician

12/2023 UC Santa Barbara NLP Lab
11/2023 CCS Research & Creative Activities Conference (RACA-CON)
8/2023 Caltech SURF Seminar Day

Teaching Experience

Fall 2025 **ME/CS/EE 133A: Robotics – Kinematics**
California Institute of Technology

Academic Services

Reviewer Conference on Neural Information Processing Systems (NeurIPS)
International Conference on Learning Representations (ICLR)
Association for Computational Linguistics Rolling Review (ARR)
Annual Meeting of the Association for Computational Linguistics (ACL)
Conference on Empirical Methods in Natural Language Processing (EMNLP)
International Joint Conference on Natural Language Processing (IJCNLP)
Asia-Pacific Chapter of the Association for Computational Linguistics (AACL)

NeurIPS Mathematical Reasoning and AI (MATH-AI) Workshop

NeurIPS Workshop on Deep Learning for Code (DL4C)
NeurIPS Workshop on Behavioral Machine Learning
ICLR VerifAI: AI Verification in the Wild Workshop
ICLR Workshop on Representational Alignment (Re-Align)
ICML AI for Math (AI4MATH) Workshop
ICML Workshop on LLMs and Cognition (LLM-Cognition)
ICML Workshop on Assessing World Models
ICML Workshop on on Models of Human Feedback for AI Alignment (MoFA)

Caltech Admissions Ambassador @ Caltech Undergraduate Admissions Office
First-Year Caltech Connector (FCC) @ Student & Family Engagement Office

Organizing Staff Agentic AI Summit 2025 @ UC Berkeley

Languages

Programming Python, C++, Lean 4, Java, C, PASCAL, OCaml, C#

Natural English (TOEFL 117/120), Chinese (Native)