SHANGYIN TAN

shangyin@berkeley.edu ◇ https://shangyit.me

EDUCATION

University of California, Berkeley Aug 2022 - May 2028
Doctor of Philosophy (Ph.D.) in Computer Science *Berkeley, US*
- Research Interests: agents, compound AI systems, programming languages
- Advised by *Koushik Sen* and *Matei Zaharia*

Purdue University Aug 2018 - Dec 2021
Bachelor of Science in Computer Science Honors *West Lafayette, US*
- Graduated with Highest Distinction, GPA: 3.98/4.0, Major GPA: 4.0
- CRA Outstanding Undergraduate Researcher Honorable Mention
- Outstanding Research Effort by an Undergraduate Student Award

WORK EXPERIENCES

Letta Inc. Mar 2025 - Aug 2025
Research Fellow (manager: *Kevin Lin*) *San Francisco, US*
- Contribute to Memory-based Agents, Terminal-use Agents, and general benchmarking

Google DeepMind May 2023 - Present
Student Researcher (manager: *Dan Zheng*, *Ningning Xie*, and *Gordon Plotkin*) *Mountain View, US*
- Build a new language for choice-based learning in JAX

Microsoft Research, Asia March 2022 - Aug 2022
Research Assistant (manager: *Quanlu Zhang* and *Fan Yang*) *Beijing, China*
- Build a program synthesis system for deep learning config file generation
- Develop a deep learning testing framework

Purdue University June 2020 - Dec 2021
Undergraduate Researcher (advised by *Guannan Wei* and *Tiark Rompf*) *West Lafayette, US*
- Compile efficient symbolic executions via multi-stage programming
- Lead the development of multiple *LLVM* symbolic execution compilers

PREPRINTS

1. Anonymous Authors. Terminal-bench: Benchmarking agents on hard, realistic tasks in command line interfaces, 2025. ICLR 2026 Conference Submission
2. **Shangyin Tan**, Kevin Lin, Koushik Sen, and Matei Zaharia. Recovery-bench: Evaluating agentic recovery from mistakes. In *NeurIPS LLM Evaluation Workshop*, 2025
3. Lakshya A. Agrawal, **Shangyin Tan**, Dilara Soylu, Noah Ziems, Rishi Khare, Krista Opsahl-Ong, Arnav Singhvi, Herumb Shandilya, Michael J. Ryan, Meng Jiang, Christopher Potts, Koushik Sen, Alexandros G. Dimakis, Ion Stoica, Dan Klein, Matei Zaharia, and Omar Khattab. Gepa: Reflective prompt evolution can outperform reinforcement learning, 2025. Preprint

4. **Shangyin Tan***, Dan Zheng*, Gordon Plotkin, and Ningning Xie. Choix: Choice-based learning in jax. In *Workshop on ML for Systems at NeurIPS*, 2023
5. Arnav Singhvi*, Manish Shetty*, **Shangyin Tan***, Christopher Potts, Koushik Sen, Matei Zaharia, and Omar Khattab. Dspy assertions: Computational constraints for self-refining language model pipelines, 2024

PUBLICATIONS

1. **Shangyin Tan**, Guannan Wei, Koushik Sen, and Matei Zaharia. Programming large language models with algebraic effect handlers and the selection monad. In *The 1st Workshop on Language Models for Programming Languages (LMPL)*, 2025
2. **Shangyin Tan**, Lakshya A. Agrawal, Arnav Singhvi, Liheng Lai, Michael J. Ryan, Dan Klein, Omar Khattab, Koushik Sen, and Matei Zaharia. Langprobe: a language programs benchmark. In *Empirical Methods in Natural Language Processing (EMNLP)*, 2025
3. Chaofan Shou, **Shangyin Tan**, and Koushik Sen. Ityfuzz: Snapshot-based fuzzer for on-chain smart contract auditing. In *ISSTA*. ACM, 2023
4. Guannan Wei, Songlin Jia, Ruiqi Gao, Haotian Deng, **Shangyin Tan**, Oliver Bračevac, and Tiark Rompf. Compiling parallel symbolic execution with continuations. In *ICSE*. IEEE/ACM, 2023
5. Zhanhui Zhou, Man To Tang, Qiping Pan, **Shangyin Tan**, Xinyu Wang, and Tianyi Zhang. INTENT: interactive tensor transformation synthesis. In *UIST*, pages 89:1–89:16. ACM, 2022
6. **Shangyin Tan**, Guannan Wei, and Tiark Rompf. Towards partially evaluating symbolic interpreters for all (short paper). In *PEPM at POPL*. ACM, 2022
7. Guannan Wei, **Shangyin Tan**, Oliver Bracevac, and Tiark Rompf. LLSC: a parallel symbolic execution compiler for LLVM IR. In *ESEC/SIGSOFT FSE*, pages 1495–1499. ACM, 2021
8. Guannan Wei, Oliver Bracevac, **Shangyin Tan**, and Tiark Rompf. Compiling symbolic execution with staging and algebraic effects. *Proc. ACM Program. Lang.*, 4(OOPSLA):164:1–164:33, 2020

PRESENTATIONS

1. **ACM SIGPLAN Workshop on Partial Evaluation and Program Manipulation**
   Towards Partially Evaluating Symbolic Interpreters for All — Jan 2022
2. **SPLASH 2021 SIGPLAN Papers Track**
   Compiling Symbolic Execution with Staging and Algebraic Effects — Oct 2021
3. **PurPL Reading Group**
   Data types a la carte — Aug 2020

AWARDS

1. CRA Outstanding Undergraduate Researcher Honorable Mention — 2022
2. Outstanding Research Effort by an Undergraduate Student, Purdue University — 2022
3. Corporate Partner Scholarship, Purdue University — 2020

PROFESSIONAL ACTIVITIES

Artifact Evaluation Committee Member

· ACM SIGPLAN Conf. on Programming Language Design and Implementation (PLDI) 2022

Sub-reviewer

· The ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA) 2023

Student Volunteer

· The ACM SIGPLAN conference on Systems, Programming, Languages, and Applications: Software for Humanity (SPLASH) 2021, 2020

OTHER EXPERIENCES

Undergraduate Teaching Assistant Jan 2019 - Jan 2021
Discrete Math, System Programming, Algorithms Analysis, ... *West Lafayette, US*

· Conduct recitations to help students with problem solving
· Advise students in lab debugging
· Monitor online Q&A forums like Piazza

Selected Coding Contests 2018 - 2020
Higher Ranked Participant *Midwest, US*

· 3^{rd} in Tech Challenge Google 2019, Chicago
· 2^{nd} in Sandia Coding Challenge 2018, West Lafayette

SKILLS

Familiar with	C, Scala, Python, C++
Have worked with	Haskell, Coq, X86-64, Java, Javascript, Scheme, LaTeX, LLVM, MatLab
Tools	GDB, Git, QuickCheck, SAT/SMT solvers (Minisat, STP, Z3)

(Skills in the same row are in random order)