

Xujie Si

Research Interests

My research interests span programming languages, software engineering, security, and artificial intelligence. My thesis research focuses on developing deep learning and reinforcement learning techniques for challenging problems in program reasoning that span verification, synthesis, and testing. I am also interested in using formal methods and symbolic reasoning techniques to analyze the robustness and fairness of machine learning models.

Education

- August 2020 **Ph.D., Computer and Information Science.**
University of Pennsylvania, Philadelphia, PA
Advisor: Prof. Mayur Naik
- August 2014 **M.S., Electrical Engineering and Computer Science.**
Vanderbilt University, Nashville, TN
Advisor: Prof. Yuan Xue
- July 2011 **B.E. (with Honors), Software Engineering.**
Nankai University, Tianjin, China

Publications

- CAV'20 **Xujie Si**, Aaditya Naik, Hanjun Dai, Mayur Naik, and Le Song. [Code2Inv: A Deep Learning Framework for Program Verification.](#) In *Proceedings of the 32nd International Conference on Computer Aided Verification, CAV 2020.*
- IJCAI'19 **Xujie Si**, Mukund Raghothaman, Kihong Heo, Mayur Naik. [Synthesizing Datalog Programs using Numerical Relaxation.](#) In *Proceedings of the 28th International Joint Conference on Artificial Intelligence, IJCAI 2019.*
- PLDI'19 Kihong Heo, Mukund Raghothaman, **Xujie Si**, Mayur Naik. [Continuously Reasoning about Programs via Differential Bayesian Inference.](#) In *Proceedings of the ACM SIGPLAN conference on Programming Language Design and Implementation, PLDI 2019 (Distinguished Paper Award).*
- ICLR'19 **Xujie Si**, Yuan Yang, Hanjun Dai, Mayur Naik and Le Song. [Learning a Meta-Solver for Syntax-Guided Program Synthesis.](#) In *Proceedings of the International Conference on Learning Representations, ICLR 2019.*
- NeurIPS'18 **Xujie Si**, Hanjun Dai, Mukund Raghothaman, Mayur Naik and Le Song. [Learning Loop Invariants for Program Verification.](#) In *Proceedings of the Thirty-second Conference on Neural Information Processing Systems, NeurIPS 2018 (Spotlight).*
- FSE'18 **Xujie Si**, Woosuk Lee, Richard Zhang, Aws Albarghouthi, Paris Koutris and Mayur Naik. [Syntax-Guided Synthesis of Datalog Programs.](#) In *Proceedings of the ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering, FSE 2018.*
- ML4P'18 Mukund Raghothaman, Sulekha Kulkarni, Richard Zhang, **Xujie Si**, Kihong Heo, Woosuk Lee and Mayur Naik. [Difflog: Beyond Deductive Methods in Program Analysis.](#) In *Proceedings of the Workshop on Machine Learning for Programming, ML4P 2018.*
- VMCAI'18 Mayur Naik, **Xujie Si**, Xin Zhang and Radu Grigore. [Maximum Satisfiability in Software Analysis: Applications and Techniques.](#) In *Proceedings of 19th International Conference on Verification, Model Checking, and Abstract Interpretation, VMCAI Invited Tutorial 2018.*

- OOPSLA'17 Xin Zhang, Radu Grigore, **Xujie Si** and Mayur Naik. **Effective Interactive Resolution of Static Analysis Alarms**. In *Proceedings of the ACM SIGPLAN conference on Systems, Programming, Languages and Applications: Software for Humanity, SPLASH/OOPSLA 2017*.
- MAPL'17 Xin Zhang, **Xujie Si** and Mayur Naik. **Combining the Logical and the Probabilistic in Program Analysis**. In *Proceedings of the Workshop on Machine Learning and Programming Languages, MAPL 2017*.
- CAV'17 **Xujie Si**, Xin Zhang, Radu Grigore and Mayur Naik. **Maximum Satisfiability in Software Analysis: Applications and Techniques**. In *Proceedings of the 29th International Conference on Computer Aided Verification, CAV Invited Tutorial 2017*.
- USENIX Security'16 Insu Yun, Changwoo Min, **Xujie Si**, Yeongjin Jang, Taesoo Kim and Mayur Naik. **APISan: Sanitizing API Usages through Semantic Cross-checking**. In *Proceedings of the 25th USENIX Security Symposium, 2016*. (**CSAW Best Applied Research Paper Finalist**)
- CP'16 **Xujie Si**, Xin Zhang, Vasco Manquinho, Mikolas Janota, Alexey Ignatiev and Mayur Naik. **On Incremental Core-Guided MaxSAT Solving**. In *Proceedings of the 22nd International Conference on Principles and Practice of Constraint Programming, CP 2016*.
- CODASPY'14 Xiaowei Li, **Xujie Si** and Yuan Xue, **Automated Black-box Detection of Access Control Vulnerabilities in Web Applications**. In *Proceedings of 2014 ACM Conference on Data and Application Security and Privacy, 2014*.

Teaching and Mentoring

- 2019-present [Mentoring] Aaditya Naik (senior undergraduate) and Ziyang Li (first-year PhD student)
- Spring 2019 [Mentoring] Jiani Huang (first-year PhD student, female)
- Fall 2018 [Mentoring] Elizabeth Dinella (first-year PhD student, female)
- 2017-2018 [Mentoring] Richard Zhang (Masters student)
- Spring 2016 [Mentoring] Aditya Kamath (Masters student)
- Spring 2018 [TA] CIS 550 Database and Information Systems, *University of Pennsylvania*
- Fall 2017 [TA] CIS 700 Software Testing and Analysis, *University of Pennsylvania*
- Spring 2016 [TA] CS 6340 Software Testing and Analysis, *Georgia Institute of Technology*
- Fall 2015 [TA] CS 7001 Introduction to Graduate Studies, *Georgia Institute of Technology*
- Spring 2012 [TA] CS 285 Computer Networks, *Vanderbilt University*
- Fall 2011 [TA] CS 212 Discrete Structures, *Vanderbilt University*

Talks

Deep Reinforcement Learning for Program Verification and Synthesis.

- Dec. 2019 IBM Programming Languages Day. Yorktown Heights, NY, USA

Synthesizing Datalog Programs using Numerical Relaxation.

- Sep. 2019 Dagstuhl Seminar 19361 on Logic and Learning. Wadern, Germany.
- Aug. 2019 International Joint Conference on Artificial Intelligence. Macao, China.

Learning Loop Invariants for Program Verification.

- Dec. 2018 Conference on Neural Information Processing Systems. Montreal, Canada
- Sep. 2018 Penn Research in Machine Learning Seminar. Philadelphia, PA, USA

Syntax-Guided Synthesis of Datalog Programs.

- Nov. 2018 Foundations of Software Engineering (FSE). Lake Buena Vista, FL, USA
- Oct. 2018 Penn PLClub. Philadelphia, USA

Selected Awards

- 2019 PLDI Distinguished Paper Award
- 2019 Facebook Fellowship and Emerging Scholars finalist
- 2018 NeurIPS Travel Award
- 2018 ESEC/FSE Travel Fund Award
- 2016 CSAW Best Applied Research Paper finalist
- 2014 World Finalist of the ACM International Collegiate Programming Contest
- 2010 National Scholarship, Nankai University
- 2009 Tianjin Government Scholarship, Nankai University

Work Experience

- 2019 **Research Scientist Intern**, *DeepMind*, London, UK.
May-Oct. Worked on neural program synthesis and deep learning for combinatorial optimization.
- 2015 **Software Engineer Intern**, *Google*, Mountain View, CA.
May-Aug. Worked on analyses of dynamic search Ads and serving algorithms.
- 2014 **Software Engineer Intern**, *Google*, Seattle, WA.
May-Aug. Designed and built a pipeline for analyzing display Ads traffic using Flume C++.

Professional Activities

- NeurIPS'20 Reviewer
- MAPL'20 Program Committee
- POPL'19 External reviews
- CAV'18 External reviews
- CAV'19,'20 Artifact Evaluation Committee
- ICFP'19 Artifact Evaluation Committee
- PLDI'18,'19 Student Volunteer Co-Chair
- POPL'16 Student Volunteer

References

Mayur Naik

Associate Professor
Computer and Information Science
University of Pennsylvania
610 Levine, 3330 Walnut St
Philadelphia, PA 19104, USA
✉ mhnaik@cis.upenn.edu
☎ 215-573-1856

Aws Albarghouthi

Assistant Professor
Computer Sciences Department
University of Wisconsin–Madison
CS-6363, 1210 West Dayton Street
Madison, WI 53706, USA
✉ aws@cs.wisc.edu
☎ 608-262-7946

Rajeev Alur

Zisman Family Professor
Computer and Information Science
University of Pennsylvania
609 Levine, 3330 Walnut St
Philadelphia, PA 19104, USA
✉ alur@cis.upenn.edu
☎ 215-573-7483

Le Song

Associate Professor
Computational Science and Engineering
Georgia Institute of Technology
1340 Klaus, 266 Ferst Drive
Atlanta, GA 30332, USA
✉ lsong@cc.gatech.edu
☎ 404-858-5702