Xin Zhang

Postdoctoral Associate, CSAIL, MIT

32-G716, 77, Massachusetts Ave, Cambridge, MA, 02139

(706)951-7679 | xzhang@csail.mit.edu | people.csail.mit.edu/xzhang/

Research Interests

I am broadly interested in topics related to programming languages and software engineering, with an emphasis on program reasoning techniques (e.g., program analysis, formal methods, type systems) and their interplays with machine learning. On one hand, I develop ML techniques to improve the usability of program reasoning techniques in general. On the other hand, I am interested in designing new program analyses and languages to improve interpretability, safety, fairness, robustness, and generalizability of machine learning models.

Education and Postdoctoral Training

Postdoctoral Associate — *Cambridge, MA, USA*

Massachusetts Institute of Technology — *2017 - present*

- Host: Armando Solar-Lezama

Ph.D. in Computer Science — *Atlanta, GA, USA*

Georgia Institute of Technology, GPA: 3.85/4.0 — *2011 - 2017*

- Thesis: Combining Logical and Probabilistic Reasoning in Program Analysis
- Advisor: Mayur Naik

B.E. in Software Engineering — *Shanghai, China*

Shanghai Jiao Tong University, GPA: 3.7/4.0 (Ranked 1 out of 120) — *2007 - 2011*

Research Experience

Computer Science and Artificial Intelligence Laboratory, MIT — *Cambridge, MA*

Postdoctoral Associate (Host: Armando Solar-Lezama)

- *Explaining Machine Learning System Judgments via Corrections.* — *2017-present*
  Proposal of a method using corrections as actionable explanations to judgments made by a machine learning system when these judgments are undesirable. Design and Implementation of a system based on integer linear programming that realizes this method for neural networks [3].

- *Verifying Algorithmic Fairness.* — *2018-present*
  In collaboration with Osbert Bastani (a researcher at the University of Pennsylvania), design and implementation of the first scalable system for verifying algorithmic fairness [1].

- *Probabilistic Programming with Higher-Order Distribution.* — *2018-present*
  In collaboration with Zenna Tavares (a PhD student at MIT), design and implementation of *Omega*, the first probabilistic programming language that supports conditioning distributional properties via higher-order constructs. Example applications include inferring a classifier that is fair or robust.

- *Probabilistic Programming with Causal Inference.* — *2018-present*
  In collaboration with Zenna Tavares and James Koppel (PhD students at MIT), the first extension to a probabilistic programming language for supporting interventions and counterfactuals.

School of Computer Science, College of Computing, Georgia Tech *Atlanta, GA*

Research Assistant (Advisor: Mayur Naik)

- *User-Guided Program Analysis.* *2014-2017*
  Design and implementation of a framework that combines logical and probabilistic reasoning for reducing the number of false alarms produced by a program analysis through user interactions [4, 11].

- *Searching Optimum Abstraction for Program Analyses.* *2011 - present*
  Design and implementation of a framework that finds program abstractions which optimally balance the tradeoff between efficiency and accuracy for Dataflow analyses [13, 15] and analyses expressed in Datalog [14].

- *A Solver for Large MaxSAT/Markov Logic Network Problems.* *2015-2017*
  Design and implementation of Nichrome, a solver for very large MaxSAT and Markov Logic Network instances that arise from program analysis and other domains by using iterative lazy solving [10, 11], query-guided solving [10], and Incremental solving [8].

Honors & Awards

2014	**ACM SIGPLAN Distinguished Paper Award,** ACM SIGPLAN conference on Programming Language Design and Implementation (PLDI)
2015	**ACM SIGSOFT Distinguished Paper Award,** joint meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on the Foundations of Software Engineering (ESEC/FSE)
2015	**Facebook Fellowship**
2017	**Outstanding Graduate Research Award,** College of Computing, Georgia Tech

Selected Publications

1. Osbert Bastani, **Xin Zhang**, Armando Solar-Lezama. Verifying Fairness Properties via Concentration. ACM Conference on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA), 2019.
2. **Xin Zhang**, Armando Solar-Lezama, Rishabh Singh. Interpreting Neural Network Judgments via Minimal, Stable, and Symbolic Corrections. Conference on Neural Information Processing Systems (NeurIPS), 2018.
3. **Xin Zhang**, Ravi Mangal, Mayur Naik, and Aditya Nori. Query-Guided Maximum Satisfiability. ACM Symposium on Principles of Programming Languages (POPL), 2016.
4. Ravi Mangal, **Xin Zhang**, Mayur Naik, and Aditya Nori. A User-Guided Approach to Program Analysis. ACM Symposium on Foundations of Software Engineering (FSE), 2015. **Distinguished Paper Award.**
5. **Xin Zhang**, Ravi Mangal, Radu Grigore, Mayur Naik, Hongseok Yang. On Abstraction Refinement for Program Analyses in Datalog. ACM Conference on Programming Language Design and Implementation (PLDI), 2014. **Distinguished Paper Award.**

Full Publication List

Conference Papers

1. Osbert Bastani, **Xin Zhang**, Armando Solar-Lezama. Verifying Fairness Properties via Concentration. ACM Conference on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA), 2019.
2. **Xin Zhang**, Armando Solar-Lezama, Rishabh Singh. Interpreting Neural Network Judgments via Minimal, Stable, and Symbolic Corrections. Conference on Neural Information Processing Systems (NeurIPS), 2018.
3. **Xin Zhang**, Radu Grigore, Xujie Si, Mayur Naik. Effective Interactive Resolution of Static Analysis Alarms. ACM Conference on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA), 2017.
4. Sulekha Kulkarni, Ravi Mangal, **Xin Zhang**, Mayur Naik. Accelerating Program Analyses by Cross-Program Training. ACM Conference on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA), 2016.
5. Xujie Si, **Xin Zhang**, Vasco Manquinho, Mikolas Janota, Alexey Ignatiev, Mayur Naik. On Incremental Core-Guided MaxSAT Solving. International Conference on Principles and Practice of Constraint Programming (CP), 2016.

6. Ravi Mangal, **Xin Zhang**, Aditya Kamath, Aditya Nori, and Mayur Naik. Scaling Relational Inference Using Proofs and Refutations. Conference on Artificial Intelligence (AAAI), 2016.
7. **Xin Zhang**, Ravi Mangal, Mayur Naik, and Aditya Nori. Query-Guided Maximum Satisfiability. ACM Symposium on Principles of Programming Languages (POPL), 2016.
8. Ravi Mangal, **Xin Zhang**, Aditya Nori and Mayur Naik. Volt: A Lazy Grounding Framework for Solving Very Large MaxSAT Instances. International Conference on Theory and Applications of Satisfiability Testing (SAT), 2015.
9. Jongse Park, Hadi Esmaeilzadeh, **Xin Zhang**, Mayur Naik, and Bill Harris. FlexJava: Language Support for Safe and Modular Approximate Programming. ACM Symposium on Foundations of Software Engineering (FSE), 2015.
10. Ravi Mangal, **Xin Zhang**, Mayur Naik, and Aditya Nori. A User-Guided Approach to Program Analysis. ACM Symposium on Foundations of Software Engineering (FSE), 2015. **Distinguished Paper Award.**
11. **Xin Zhang**, Ravi Mangal, Radu Grigore, Mayur Naik, Hongseok Yang. On Abstraction Refinement for Program Analyses in Datalog. ACM Conference on Programming Language Design and Implementation (PLDI), 2014. **Distinguished Paper Award.**
12. **Xin Zhang**, Ravi Mangal, Mayur Naik, Hongseok Yang. Hybrid Top-down and Bottom-up Interprocedural Analysis. ACM Conference on Programming Language Design and Implementation (PLDI), 2014.
13. **Xin Zhang**, Mayur Naik, Hongseok Yang. Finding Optimum Abstractions in Parametric Dataflow Analysis. ACM Conference on Programming Language Design and Implementation (PLDI), 2013.
14. Cheng Zhang, Juyuan Yang, Yi Zhang, Jing Fan, **Xin Zhang**, Jianjun Zhao, Peizhao Ou. Automatic Parameter Recommendation for Practical API Usage. International Conference on Software Engineering (ICSE), 2012.

Workshop Papers

1. Daniel Jackson, Jonathan DeCastro, Soonho Kong, Dimitrios Koutentakis, Angela Leong Feng Ping, Armando Solar-Lezama, Mike Wang, and **Xin Zhang**. Certified Control for Self-Driving Cars. Workshop on the Design and Analysis of Robust Systems (DARS), 2019.
2. **Xin Zhang**, Xujie Si, and Mayur Naik. Combining the Logical and the Probabilistic in Program Analysis. ACM SIGPLAN Workshop on Machine Learning and Programming Languages (MAPL), 2017.
3. Jongse Park, Kangqi Ni, **Xin Zhang**, Hadi Esmaeilzadeh, Mayur Naik. Expectation-Oriented Framework for Automating Approximate Programming. Workshop on Approximate Computing Across the System Stack (WACAS) in conjunction with ASPLOS, 2014.

Invited Tutorials

1. Mayur Naik, Xujie Si, **Xin Zhang**, Radu Grigore. Maximum Satisfiability in Program Analysis: Applications and Techniques. International Conference on Verification, Model Checking, and Abstract Interpretation (VMCAI), 2018.
2. Xujie Si, **Xin Zhang**, Radu Grigore, Mayur Naik. Maximum Satisfiability in Software Analysis: Applications and Techniques. International Conference on Computer Aided Verification (CAV), 2017.

Manuscripts

1. Zenna Tavares, **Xin Zhang**, Javier Burroni, Edgar Minasyan, Rajesh Ranganath, Armando Solar-Lezama. The Random Conditional Distribution for Higher-Order Probabilistic Inference. 2019.
2. Zenna Tavares, James Koppel, **Xin Zhang**, Armando Solar-Lezama. A Language for Counterfactual Generative Models. 2019.

Research Talks

Adaptive Program Analyses via Online Learning

Peking University *August 2019*

Program Analysis for Software 2.0

University of Pennsylvania — *May 2019*

Interpreting Neural Network Judgments via Minimal, Stable, and Symbolic Corrections

Peking University — *January 2019*
IBM Thomas J. Watson Research Center — *March 2018*

A User-Guided Approach to Program Analysis

IBM Programming Languages Day — *December 2016*
New Jersey Programming Languages and Systems Seminar — *September 2016*

Petablox: Declarative Program Analysis for Big Code

Google, Mountain View — *August 2016*
UC Berkeley — *August 2016*
Facebook Fellows Workshop — *July 2016*

Architectures and Systems for Mobile-Cloud Computing: A Workload-Driven Perspective

Qualcomm Innovation Fellowship Finalist Presentation — *March 2014*

Positions Held

Massachusetts Institute of Technology — *Cambridge, MA*
Postdoctoral Associate — *2017 - present*

University of Pennsylvania — *Philadelphia, PA*
Visiting Scholar — *2016 - 2017*

Microsoft Research Cambridge — *Cambridge, UK*
Research Intern — *2013*

Georgia Institute of Technology — *Georgia, GA*
Research Assistant — *2011-2017*

Teaching Experience

Kaufman Teaching Certificate Program (KTCP) — *MIT*
Trainee (Passed) — *Spring 2019*

6.820: Foundations of Program Analysis (Class Size: 30) — *MIT*
Guest Lecturer — *Fall 2017, Fall 2019*

CETL8713: Fundamentals of Teaching and Learning in Higher Education — *Georgia Tech*
Student — *Spring 2016*

CS6340: Software Analysis and Testing (Class Size: 30) — *Georgia Tech*
Teaching Assistant — *Fall 2014*

CS4400: Introduction to Database Systems (Class Size: 50) *Georgia Tech*
Teaching Assistant *Spring 2013*

Mentorship

Zenna Tavares and James Koppel, Ph.D. students, MIT.

- The Random Conditional Distribution for Higher-Order Probabilistic Inference (In Submission).
- A Language for Counterfactual Generative Models (In Submission).

Yifan Chen, Ph.D. student, Peking University. Learning Equivalence Class Abstraction for Accelerating Datalog Analysis (In Progress).

Xujie Si, Ph.D. student, University of Pennsylvania. On Incremental Core-Guided MaxSAT Solving, CP 2016.

Sulekha Kulkarni, Ph.D. student, University of Pennsylvania. Accelerating Program Analyses by Cross-Program Training, OOPSLA 2016.

Aditya Kamath, M.S. student, Georgia Tech. Scaling Relational Inference Using Proofs and Refutations, AAAI 2016.

References

Armando Solar-Lezama
Associate Professor
Computer Science & Artificial Intelligence Laboratory
MIT
asolar@csail.mit.edu
+1 617 528 9727

Mayur Naik
Associate Professor
Department of Computer and Information Science
University of Pennsylvania
mhnaik@cis.upenn.edu
+1 215 573 1856

Rajeev Alur
Zisman Family Professor
Department of Computer and Information Science
University of Pennsylvania
alur@cis.upenn.edu
+1 215 573 7483

Hongseok Yang
Professor
School of Computing
KAIST
hongseok00@gmail.com
+82 42 350 3573

Isil Dillig
Associate Professor
Department of Computer Science
University of Texas, Austin
isil@cs.utexas.edu
+1 512 471 9794