

# Xujie Si

Bahen 4242, 40 St George St

Toronto, Ontario M5S 2E4

☎ (+1) 437-217-8940

✉ [six@cs.toronto.edu](mailto:six@cs.toronto.edu)

🌐 <https://www.cs.toronto.edu/~six/>

## Research Interests

My research interests span programming languages, formal methods, and artificial intelligence. Specifically, my research focuses on developing machine learning techniques to address program reasoning challenges such as program verification, program synthesis, and software testing. My research also concerns neuro-symbolic techniques to improve learning efficiency and robustness, and enable logic reasoning on perception tasks.

## 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*

## Employment

- 2023 Jan. – **Assistant Professor, University of Toronto.**
- 2021 Jan. – **Canada CIFAR AI Chair, Mila – Quebec AI Institute.**
- 2021 – 2022 **Assistant Professor, McGill University.**
- 2019 **Research Scientist Intern, DeepMind, London.**
- 2015 **Software Engineer Intern, Google, Mountain View.**
- 2014 **Software Engineer Intern, Google, Seattle.**

## Publications

- NeurIPS'21 Sever Topan, David Rolnick, Xujie Si. **Techniques for Symbol Grounding with SATNet.** In *Proceedings of the thirty-fifth Conference on Neural Information Processing Systems, NeurIPS 2021 (Spotlight)*.
- NeurIPS'21 Jiani Huang, Ziyang Li, Binghong Chen, Karan Samel, Mayur Naik, Le Song and Xujie Si. **Scallop: From Probabilistic Deductive Databases to Scalable Differentiable Reasoning.** In *Proceedings of the thirty-fifth Conference on Neural Information Processing Systems, NeurIPS 2021*.
- FMCAD'21 Nham Le, Xujie Si, Arie Gurfinkel. **Data-driven Optimization of Inductive Generalization.** In *Proceedings of the 21st International Conference on Formal Methods in Computer-Aided Design*.
- SIGCSE'21 Alana Ceci, Hanneli C. A. Tavante, Brigitte Pientka, Xujie Si. **Data Collection for the Learn-OCaml Programming Platform: Modelling How Students Develop Typed Functional Programs.** In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education, SIGCSE 2021*.
- 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.

## Supervision

### PhD students at McGill & Mila

- 2021/09 - Allen Geng, Zhaoyu Li, Ray Luo
- 2021/06 - Breandan Considine (co-supervised with Prof. Jin Guo)

### Masters students at McGill

- 2021/09 - Xiaojie Xu, Mingchuan Zhang
- 2020/09 - Sever Topan

### Undergraduate students at McGill

2021/09 - Hongye Chen, Owen Smith, Yingjie Xu, Hanwen Zhu  
2021/05 - Yue Wu  
2021/01 - Adam Weiss  
2021/01 - 05 Zhenghua Chen, Kangrui Ren

---

## Teaching

Fall 2021 [Instructor] COMP 597: Automated Reasoning with Machine Learning  
Winter 2021 [Instructor] COMP 302: Programming Languages and Paradigms  
Summer 2019 [TA] CIS 547 Software Analysis (Online Course) , *University of Pennsylvania*  
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

### **Program Reasoning: An Interesting AI Challenge**

Sep. 2021 Mila - Quebec AI Institute, Montreal, QC, Canada  
Oct. 2021 Aliibaba Damo Academy, Virtual  
Nov. 2021 Nankai University, Virtual

### **End-to-end Learning for Program Verification.**

Sep. 2020 Forum on specification & Design Languages (FDL 2020), Virtual  
**Code2Inv: A Deep Learning Framework for Program Verification**  
July 2020 International Conference on Computer-Aided Verification, Virtual  
**Learning-aided Program Reasoning.**  
Jan. 2020 University of Waterloo, Waterloo, ON, Canada  
Feb. 2020 McGill University, Montreal, QC, Canada  
Mar. 2020 Swiss Federal Institute of Technology Lausanne (EPFL), Virtual  
Mar. 2020 National University of Singapore(NUS), Virtual

### **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

- 2021 Canada CIFAR AI Chair
- 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

---

## Professional Activities

- CAV'22 Program Committee
- PLDI'22 Program Committee
- ICLR'21,'22 Reviewer
- AAAI'21,'22 Reviewer
- NeurIPS'20,'21 Reviewer
- APLAS'21 Program Committee
- MAPL'20 Program Committee
- CAV'19,'20 Artifact Evaluation Committee
- ICFP'19 Artifact Evaluation Committee
- PLDI'18,'19 Student Volunteer Co-Chair
- POPL'16 Student Volunteer