Ziyang Li

https://liby99.github.io/

Email : liby99@seas.upenn.edu
Mobile : +1-858-699-3237

Education

- **University of Pennsylvania** — Philadelphia, PA
  Ph.D. Computer and Information Science, Advisor: Mayur Naik; GPA: 4.0/4.0 — *Jul 2019 – Present*
- **University of California – San Diego** — La Jolla, CA
  B.S. Computer Science (GPA: 3.9/4.0); B.S. Mathematics (GPA: 3.7/4.0) — *Sep 2015 – Jun 2019*

Research Interest

I am a PhD candidate in Computer Science at the University of Pennsylvania, advised by Prof. Mayur Naik. My primary area of research is programming languages. My research also spans the fields of machine learning and security. My dissertation research is focused on neurosymbolic programming, an emerging paradigm that bridges the gap between deep learning and logical reasoning. To this end, I have developed Scallop, a general-purpose neurosymbolic programming language and compiler toolchain. Scallop has been used to develop diverse applications in the domains of natural language processing (NLP), computer vision (CV), cybersecurity, clinical-decision making, and bioinformatics. I was awarded the AWS Fellowship in 2023 for my research on trustworthy AI.

Publications

Pre-prints:

- **Lobster: A GPU-Accelerated Framework for Neurosymbolic Programming**
  Paul Biberstein, Ziyang Li, Joseph Devietti, Mayur Naik
  In Submission, [Preprint]
- **LLM-Assisted Static Analysis for Detecting Security Vulnerabilities**
  Ziyang Li, Saikat Dutta, Mayur Naik
  In Submission, [arXiv]
- **LASER: A Neuro-Symbolic Framework for Learning Spatio-Temporal Scene Graphs with Weak Supervision**
  Jiani Huang, Ziyang Li, Mayur Naik, Ser-Nam Lim
  In Submission, [arXiv]

Monographs:

- **Neurosymbolic Programming in Scallop: Principle and Practice**
  Ziyang Li, Jiani Huang, Jason Liu, Mayur Naik
  Invited Monograph, Foundations and Trends in Programming Languages, 2024

Conferences and Journals:

- **Understanding the Effectiveness of Large Language Models in Detecting Security Vulnerabilities**
  Avishree Khare, Saikat Dutta, Ziyang Li, Alaia Solko-Breslin, Rajeev Alur, Mayur Naik
  ICST 2025, [arXiv]
- **Data-Efficient Learning with Neural Programs**
  Alaia Solko-Breslin, Seewon Choi, Ziyang Li, Neelay Velingker, Rajeev Alur, Mayur Naik, Eric Wong
  NeurIPS 2024

- **Crowd-sourced machine learning prediction of Long COVID using data from the National COVID Cohort Collaborative**
  Timothy Bergquist et al., . . . , Neelay Velingker, Ziyang Li, Yinjun Wu, Jiani Huang, Adam Stein, Emily J. Getzen, Qi Long, Mayur Naik, Ravi B. Parikh, . . .
  eBioMedicine 2024, **NIH L3C Honorable Mention Award**

- **TYGR: Type Inference on Stripped Binaries using Graph Neural Networks**
  Ziyang Li, Chang Zhu*, Anton Xue, Ati Priya Bajaj, William Gibbs, Yibo Liu, Rajeev Alur, Tiffany Bao, Hanjun Dai, Adam Doupé, Mayur Naik, Yan Shoshitaishvili, Ruoyu Wang, Aravind Machiry*
  USENIX Security 2024

- **DISCRET: Synthesizing Faithful Explanations For Treatment Effect Estimation**
  Yinjun Wu, Mayank Keoliya, Kan Chen, Neelay Velingker, Ziyang Li, Emily J Getzen, Qi Long, Mayur Naik, Ravi B Parikh, Eric Wong
  ICML 2024, **Spotlight**

- **Relational Programming with Foundation Models**
  Ziyang Li, Jiani Huang, Jason Liu, Felix Zhu, Eric Zhao, William Dodds, Neelay Velingker, Rajeev Alur, Mayur Naik
  AAAI 2024

- **Improved Logical Reasoning of Language Models via Differentiable Symbolic Programming**
  Jiani Huang, Hanlin Zhang*, Ziyang Li, Mayur Naik, Eric Xing*
  ACL-Findings 2023

- **Scallop: a Language for Neurosymbolic Programming**
  Ziyang Li, Jiani Huang, Mayur Naik
  PLDI 2023

- **Scallop: From Probabilistic Deductive Databases to Scalable Differentiable Reasoning**
  Jiani Huang, Ziyang Li*, Binghong Chen, Karan Samel, Mayur Naik, Le Song, Xujie Si*
  NeurIPS 2021

- **ARBITRAR: User-Guided API Misuse Detection**
  Ziyang Li, Aravind Machiry, Binghong Chen, Mayur Naik, Ke Wang, Le Song
  IEEE S&P 2021

- **HOPPITY: Learning Graph Transformations to Detect and Fix Bugs in Programs**
  Elizabeth Dinella, Hanjun Dai, Ziyang Li, Mayur Naik, Le Song, Ke Wang
  ICLR 2020, **Spotlight**

Research Experiences

- **Research Assistant, UPenn** — University of Pennsylvania
  Advisor: Mayur Naik — *July 2019 – Present*

- **Undergraduate Research Assistant, UCSD PL** — University of California – San Diego
  Advisor: Sorin Lerner — *Sep 2018 – Jun 2019*

- **Undergraduate Research Assistant, UCSD VISCOMP** — University of California – San Diego
  Advisor: Ravi Ramamoorthi — *Mar 2018 – Jun 2019*

- **Undergraduate Research Intern, UCSD Design Lab** — University of California – San Diego
  Advisor: Scott Klemmer — *Jun 2017 – Jun 2018*

Invited Talks and Tutorials

- **Scallop Tutorial** *Summer School of Neurosymbolic Programming* — Salem, MA, *Jun 2024*
- **Invited Talk on Scallop** *Columbia University PL Seminar* — New York, NY, *Oct 2024*
- **Invited Talk on Scallop** *UT Austin PL Seminar* — Austin, TX, *Sep 2024*
- **Neurosymbolic AI** *Guest lecture in Trustworthy AI, UPenn CIS* — Philadelphia, PA, *Apr 17, 2024*
- **Invited Talk on Scallop** *Peking University PL Seminar* — Peking, China, *Dec 10, 2023*
- **Invited Talk on Scallop** *Purdue University PL Seminar* — West Lafayette, IN, *Nov 9, 2023*
- **Invited Talk on Scallop** *KDD'2023* — Los Angeles, CA, *Aug 7, 2023*
- **Tutorial on Scallop** *PLDI'2023* — Orlando, FL, *Jun 17, 2023*
- **Tutorial on Neurosymbolic Methods** *LOG'2022* — Online, *Oct, 2022*
- **Tutorial on Scallop** *Summer Shool of Formal Techniques* — Mountain View, CA, *Jun 1, 2022*

Working Experiences

- **Relational AI** *Research Intern, Mentor: Hung Q. Ngo* — Virtual, *May 2021 – August 2021*
- **Visa, Inc.** *Research Intern, Mentor: Ke Wang* — Virtual, *May 2020 – July 2020*
- **Coursera, Inc.** *Front-end Engineer Intern* — Mountain View, CA, *Jun 2018 - Sep 2018*
- **Deep Media, Ltd.** *Full-stack Engineer Intern* — Shenzhen, China, *Sep 2016 – Jan 2017*
- **Yobs Technology** *Full-stack Engineer Intern* — Los Angeles, CA, *Jan 2016 – Sep 2016*
- **Easyhin** *Front-end Engineer Intern* — Shenzhen, China, *Aug 2015 – Sep 2015*

Fellowships

- **Amazon Web Service Fellowship** *For work on Trustworthy AI* — Amazon, *May 2023*
- **KPCB Fellows 2018** *Engineering Fellows (<2%)* — San Francisco, *June 2018*

Teaching Experiences

- **Teaching Assistant** *CIS 7000, Large Language Models* — University of Pennsylvania, *2024*
- **Teaching Assistant** *CIS 5470, Software Analysis* — University of Pennsylvania, *2020;21;22;23;25*
- **Tutor** *CSE 190, Virtual Reality Technology* — University of California – San Diego, *Spring 2019*
- **Tutor** *CSE 165, 3D User Interaction* — University of California – San Diego, *Winter 2019*
- **Tutor** *CSE 130, Programming Language* — University of California – San Diego, *Fall 2018*
- **Tutor** *CSE 163, Advanced Computer Graphics* — University of California – San Diego, *Spring 2018*
- **Tutor** *CSE 167, Intro to Computer Graphics* — University of California – San Diego, *Winter 2018*
- **Tutor** *CSE 12, Data Structure* — University of California – San Diego, *Winter 2017*

Academic Service

- **Workshop Co-organizer** — *TACPS Workshop 2025*
- **Reviewer** — *ICLR 2024-2025, NeurIPS 2023-2024, ICML 2024, AAAI 2024-2025, ACL ARR 2023-2024*
- **Artifact Evaluation Committee** — *USENIX Security 2025*

Selected Side-Projects

Programming Languages and Program Analysis Tools

- **Probabilistic DataLog Engine**: A probabilistic datalog engine with high performance optimizations oriented towards machine learning applications. Written in Rust.
- **Under-constrained Symbolic Execution Engine**: High performance under-constrained symbolic execution engine for LLVM IR written in Rust. Used in Arbitrar.
- **LLVM IR Binding for Rust**: Safe LLVM Binding for Rust. Used in Arbitrar. [Github]
- **Menhera**: A TypeScript-like functional programming language compiler written in OCaml. [Github]

Rendering, Animation, and Simulations:

- **Fourier Depth of Field**: Fourier transform based depth of field analysis for RayTracer. [Github]
- **Rotamina**: Character animator and simulator with GUI. Written in C++. [Github]
- **MPM-RS**: Material point method for simulating fluid and soft-body dynamics. Written in Rust. [Github]
- **Geometry Sketchpad**: Geometry sketching GUI application written in Rust. [Github]
- **AoSoA Storage**: Array-of-struct-of-array storage system for high performance parallel computing with Kokkos and Cabana. Designed for physics simulation applications. Used by UPenn CG Group. [Github]

Video Games and VR Applications:

- **VR Piano**: VR Application for recording virtual characters playing the Piano, connecting MIDI keyboards and body tracking systems. Written in Unity.
- **Naruhodo**: An 3D story puzzle game engine made in Unity for easy level design. [Github]
- **Neon Ping Pong**: VR Ping Pong Game written in C++. [Website] [Video]
- **Space Escape**: VR Room Escape Puzzle Game settled in Space Station. Developed in Unity. [Website] [Video]

Web Applications:

- **inso.link**: A mirror download site for OSU! beatmaps for Chinese players. Hosted and maintained since 2016 and has 30K users while supporting >2M downloads. [Website] [Status Site]
- **saemanga.com**: A minimalistic online manga reader. Had >1K users. (2016-2020, currently out-of-service)

Skills

- **Languages**: Rust, C++/C, Python, C#, TypeScript/JavaScript, OCaml, Java, Haskell, Coq
- **Libraries/Engines/Tools**: PyTorch, Unity, Unreal Engine 4/5, React, ExpressJs, Asp.net
- **Design**: Adobe Photoshop, Final Cut Pro, Premiere, After Effects, Illustrator, Blender, Cinema 4D
- **Audio/Music**: Logic Pro, Ableton

Last update: Jan 6, 2024