

Shraddha Barke

Microsoft Research Redmond Lab
Microsoft Building 99
✉ sbarke [at] microsoft.com
shraddhabarke.github.io

Research Agenda

My research is at the intersection of software engineering and artificial intelligence and focuses on making AI trustworthy. Lately I have been working on post-training models for complex program reasoning tasks and building agentic pipelines for trustworthy code generation. I also work on ensuring the safety and reliability of tool-calling agents through invariant generation Formal methods, Artificial Intelligence, Human Computer Interaction.

Employment

- 2024-Present **Senior Researcher, RiSE (Research in Software Engineering), Microsoft Research**, Redmond, WA.
Research work includes AI for proof generation, training AI models for program-reasoning tasks using RL and improving the reliability of AI agents.
- 2018-2024 **PhD Candidate, University of California San Diego**, San Diego, CA.
Advisor: Nadia Polikarpova. Graduate research in programming languages and artificial intelligence.
- Fall 22–Winter 23 **PhD Resident, Pitchfork, Google Labs**, Mountain View, CA.
Collaborators: Joel Galenson, Jonathan Malmaud, Rishabh Singh. Project: Automatic code migration system driven by LLMs that upgrades deprecated library code and code translation.
- Summer 2022 **Graduate Research Intern, Calc Intelligence Team, Microsoft Research**, Cambridge, UK (Remote).
Collaborators: Christian Poelitz, Andrew Gordon, Ben Zorn, Jack Williams, José Cambronero, Carina Negreanu. Project: System that synthesizes spreadsheet programs over tabular data from natural-language and explores LLM-based code generation for data-centric tasks.
- Summer 2019 **Graduate Research Intern, PROSE, Microsoft Research**, Redmond, WA.
Collaborators: Xiang Gao, Alan Leung, Arjun Radhakrishnan, Gustavo Soares, Ashish Tiwari, Sumit Gulwani. Project: Extension of BluePencil, an on-the-fly synthesis engine that automatically detects repetitive edits and synthesizes transformations for other locations in the code, based on only one instance of the edit. Resulted in an OOPSLA paper, and a US patent.
- Summer 2017 **Student Researcher, Technische Universität Kaiserslautern**, Kaiserslautern, Germany.
Collaborators: Annette Bieniusa and Peter Zeller. Project: Extension of Repliss, a formal verification system that supports conflict resolution data types (CRDTs) to manage the state of replicas and guarantee safety of applications built on weakly consistent databases. Developed a collaborative editing application as a testing prototype for Antidote, a distributed database with replicated data types.
- Summer 2016 **Research Intern, Tata Institute of Fundamental Research**, Mumbai, India.
Collaborators: Ashutosh Gupta. Thesis Project: A fence synthesis system for weak memory models that uses a constraint-based approach to guarantee minimal number of inserted fences. Evaluated the tool, Orbis against state of the art fence synthesis tools and achieved optimal fence placement. **Skills: SMT Solvers, C++**
- Winter 2015 **Software Development Intern, Outreachy, Linux Kernel**, Remote.
Collaborators: Greg Hartman. Project: Linux kernel staging drivers cleanup, contributed over 500 patches to kernel code that involved updating API interfaces and fixing bugs using static analysis tools, Coccinelle and Checkpatch. Ranked 6th among the most active developers in terms of patches for 4.4 Linux version by LWN.net.

Education

- 2018–2024 **Ph.D. Computer Science, University of California**, San Diego, CA, USA.
Advisor: Dr. Nadia Polikarpova. Thesis: NeuroSymbolic Program Synthesis for Low-Resource Languages
- 2018–2020 **M.S. Computer Science, University of California**, San Diego, CA, USA.
Advisor: Dr. Nadia Polikarpova. Thesis: Guided Program Synthesis Using Probabilistic Models.
- 2013–2017 **B.E. Electronics Engineering, Birla Institute of Technology and Science (BITS) Pilani**, Goa, India.
Advisor: Dr. Ashutosh Gupta. Thesis: Fence Synthesis for Weak Memory Models.

Peer-reviewed Publications

Published Conference Papers

- NeurIPS 2024 **Shraddha Barke**, Emmanuel Anaya-Gonzalez, Saketh Kasibatla, Taylor Berg-Kirkpatrick, Nadia Polikarpova. HYSYNTH: Context-Free LLM Approximation for Guiding Program Synthesis. *2024 Conference on Neural Information Processing Systems Event* in Vancouver, Canada.
- NAACL 2024 **Shraddha Barke**, Christian Poelitz, Carina Suzana Negreanu, Benjamin Zorn, José Cambronero, (Findings) Andrew D. Gordon, Vu Le, Elnaz Nouri, Nadia Polikarpova, Advait Sarkar, Brian Slininger, Neil Toronto, Jack Williams. Solving Data-Centric Tasks with Large Language Models. *North American Chapter of the Association for Computational Linguistics*, 2024 in Mexico City, Mexico.
- OOPSLA 2023 **Shraddha Barke**^{*}, Michael James^{*}, Nadia Polikarpova. Grounded Copilot: How Programmers Interact with Code-Generating Models. *Object-Oriented Programming, Systems, Languages, and Applications*, 2023 in Cascais, Portugal. **Distinguished Paper Award**
- OOPSLA 2021 Kusra Ferdowsi, **Shraddha Barke**, Hila Peleg, Nadia Polikarpova, Sorin Lerner. LooPy: Interactive Program Synthesis with Control Structures. *Object-Oriented Programming, Systems, Languages, and Applications*, 2021 in Chicago, USA.
- OOPSLA 2020 **Shraddha Barke**, Hila Peleg, Nadia Polikarpova. Just-in-Time Learning for Bottom-Up Enumerative Synthesis. *Object-Oriented Programming, Systems, Languages, and Applications*, 2020 (Remote).
- OOPSLA 2020 Xiang Gao, **Shraddha Barke**, Arjun Radhakrishna, Gustavo Soares, Sumit Gulwani, Alan Leung, Nachiappan Nagappan, Ashish Tiwari. Feedback-Driven Semi-Supervised Synthesis of Program Transformations. *Object-Oriented Programming, Systems, Languages, and Applications*, 2020.
- EMNLP 2019 **Shraddha Barke**, Rose Kunkel, Eric Meinhardt, Nadia Polikarpova, Eric Bakovic, and Leon Bergen. Constraint-based Learning of Phonological Processes. *Conference on Empirical Methods in Natural Language Processing*, 2019 in Chek Lap Kok, Hong Kong.

Book Chapters

- 2018 Annette Bieniusa, Peter Zeller, and **Shraddha Barke**. Collaborative Work Management with a Highly-Available Kanban Board. *Principled Software Development - Essays Dedicated to Arnd Poetzsch-Heffter on the Occasion of his 60th Birthday*, 2018.

Awards and Scholarships

- 2023 ACM SIGPLAN Distinguished Paper Award at OOPSLA 2023
- 2022 ACM SIGPLAN PAC Award
- 2022 CRA-WP Grad Cohort for IDEALS Invite, San Diego, CA.
- 2019 CSE Award for Contributions to Diversity - UC San Diego
- 2019 PLMW Scholarship for PLDI 2019, Phoenix
- 2017 PLMW Scholarship for POPL 2017, Paris
- 2016 Ranked 6th most active developer for Linux 4.4 by LWN.net

Talks

- 2024 “HySynth: Context-Free LLM Approximation for Guiding Program Synthesis”, NeurIPS 2024, Canada.
- 2024 “Solving Data-centric Tasks using Large Language Models”, NAACL 2024, Mexico City.
- 2024 “Better Learning through Programming Languages”, Microsoft Research, Redmond.
- 2023 “How Large Language Models Revolutionized Program Synthesis”, University of Utah.
- 2022 “Grounded Copilot: How Programmers Interact with Code-Generating Models”, MSR Redmond.
- 2022 “Constraint-based Learning of Phonological Processes,” CSE Research Open House, UC San Diego.
- 2022 “Just-in-Time Learning for Bottom-Up Enumerative Synthesis,” PLDI 2022, San Diego, CA.
- 2020 “Just-in-Time Learning for Bottom-Up Enumerative Synthesis,” OOPSLA 2020, Remote.
- 2019 “Constraint-based Learning of Phonological Processes,” EMNLP 2019, Hong Kong.
- 2016 “Fence Synthesis for Weak Memory Models,” Indian Institute of Science, Bangalore.
- 2016 “Diving into Open Source with Linux Kernel” Workshop, Grace Hopper 2016, Bangalore.

Teaching

- Spring 2024 Teaching Assistant, CSE 130 Programming Languages, UC San Diego
- Fall 2023 Guest Lecturer, CS 6353 Deep Learning, University of Utah
- Spring 2023 Teaching Assistant, CSE 130 Programming Languages, UC San Diego
- Winter 2023 Guest Lecturer, CSE 291 Program Synthesis, UC San Diego
- Spring 2022 Teaching Assistant, CSE 130 Programming Languages, UC San Diego
- Fall 2020 Teaching Assistant, CSE 130 Programming Languages, UC San Diego
- Fall 2019 Teaching Assistant, CSE 130 Programming Languages, UC San Diego

US Patents

- 2024 Benjamin Goth Zorn, Carina Suzana Negreanu, Neil Blunt Toronto, Brian Paul Slininger, Andrew Donald Gordon, Advait Sarkar, Elnaz Nouri, Vu Minh LE, Christian Leopold Bejamin Poelitz, Shraddha Govind Barke, Sruti Srinivasa Ragavan. Model capability extraction, *US Patent Number 17/969922*, 2024.
- 2022 Shraddha Govind Barke, Xiang Gao, Sumit Gulwani, Alan Thomas Leung, Nachiappan Nagappan, Arjun Radhakrishna, Gustavo Araujo Soares, Ashish Tiwari, Mark Alistair Wilson-Thomas. Feedback-driven semi-supervised synthesis of program transformations, *US Patent Number 11513773*, 2022.

Mentorship

- 2023-Present Mentor, SIGPLAN-M
- 2023 Mentor, Undergraduate thesis: Recursive phonological rule discovery, UC San Diego
- 2019-2020 Mentor, Early Research Scholars Program (ERSP), UC San Diego
- 2019-2020 Mentor, GradWIC Mentorship Program, UC San Diego

Service and Leadership

- 2026 Object-oriented Programming, Systems, Languages, and Applications (OOPSLA) Program Committee
- 2026 Foundations of Software Engineering (FSE) Program Committee
- 2025 Reviewer at the ACM Transactions on Programming Languages and Systems
- 2025 Reviewer at the ACM Transactions on Computing Education
- 2025 NeurIPS 2025 Program Committee
- 2025 SANER 2025 Program Committee
- 2024 Workshop Program Committees: HATRA 2025, HATRA 2024, PLATEAU 2024
- 2023-2024 POPL 2023-2024 Student Volunteer Co-Chair
- 2023, 2022 ACL ARR Program Committee
- 2021-2023 Artifact Evaluation Committee: OOPSLA 2021, CAV 2021, ICFP 2023
- 2022-2024 SIGPLAN Long-Term Mentoring Committee
- 2023 ICLR 2023 Deep Learning for Code Program Committee
- 2019-2022 Student Volunteer Committee: ASE 2019, PLDI 2019, PLDI 2022
- 2019-2020 Graduate Women in Computing Coordinator, UC San Diego
- 2018-2020 CSE Diversity Equity and Inclusion (DEI) Committee, UC San Diego
- 2017-2019 Outreachy Linux Kernel Coordinator

Summer Schools and Seminars

- 2023 AI to Assist Mathematical Reasoning: A Workshop (Remote)
- 2022 Summer school on Neurosymbolic Programming, Caltech, Pasadena, CA
- 2021 Synthesis of Models and Systems workshop, Simons Institute for Theory of Computing (Remote)
- 2021 Galois Summer School for Trustworthy Machine Learning, Artificial Intelligence, and Data Science
- 2021 OPLSS Oregon Programming Languages Summer School (Remote)
- 2020 PL + HCI Swimmer School (Remote)