Shuyang Liu

🌐 shuyang-liu.github.io   ✉ liushuyang860@gmail.com

EDUCATION

Huazhong University of Science and Technology   **GPA: 3.92/4.00**
B.E. in Computer Science   *Sept. 2020 - Jun. 2024*

PUBLICATION

Symmetry-Preserving Program Representations for Learning Code Semantics
Kexin Pei, Weichen Li*, Qirui Jin*, **Shuyang Liu**, Scott Geng, Lorenzo Cavallaro, Junfeng Yang, Suman Jana
Accepted by Symposium on Machine Programming (MAPS) workshop at ESEC/FSE 2023
Submitted to International Conference on Learning Representations (ICLR), 2024. Under Review.

RESEARCH EXPERIENCE

Exploiting Code Symmetries for Learning Program Semantics   **Advisor: Suman Jana**
Columbia University   *Mar. 2023 - Aug. 2023*

- Employed Tree-sitter to construct Program Dependence Graphs (PDGs) based on data and control dependencies between statements.
- Implemented nine types of semantic-preserving source transformations beyond PDG automorphisms.
- Evaluated a range of traditional and LLM baseline models for method name prediction.

Enhancing Code Semantics Learning with Fine-Grained PDGs   **Advisor: Kexin Pei**
University of Chicago   *Sept. 2023 - Present*

Extended our existing work from inter-statement analysis to intra-statement analysis, constructing fine-grained Program Interpretation Graphs for LLMs to capture code semantics.

- Static typed language: Employed JavaParser to extract dependencies between tokens for Java.
- Dynamic typed language: Applied Pytype to perform static type inference for Python.

Automatic Identification of Bug Inducing Commits   **Advisor: Ming Wen**
CASTLE Lab, Hong Kong University of Science and Technology   *Ongoing*

- Systematically validated bug-fixing and associated bug-inducing commits for 237 bugs across five large open-source Java projects via bisect testing on version histories.
- Reproduced SZZ Unleashed algorithm on Defects4J database.
- Improved BIC candidate detection by integrating Failure Coverage analysis with SZZ Unleashed.

EXCHANGE PROGRAM

Real-time Traffic Sign Recognition with Adversarial Training   **Advisor: Terence Sim**
School of Computing, National University of Singapore   *May. 2022 - Aug. 2022*

- Role: Student Team **Leader**
- Constructed and tuned a CNN; Implemented a suite of innovative data augmentation techniques to improve adversarial robustness and generalize to diverse scenarios.

TECHNICAL SKILLS

- **Languages**: C/C++, Java, Python (Pytorch, Tensorflow)
- **Tools**: Git, Linux, LaTeX, Soufflé, CodeQL, Tree-sitter, and JavaParser
- **Areas of Interest**: Software Engineering, Security, Programming Languages, and Machine Learning