Shraddha Barke

Computer Science Department
University of California, San Diego
✉ *sbarke [at] eng.ucsd.edu*
`shraddhabarke.github.io`

Research Agenda

My research focuses on building human-like AI systems that aim to automate scientific discovery and learn causal theories of the world. To that end, I draw on ideas from program synthesis to develop systems that generalize well, require less data and produce interpretable results. Further, my research also investigates how powerful AI large language models have changed the landscape of programming through the lens of human computer interaction and privacy. Keywords: Program Synthesis and Verification, Artificial Intelligence, Human Computer Interaction.

Education

2018–2024 (Expected) **Ph.D. Computer Science**, *University of California*, San Diego, CA, USA.
Advisor: Dr. Nadia Polikarpova. Thesis: Stay tuned!

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

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.

OOPLSA 2021 Kasra Ferdowsi, **Shraddha Barke**, Hila Peleg, Nadia Polikarpova, Sorin Lerner. LooPy: Interactive Program Synthesis with Control Structures. *Object-Oriented Programming, Systems, Languages, and Applications*, 2021.

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.

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.

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.

In Submission/ Under Preparation Papers

ASE 2023 **Shraddha Barke**, Christian Poelitz, Carina Suzana Negreanu, Benjamin Zorn, José Cambronero, Andrew D. Gordon, Vu Le, Elnaz Nouri, Nadia Polikarpova, Advait Sarkar, Brian Slininger, Neil Toronto, Jack Williams. CodeXData: Solving Data-Centric Tasks with Large Language Models

2019 John Sarracino, **Shraddha Barke**, Hila Peleg, Nadia Polikarpova, and Sorin Lerner. Targeted Program Synthesis for Programming with Invariants.

Professional Employment

2018-Present **University of California San Diego**, *PhD Candidate*, San Diego, CA.
Advisor: Nadia Polikarpova. Graduate research in programming languages and artificial intelligence.

2023-Present **RISE, Microsoft Research**, *Part-time Researcher*, Redmond, WA.
Collaborators: Christian Poelitz, Andrew Gordon, Ben Zorn, Jack Williams, José Cambronero, Carina Negreanu. Project (Ongoing): Generation of synthetic tabular data for LLMs that provides differential privacy guarantees.

Fall 22–Winter 23 **Google X - the moonshot factory**, *PhD Resident, Pitchfork*, Mountain View, CA.
Collaborators: Joel Galenson, Jonathan Malmaud, Rishabh Singh. Project (details confidential): Automatic code migration system driven by LLMs that upgrades deprecated library code and translates across languages.

Summer 2022 **Calc Intelligence Team, Microsoft Research**, *Graduate Research Intern*, Cambridge, UK (Remote).
Collaborators: Christian Poelitz, Andrew Gordon, Ben Zorn, Jack Williams, José Cambronero, Carina Negreanu. Project: CodeXData, a system that synthesizes spreadsheet manipulation programs over tabular data from natural-language description and explores LLM-based code generation for data-centric tasks. Resulted in a paper submission and the filing of a patent.

Summer 2019 **PROSE, Microsoft Research**, *Graduate Research Intern*, Redmond, WA.
Collaborators: Alan Leung, Arjun Radhakrishnan, Gustavo Soares. 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 **Technische Universität Kaiserslautern**, *Student Researcher*, 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 **Tata Institute of Fundamental Research**, *Undergraduate Research Intern*, Mumbai, India.
Collaborators: Ashutosh Gupta. Bachelor's 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 (Glue and Memorax) and achieved optimal fence placement.

Winter 2015 **Outreachy, Linux Kernel**, *Software Development Intern*, 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.

Invited Talks

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
2021 "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 Positions

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

Mentorship

2023 – 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

2023 POPL 2023 Student Volunteer Co-Chair, Boston, MA
2023 ICFP 2023 Artifact Evaluation Committee
2022-Present SIGPLAN Long-Term Mentoring Committee
2023 ICLR 2023 Deep Learning for Code Program Committee
2022 ACL ARR 2022 Program Committee
2022 Faculty Hiring Student Representative, UC San Diego
2022 PLDI 2022 Student Volunteer, San Diego, CA
2021 OOPSLA 2021 Artifact Evaluation Committee
2021 CAV 2021 Artifact Evaluation Committee
2019 ASE 2019 Student Volunteer, San Diego, CA
2019 PLDI 2019 Student Volunteer, Phoenix, AZ
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, UC San Diego
2018 PhD Admissions Student Committee, UC San Diego

Awards and Scholarships

2019 CSE Award for Contributions to Diversity - UC San Diego
2019 PLMW Scholarship for PLDI 2019, Phoenix
2017 PLMW Scholarship for POPL 2017, Paris
2017 VMW Scholarship for CAV 2017, Heidelberg
2016 Ranked 6th most active developer for Linux 4.4 by LWN.net

Summer Schools and Seminars

2022 Summer school on Neurosymbolic Programming, Caltech, Pasadena, CA
2022 CRA-WP Grad Cohort Workshop for IDEALS, San Diego, CA.
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

US Patents

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.