Yuan Si

samsi0322.github.io | github.com/SamSi0322 | Google Scholar | y3si@uwaterloo.ca | +1(437)777-5619

Research Interests

AI for software engineering (AI4SE), LLM-based program repair, multimodal debugging, AI for computing education; analytic number theory, probability theory.

Publications

- **Yuan Si**, Daming Li, Hanyuan Shi, Jialu Zhang. *ViScratch: Using Large Language Models and Gameplay Videos for Automated Feedback in Scratch*. ACM International Conference on the Foundations of Software Engineering (FSE) 2026 (accepted, open review).
- **Yuan Si**, Simeng Han, Daming Li, Hanyuan Shi, Jialu Zhang. *ScratchEval: A Multimodal Evaluation Framework for LLMs in Block-Based Programming*., ACM International Symposium on Software Testing and Analysis (ISSTA) 2026 (accepted).
- **Yuan Si**, Ming Wang, Daming Li, Hanyuan Shi, Jialu Zhang. *EcoScratch: Cost-Effective Multimodal Repair for Scratch Using Execution Feedback*., 2026.
- **Yuan Si**, Kyle Qi, Daming Li, Hanyuan Shi, Jialu Zhang. *Stitch: Step-by-step LLM Guided Tutoring for Scratch*., 2025.
- **Yuan Si**. *A Microcanonical Phase Transition for the Collatz Affine Random Model* (2026). Submitted to *Forum of Mathematics, Sigma*.
- **Yuan Si**. *Mixed Parity, Diagonal Denominator, and the Pell-Chord Genus-Five Obstruction for the Four-Corner Rational Distance Problem* (2026). Submitted to *J. Number Theory*.
- **Yuan Si**. *Elliptic Decomposition of the Pell-Chord Genus-Five Obstruction for the Four-Corner Rational Distance Problem* (2026).
- **Yuan Si**. *A Gentle Introduction to Optimization* (2024). Adopted as required reading in 3 university courses.

Experience

- **University Researcher - AI in Multidimensional Input** — Remote, Waterloo, Canada
  UWaterloo, ECE Department — *July 2025 – present*
  - Led 4 first-authored papers on LLM-driven debugging, tutoring, and evaluation for Scratch (ViScratch accepted at FSE 2026, ScratchEval at ISSTA 2026).
  - Built multimodal repair systems fusing gameplay video, execution feedback, and project JSON to localize bugs and synthesize fixes via LLM-guided loops.
  - Developed an interactive tutoring system (Stitch) and a 100-project executable benchmark (ScratchEval) for evaluating LLM repair quality on block-based code.

- **University Researcher - Game Theory & Graph Theory** — On-site, Waterloo, Canada
  University of Waterloo, Combinatorics & Optimization Department — *Dec. 2024 – March 2025*
  - Derived Nash equilibria for multiplayer RPS via LP; modeled Tesla charging placement via independent dominating sets; analyzed Public Goods Games.

- **Microsoft Researcher - AI Trends** — Remote, Vancouver, Canada
  Microsoft - Microsoft Azure & AI, Research — *May 2024 – Aug. 2024*
  - Investigated AI scribe technologies for healthcare and improved an insurance chatbot via systematic A/B analysis of user interaction data.

Education

University of Waterloo — Waterloo, Canada
HBM, Math / Data Science & Combinatorics Optimization & Pure Math (co-op) — *Sept. 2023 – April 2028*

Projects

- **ResearchOS**: Local-first research decision and execution system; turns a brief into runnable experiments, evidence-checked claims, and a packaged manuscript draft. FastAPI + React, provider-agnostic LLM adapters (OpenAI / Anthropic / mock), dual-agent code worker, HITL approval gates, encrypted secret store. AGPL-3.0.
- **DevToolkit**: 43 zero-dependency single-file Python CLI tools across 9 categories (web, data, process, MCP, security, scaffolding).

Skills

- **Research**: Experiment design, human evaluation, ablation studies, benchmark construction, statistical analysis
- **AI/LLM**: Multimodal repair pipelines, LLM evaluation, prompt engineering, MCP tooling
- **Languages & Tools**: Python, JavaScript/TypeScript, C/C++, Java, Git, Linux, LaTeX, Magma, SageMath