

XIYU ZHAI

249 Thrid Street, Cambridge, MA 02142

(+1)6178003895 \diamond xiyuzhai@mit.edu

EDUCATION

Massachusetts Institute of Technology

August 2018 - Present

PhD Candidate, EECS Department

Major: Programming Language for AI

University of Cambridge

October 2017 - June 2018

Master of Advanced Studies in Part III Mathematics

Major: Pure Mathematics and Statistics

Grades: Distinction

University of Science and Technology of China

August 2013 - July 2017

Bachelor of Science

Major: Pure Mathematics

Department of Mathematical Sciences

GPA: 4.11/4.3, Rank: #1

RESEARCH

- Programming Language for AI.

I'm implementing an programming language (<https://github.com/ancient-software/husky>) that enables advanced learning-augmented programming across multiple domains, including pattern recognition in computer vision, learning-based rendering, video compression, RL, theorem proving, etc.

The immediate goal now is to make a blazingly fast, fully interpretable, provably robust image classifiers (which shall be different from deep learning in every possible way).

Have started pursuing this goal even before PhD. It started theoretical, but it has evolved into a programming language project.

Expect to finish the first round of results in one or two months.

- Number Theory.

I'm simultaneously interested in Number Theory itself and rewriting existing proofs into machine checkable ones, just like Xena project.

Still learning the basics.

PROGRAMMING TECHNIQUE

- Rust. My major language. I'm use Rust mostly in my PhD project, 60k+ lines of code currently.
- Lean4. I just begin to write specs in Lean4. Could be writing this a lot.
- C++. Have experience writing and managing a codebase of 30k+ lines of code. But I know the language well enough to prefer Rust.
- Javascript/Typescript. I use it for gui development. I had experience with React but now I'm using Svelte.
- Python. Have experience with it during internship and past immature attempts of my PhD project.

- Haskell. Didn't programmed much in it, but immitate many of its features in my project.

PUBLICATION

I've not yet finish anything publishable in my PhD project. So the following doesn't reflect much about what I'm doing right now.

1. Wenlong Mou, Liwei Wang, Xiyu Zhai, and Kai Zheng. Generalization bounds of sgld for non-convex learning: Two theoretical viewpoints. *COLT 2018*
2. Alexander Rakhlin and Xiyu Zhai. Consistency of interpolation with laplace kernels is a high-dimensional phenomenon. *COLT 2019*
3. Simon S. Du, Xiyu Zhai, Barnabás Póczos, and Aarti Singh. Gradient descent provably optimizes over-parameterized neural networks. *ICLR 2018*
4. Simon S. Du, Yining Wang, Xiyu Zhai, Sivaraman Balakrishnan, Ruslan R. Salakhutdinov, and Aarti Singh. How many samples are needed to learn a convolutional neural network? *NeurIPS 2018*
5. Simon S. Du, Jason D. Lee, Haochuan Li, Liwei Wang, and Xiyu Zhai. Gradient descent finds global minima of deep neural networks. *ICML 2019*
6. Tengyuan Liang, Alexander Rakhlin, and Xiyu Zhai. On the risk of minimum-norm interpolants and restricted lower isometry of kernels. *arXiv19*