Chuqin (Allen) Geng

Updated November 16, 2026

Email: chuqin.geng@mail.mcgill.ca
Address: 9 Delphinium Ave, Richmond Hill, ON L4E 4V1
Phone: (514) 652-1880
Website: allengeng123.github.io

Research interests

Transparent and reliable AI systems: verifying, explaining, and steering model behavior. Neuro-symbolic AI, with applications in AI for science and drug discovery. Mechanistic interpretability of LLMs and AI agents, with the goal of understanding their internal mechanisms and improving alignment and safety.

Education

University of Toronto — Toronto, ON, CA
Visiting Researcher — *Jan* 2023 – Present
Advisor: Prof. Xujie Si

McGill University, Mila - Quebec AI Institute — Montreal, QC, CA
PhD in Computer Science — *Sep* 2021 – Present
Advisor: Prof. Xujie Si, *GPA: 4.0/4.0*

Georgia Institute of Technology — Atlanta, GA, USA
MSc in Computer Science, *GPA: 4.0/4.0* — *Sep* 2019 – *Apr* 2021

University of Reading — Reading, Berkshire, UK
MSc in Finance, *GPA: 4.0/4.0* — *Sep* 2016 – *Dec* 2017

University of Toronto — Toronto, ON, CA
Honours BSc in Math and Statistics — *Sep* 2011 – *Aug* 2015

Selected Papers

Are We Recovering Mechanisms? Objective-Level Recovery Gaps in Mechanistic Interpretability
Chuqin Geng, Li Zhang, Haolin Ye, Mark Zhang, Luke Zhang, Xujie Si
In submission

Rethinking Circuit Evaluation: Do Circuits Explain Model Errors?
Li Zhang*, **Chuqin Geng***, Mark Zhang, Chen Yang, Luke Zhang, Haolin Ye, Xujie Si
In submission

VisionLogic: Discovering and Grounding Decision-Relevant Visual Concepts
Chuqin Geng, Yuhe Jiang, Li Zhang, Zhaoyue Wang, Haolin Ye, Mark Zhang, Jingkai Xu, Xujie Si
In submission

Beyond Message Passing: A Symbolic Alternative for Expressive and Interpretable Graph Learning
Chuqin Geng, Li Zhang, Haolin Ye, Ziyu Zhao, Yuhe Jiang, Tara Saba, Xinyu Wang, Xujie Si
In submission

CROSSREF-VQA: Benchmarking Cross-modal Reference Resolution in Structured Graphics
Yuhe Jiang, Youyuan Zhang, Hangrui Bi, **Chuqin Geng**, Xujie Si
In submission

Neural Proposals, Symbolic Guarantees: Neuro-Symbolic Graph Generative Modeling
Chuqin Geng, Li Zhang, Mark Zhang, Zhaoyue Wang, Haolin Ye, Xujie Si
NeurIPS 2026

DreamProver: Evolving Transferable Lemma Libraries via a Wake-Sleep Theorem-Proving Agent
Youyuan Zhang*, Jialiang Sun*, Hangrui Bi, **Chuqin Geng**, Wenjie Ma, Zhaoyu Li, Xujie Si
COLM 2026

LogicXGNN: Grounded Logical Rules for Explaining Graph Neural Networks
Chuqin Geng, Ziyu Zhao, Zhaoyue Wang, Haolin Ye, Yuhe Jiang, Xujie Si
ICLR 2026 ***(Top ~2%)***

NEUROLOGIC: From Neural Representations to Interpretable Logic Rules
Chuqin Geng, Anqi Xing, Li Zhang, Ziyu Zhao, Yuhe Jiang, Xujie Si
UCRL @ ICLR 2026

Learning Minimal Neural Specifications
Chuqin Geng, Zhaoyue Wang, Haolin Ye, Xujie Si
NeuS 2025 ***(Oral)***

Towards Robust Saliency Maps
Van Nham Le*, **Chuqin Geng***, Xujie Si, Arie Gurfinkel
ACML 2024

Towards Reliable Neural Specifications
Chuqin Geng*, Van Nham Le*, Xiaojie Xu, Zhaoyue Wang, Arie Gurfinkel, Xujie Si
ICML 2023 ***(Oral)***

Identifying Different Student Clusters in Functional Programming Assignments: From Quick Learners to Struggling Students
Chuqin Geng, Wenwen Xu, Yingjie Xu, Brigitte Pientka, Xujie Si
SIGCSE 2023 TS

TorchProbe: Fuzzing Dynamic Deep Learning Compilers
Qidong Su, **Chuqin Geng**, Gennady Pekhimenko, Xujie Si
APLAS 2023

Scalar Invariant Networks with Zero Bias
Chuqin Geng, Xiaojie Xu, Haolin Ye, Xujie Si
NeuReps @ NeurIPS 2023

Novice Type Error Diagnosis with Natural Language Models
Chuqin Geng, Haolin Ye, Yixuan Li, Tianyu Han, Brigitte Pientka, Xujie Si
APLAS 2022

SAT-DIFF: A Tree Diffing Framework Using SAT Solving
Chuqin Geng, Haolin Ye, Yihan Zhang, Brigitte Pientka, Xujie Si
preprint, arxiv.org/abs/2404.04731

Can ChatGPT Pass An Introductory Level Functional Language Programming Course?
Chuqin Geng, Yihan Zhang, Brigitte Pientka, Xujie Si
preprint, arxiv.org/abs/2305.02230

Selected awards

Top Reviewer (NeurIPS)	2026
SOCS Grad Stimulus Initiative (McGill University)	2025
Mila Research Scholarship (Mila - Quebec AI Institute)	2023–2025
Graduate Excellence Award (McGill University)	2021–2025
Faculty of Science Grad Supplement (McGill University)	2023
Apple Scholars nomination (McGill University)	2023
Computer Science Top-up Award (McGill University)	2022
Academic Distinction Award (University of Reading)	2017
Second prize in mathematics competition (Shaanxi province, China)	2009

Research Experience

Explainable AI, Mechanistic Interpretability, and AI for Science
Advisor: Prof. Xujie Si (University of Toronto) *Jan* 2025 – Present
Developed grounded logical explanations for graph neural networks, revealing molecular substructures associated with their predictions. Combined neural graph generation with symbolic solving to generate molecules satisfying hard constraints, enabling controllable design for drug and materials discovery. Identified limitations in current mechanistic interpretability methods for large language models.

Neural Network Verification, Robustness, and AI Safety
Advisor: Prof. Xujie Si (University of Toronto) *Sep* 2022 – *May* 2025
Proposed neural activation patterns as verification specifications, extending certifiable input regions beyond hand-chosen local neighborhoods. Developed minimal neural specifications for concise, interpretable verification. Explored methods to improve model robustness, fairness, and safety.

Computer science education research *Sep* 2021 – *Dec* 2022
Advisors: Profs. Xujie Si and Brigitte Pientka (McGill University)
Studied transformer-based language models for diagnosing type errors made by novice programmers and conducted education research using statistical analysis of student activity logs, with the goal of auditing and improving the student learning experience.

Multimodal clinical prediction and reinforcement learning
Graduate research (Georgia Institute of Technology) *Sep* 2019 – *Apr* 2021
Conducted research projects on multimodal mortality prediction using MIMIC clinical data and reinforcement learning for robot control.

Disproof of a conjecture in biometric security optimization
Mentors: Prof. Steven Rayan (University of Toronto) *Jan* 2015 – *Aug* 2015
Disproved a conjecture regarding optimal solutions for biometric privacy-security trade-offs, providing new bounds for secure system optimization.

Teaching experience

Head teaching assistant, McGill University Winter 2022
COMP 302: Programming Languages and Paradigms
Conducted weekly office hours and tutorials, designed and graded exams, developed auto-graders for assignments, implemented mutation testing, and utilized Moss for plagiarism detection.

Industry experience

Cirqvia (startup project)
Founder *May* 2026 – Present
Translating mechanistic interpretability research to address the practical need to understand model behavior in high-stakes domains and improve the alignment, safety, and reliability of LLMs and AI agents. Leading research and product development; received early-stage funding.

FITFI Inc. Toronto, CA
Senior Data Scientist *Jan* 2018 – *Sep* 2019
Developed deep learning methods for automatic motion and exercise detection through modeling of multimodal biometric data, supporting applications in AI for health. Led the algorithm team, secured a demo opportunity at the 2019 Collision Conference, and helped raise over CAD 2 million in funding.

SHAREWIN SOFTWARE Beijing, China
Algorithm Engineer *Sep* 2015 – *Sep* 2016
Developed seismic-interpretation algorithms for oil and gas exploration software, using preprocessing, filtering, and ant-tracking methods to delineate geological faults in 3D seismic volumes.

Talks

Learning Minimal Neural Specifications *May* 2025
NeuS 2025

Towards Reliable Neural Specifications
ICML 2023 *Jul 2023*

A study on student performance clusters *Mar* 2023
SIGCSE 2023, SPLICE Workshop

Novice Type Error Diagnosis with Natural Language Models *Dec* 2022
APLAS 2022

Service

I contribute to academic service through conference reviewing and student selection:

- **Machine Learning:** ICML (2023–2026), ICLR (2024–2027), NeurIPS (2023–2026), AAAI (2024–2027)
- **Computer Vision:** CVPR (2024, 2025), ECCV 2026
- **Formal Methods:** TACAS 2025
- **HCI & Education:** SIGCSE 2023, CHI 2023
- **Supervision Evaluation Committee Member:** Mila Supervision Request Process 2027

Mentoring

I have mentored students from the following institutions:

- **University of Toronto:** Li Zhang, Mark Zhang, Eric Liang, Luke Zhang
- **McGill University:** Ziyu Zhao, Haolin Ye, Yihan Zhang, Zhaoyue Wang, Xiaojie Xu, Yixuan Li, Tianyu Han, Yingjie Xu
- **McMaster University:** Jingkai Xu
- **Tsinghua University:** Chen Yang