

Xujie Si

Bahen 4242, 40 St George St

Toronto, Ontario M5S 2E4

☎ (+1) 437-217-8940

✉ 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 various reasoning challenges such as auto-(in)formalization, theorem proving, program verification, and program synthesis. My research also concerns neuro-symbolic techniques to improve learning efficiency and robustness.

Education

2016/08– **Ph.D., Computer and Information Science.**

2020/08 *University of Pennsylvania*, Philadelphia, PA Advisor: Prof. Mayur Naik

2011/08– **M.S., Electrical Engineering and Computer Science.**

2014/08 *Vanderbilt University*, Nashville, TN Advisor: Prof. Yuan Xue

2007/09 – **B.E. (with Honors), Software Engineering.**

2011/06 *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*.

2019/06–10 **Research Scientist Intern**, *DeepMind*, London.

Selected Publications

- NeurIPS'24 Bowen Li, Zhaoyu Li, Qiwei Du, Jinqi Luo, Wenshan Wang, Yaqi Xie, Simon Stepputtis, Chen Wang, Katia P. Sycara, Pradeep Kumar Ravikumar, Alexander G. Gray, **Xujie Si**, Sebastian Scherer. **LogiCity: Advancing Neuro-Symbolic AI with Abstract Urban Simulation**. In *Proceedings of the thirty-seven Conference on Neural Information Processing Systems*, 2024.
- NeurIPS'24 Hao Tang, Keya Hu, Jin Peng Zhou, Sicheng Zhong, Wei-Long Zheng, **Xujie Si**, Kevin Ellis. **Code Repair with LLMs gives an Exploration-Exploitation Tradeoff**. In *Proceedings of the thirty-seven Conference on Neural Information Processing Systems*, 2024.
- ICML'24 Logan Murphy, Kaiyu Yang, Jialiang Sun, Zhaoyu Li, Anima Anandkumar, **Xujie Si**. **Autoformalizing Euclidean Geometry**. In *Proceedings of the 41st International Conference on Machine Learning*, 2024.
- COLM'24 Zhaoyu Li, Jialiang Sun, Logan Murphy, Qidong Su, Zenan Li, Xian Zhang, Kaiyu Yang, **Xujie Si**. **A Survey on Deep Learning for Theorem Proving** In *Proceedings of the First Conference on Language Modeling*, 2024.
- FMCAD'24 Max Kopinsky, Brigitte Pientka and **Xujie Si**. **Modernizing SMT-Based Type Error Localization** In *Proceedings of the 24th International Conference on Formal Methods in Computer-Aided Design*
- NeurIPS'23 Zhaoyu Li, Jinpei Guo, Yuhe Jiang, **Xujie Si**. **Learning Reliable Logical Rules with SATNet**. In *Proceedings of the thirty-six Conference on Neural Information Processing Systems*, 2023

- ICML'23 Chuqin Geng, Nham Le, Xiaojie Xu, Zhaoyue Wang, Arie Gurfinkel, **Xujie Si**. **Towards Reliable Neural Specifications**. In *Proceedings of the 40th International Conference on Machine Learning*, 2023 (**Oral**).
- SIGCSE'23 Chuqin Geng, Wenwen Xu, Yingjie Xu, Brigitte Pientka, **Xujie Si**. **Identifying different student clusters in functional programming assignments: From quick learners to struggling students**. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education*, 2023.
- NeurIPS'22 Zhaoyu Li, **Xujie Si**. **NSNet: A General Neural Probabilistic Framework for Satisfiability Problems**. In *Proceedings of the thirty-six Conference on Neural Information Processing Systems*, 2022.
- 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*.
- 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**).

Supervision

PhD students at University of Toronto

- 2023/01 - Zhaoyu Li
- 2024/06 - Honghua Dong
- 2024/12 - Hangrui Bi

Masters students at University of Toronto

- 2023/09 - Sissi Jiang, Steven Zhong

Undergraduate students at University of Toronto

- 2024/05 - Sarah Walker, Scott Cui
- 2023 - 2024 Shujie Deng, Jingyu Hu, Sophia Liao

PhD students at McGill & Mila

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

Masters students at McGill

- 2022 - 2024 Rebecca Wang (co-supervised with Prof. Clark Verbrugge)
- 2021 - 2023 Xiaojie Xu
- 2020 - 2022 Sever Topan

Undergraduate students at McGill

- 2021/09 - 12 Hongye Chen, Owen Smith, Yingjie Xu, Hanwen Zhu
- 2021/05 - 12 Yue Wu

2021/01 - 12 Adam Weiss
2021/01 - 05 Zhenghua Chen, Kangrui Ren

Research Grants

2023-2027 **Start-up fund**, *University of Toronto*, \$500,000.
2021-2025 **Learning-aided Program Reasoning**, *NSERC Discovery*, \$135,000.
2021-2025 **Intelligent Reasoning System by Combining Machine Learning and Formal Methods**, *Canada CIFAR AI Chairs Program*, \$1,048,500, (*restricted to Mila graduate students*).
2022-2023 **Machine learning for software repair**, *FPT Software (Mila Industrial Partner)*, \$40,500, (*restricted to Mila graduate students*).

Teaching

Fall'24 CSC 2108: Automated Reasoning with Machine Learning, *University of Toronto*
Fall'24 CSC 324: Principles of Programming Languages, *University of Toronto*
Winter'24 CSC 324: Principles of Programming Languages, *University of Toronto*
Winter'23,24 Topics in Machine Learning: Automated Reasoning with Machine Learning, *University of Toronto*
Fall'21,22 COMP 597: Automated Reasoning with Machine Learning, *McGill University*
Winter'21,22 COMP 302: Programming Languages and Paradigms, *McGill University*

Selected Awards

2021 Canada CIFAR AI Chair
2019 PLDI Distinguished Paper Award
2019 Facebook Fellowship and Emerging Scholars finalist
2016 CSAW Best Applied Research Paper finalist

Professional Activities

Area Chair ICLR'25
Reviewer NeurIPS'20-24; ICML'22-24; ICLR'21-24; AAAI'21-23; CVPR'23; ICCV'23
CAV'24 Local Chair and Program Committee
TACAS'24 Program Committee
SPLASH'23 Program Co-Chair of Student Research Competitions
SPLASH'22 Program Co-Chair of Student Research Competitions
CAV'22 Program Committee
PLDI'22 Program Committee
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