Yuhao Zhang

(608) 236-3965 · www.linkedin.com/in/yuhaoz · https://github.com/ForeverZyh

EDUCATION

University of Wisconsin-Madison Aug 2019 - Present
PhD Student in Computer Science, GPA: 3.96 Madison, WI

Peking University Sept 2015 - Jul 2019
B.S. in Computer Science and Technology, Summa Cum Laude, Outstanding Undergraduate Student Beijing, CN

PUBLICATIONS

Yuhao Zhang, Aws Albarghouthi, Loris D'Antoni, "BagFlip: A Certified Defense against Data Poisoning" in *Thirty-sixth Conference on Neural Information Processing Systems (**Neurips** 2022)*, New Orleans, the United States

Yuhao Zhang*, Yasharth Bajpai*, Priyanshu Gupta*, Ameya Ketkar*, Miltiadis Allamanis, Titus Barik, Sumit Gulwani, Arjun Radhakrishna, Mohammad Raza, Gustavo Soares, and Ashish Tiwari, "Overwatch: Learning Patterns in Code Edit Sequences" in *Object-Oriented Programming, Systems, Languages & Applications (**OOPSLA** 2022)*, Auckland, New Zealand

Yuhao Zhang, Aws Albarghouthi, Loris D'Antoni, "Certified Robustness to Programmable Transformations in LSTMs" in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (**Oral Presentation EMNLP** 2021)*, Online and Punta Cana, Dominican Republic

Yuhao Zhang, Luyao Ren, Liqian Chen, Yingfei Xiong, Shing-Chi Cheung, Tao Xie, "Detecting Numerical Bugs in Neural Network Architectures" in *Proceedings of the ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (**ACM Distinguished Paper Award ESEC/FSE** 2020)*, Online, the United States.

Yuhao Zhang, Aws Albarghouthi, Loris D'Antoni, "Robustness to Programmable String Transformations via Augmented Abstract Training" in *Proceedings of the Thirty-seventh International Conference on Machine Learning (**ICML** 2020)*, Online, Austria

Yuhao Zhang, Yifan Chen, Shing-Chi Cheung, Yingfei Xiong, and Lu Zhang, "An Empirical Study on TensorFlow Program Bugs" in *Proceedings of the 27th ACM SIGSOFT International Symposium on Software Testing and Analysis (**ISSTA** 2018)*, Amsterdam, Netherlands

SKILLS

Languages	C++, C, C#, Python, OCaml, Rust, Scheme, Java, JavaScript, HTML, CSS, PHP, Verilog
Technologies	Git, TensorFlow, Keras, Pytorch, LLVM, MySQL

PROFESSIONAL EXPERIENCE

Microsoft May 2021 - Aug 2021 & Feb 2022 - April 2022
Research Intern - PROSE (Programming by Examples and Natural Language) Team [C#/Python] Remote

· Worked on the Blue-Pencil project, one of the main projects behind **Visual Studio IntelliCode**.
· Designed the representation for storing thousands of source code editing patterns from developer traces.
· Implemented the framework that learns and applies hundreds of patterns and achieved 78% precision.
· Observed a 4X increase in the number of users using the tool because of the framework.

Microsoft Research Sept 2018 - March 2019
Research Intern - DKI (Data, Knowledge, Intelligence) Group [C#/Python] Beijing, CN

· Worked on **Ideas**, a plugin in Excel, which analyzes and provides high-level visual summaries for data analysts.
· Improved the classification accuracy of the intermediary model from 88% to 93% for six primary languages.
· Accelerated 4X column headers' matching speed with target phrases by implementing the Aho–Corasick algorithm.
· Tuned the hyperparameters of models by implementing a grid search algorithm, which is used by other groups.
· Won the **Award of Excellence** during the internship.

PROJECTS

Forward-mode Automatic Differentiation (AD) for Angora Fuzzer Feb 2020 - May 2020
Forward-mode AD computes more precise partial derivatives than the counterpart in Angora [C++/LLVM/Rust]

· Implemented an `Int` class to compute partial derivatives while keeping the original semantics of primitive int types.
· Instrumented the intermediate representation to surrogate the primitive int types with the `Int` class using LLVM.
· Registered new trace functions in compiled binary for communicating with Angora Fuzzer by proxy calls.

Course Scheduling System for Peking University Sept 2017 - May 2018
Course schedules generated by our system outperformed the dean's design on three metrics [Python/C++]

· Cleaned data and mined rules in raw data provided by the dean containing 529 majors in 39 departments.
· Designed a simulated annealing algorithm to solve thousands classroom conflicts and smooth the course density.
· Pipelined the components: data preprocessing, tabu search, simulated annealing, and generating course schedules.

PKURUNNER Application Apr 2016 - Jun 2018
The Android application is used by more than 2000 students for recording their running traces [Java]

· Implemented the GUI showing the map, the current location, the running trace, and metrics like running speed.
· Designed and implemented the logics interacting with the users to start, pause, and stop running.
· Invoked Gaode Maps APIs to get the GPS locations of the user and packed the trace for uploading to the server.

HONORS AND AWARDS

Research

· ACM Distinguished Paper Award at ESEC/FSE 2020

ACM-ICPC

· Ho-Chi-Minh City Regional **4th place**, 2017
· Xi'an Regional **Gold**, 2017
· Yangon Regional **7th place**, 2016
· Dalian Regional **Gold**, 2016
· Hefei Regional **Gold** 2015

Scholarships

· SenseTime Scholarship 2019
· Suzhou Industrial Park Scholarship 2018
· Schlumberger Scholarship 2017
· iPinYou Scholarship 2016

SERVICE

FoMLAS 2022, 2021, Program Committee
ICML 2022, 2021, Reviewer
Neurips 2022, Reviewer
VMCAI 2023, Artifact Evaluation Committee
CAV 2021, Artifact Evaluation Committee