Curriculum Vitae

Jifeng Wu
Master of Science in Computer Science @ UBC
Vancouver, BC & Bloomington, IN (US Permanent Resident)
jifengwu2k@gmail.com | abbaswu.github.io

Education

University of British Columbia *Vancouver, Canada*
MSc. in Computer Science *September 2022 - November 2024*
- Cumulative GPA: 4.00/4.00

Wuhan University *Wuhan, China*
BEng. in Software Engineering *September 2018 - June 2022*
- Cumulative GPA: 3.93/4.00

Publications

Conference Articles

QuAC: Quick Attribute-Centric Type Inference for Python
Jifeng Wu, Caroline Lemieux
Proceedings of the ACM on Programming Languages OOPSLA (2024). 2024

Journal Articles

Effective Stack Wear Leveling for NVM
Jifeng Wu, Wei Li, Libing Wu, Mengting Yuan, Chun Jason Xue, Jingling Xue, Qingan Li
IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (2023). 2023

Research Projects

ExploTest: Dynamic Unit Test Extractor for iPython *University of British Columbia*
Zihao Huang, **Jifeng Wu**, Caroline Lemieux *May 2024 - August 2024*
- We developed ExploTest, a tool that automates the extraction of unit tests from exploratory testing sessions within iPython, aiding future compatibility and regression testing. It supports complex inputs, system-level runs, and handles challenges like hard-to-serialize data and nondeterministic behaviors. Preliminary results demonstrate its effectiveness on a set of Python benchmarks.

PBTFactory: Automated Property-Based Testing Using Large Language Models *University of British Columbia*
Gary Wu, **Jifeng Wu**, Caroline Lemieux *May 2024 - August 2024*
- We developed a multi-stage pipeline for generating Property-Based Tests (PBT) using LLMs to enhance software correctness and testing efficiency.

QuAC: Quick Attribute-Centric Type Inference for Python *University of British Columbia*
Jifeng Wu, Caroline Lemieux *January 2023 - August 2024*
- We implemented QuAC, a novel Python type inference tool that leverages attribute sets and information retrieval techniques to predict types. QuAC demonstrated high accuracy and coverage, outperforming non-LLM baselines in predicting container type parameters and non-builtin types, and significantly reducing run times. Compared to LLM-based methods, QuAC is nearly two orders of magnitude faster while maintaining consistency in predictions, making it a practical and efficient solution for enhancing static checking and bug-finding tools in dynamically typed Python environments.

Impact of Synthetic Data on Image Captioning models *University of British Columbia*
EECE 571F Deep Learning with Structure (Group Project) *October 2023 - December 2023*
- Image classifiers trained on real data augmented with data from "in-the-wild" generative models achieve high accuracy and effective robustness. Inspired by such previous work, we explore whether we can achieve similar results for image captioning models.

Implementation and Comparison of Syntax-Guided Program Synthesis Techniques *University of British Columbia*
CPSC 548 Directed Studies (Personal Project) *January 2023 - April 2023*
- Program synthesis, the automatic creation of programs that meet user intent, is an important topic in programming theory. I examined core concepts and principles, implemented and compared classic syntax-guided synthesis algorithms, offering insights into their strengths, weaknesses, challenges, and future research directions.

Effective Stack Wear Leveling for NVM *Wuhan University*
Jifeng Wu, Qingan Li *August 2021 - February 2023*
- We proposed Loop2Recursion, a software-based approach for increasing the lifespan of non-volatile memory (NVM) with limited write durability, such as phase change memory (PCM), by converting wear-heavy loops in programs into recursive functions. Implemented as an LLVM pass, Loop2Recursion is applicable to a large variety of hardware architectures, operating systems, and programming languages.

Implementation and Comparison of Marker Selection Techniques *University of British Columbia*

CPSC 545 Algorithms for Bioinformatics (Group Project) *October 2022 - December 2022*

- Advances in genomics and microscopy have boosted the use of single-cell RNA sequencing (scRNA-seq) in biomedical research. However, scRNA-seq data is large-scale and high-dimensional, posing challenges in analysis and reducing model reliability for downstream tasks. Marker selection, which identifies key genes that significantly contribute to cell type classes, can address these issues. We review marker selection methods from the bioinformatics community and test them on real-world scRNA-seq datasets.

Dynamically Inspecting Python Bytecode *University of British Columbia*

CPSC 507 Software Engineering (Personal Project) *October 2022 - December 2022*

- I implemented a modified Python interpreter allowing user-defined callbacks to inspect Python bytecode during the execution of a program. This is an ideal starting point for building novel dynamic program analysis tools for Python.

Community Detection Using Social Network and Trajectories *Wuhan University*

Jifeng Wu, Yuanyuan Zhu *September 2019 - August 2021*

- Many online communities in social networks do not translate to real-world closeness. We explore using user trajectories to identify frequently meeting cohesive user groups. We developed an algorithm for calculating spatiotemporal similarity between trajectories in linear time and another algorithm for community detection integrating social cohesion and trajectory similarity. Comprehensive evaluations on two datasets demonstrated the effectiveness and efficiency of these algorithms in pinpointing real-world cohesive groups.

Conference Control System Based on Gesture Recognition *Wuhan University*

Service Outsourcing Innovation and Entrepreneurship Competition for Chinese College Students (Group Project) *January 2021 - May 2021*

- We implemented a computer vision-based application that captures video from a computer's webcam, recognize 5 hand gestures, and use the recognized gestures to control a computer — utilizing a novel, declarative pipeline parallelism framework for enhanced multicore performance.

Effective Search of Gadgets in the "Attack Lab" Experiment *Wuhan University*

Computer Systems: A Programmer's Perspective (Personal Project) *December 2020*

- To find a set of "gadgets" within a binary that could be exploited in Return-oriented programming (ROP) to complete a given task, I designed a scheme to store gadgets within a Pandas DataFrame, enabling the use of exploratory data analysis to rapidly query all possible gadgets.

Traffic Scene Smart App *Wuhan University*

China Software Cup (Group Project) *April 2020 - August 2020*

- We implemented a computer vision-based application that can identify motor vehicles, non motor vehicles, pedestrians and the license plates, brands, orientations and colors of motor vehicles, monitor traffic flow, as well as record traffic violations.

Service

Research Assistant *University of British Columbia*

Software Practices Lab *January 2023 - Present*

Conducted research with Prof. Caroline Lemieux in Software Engineering.

Teaching Assistant *University of British Columbia*

CPSC 410 Advanced Software Engineering *September 2022 - December 2022*

Gave in-depth feedback and advice regarding students' course projects in Static and Dynamic Program Analysis.

Freshman Mentor *Wuhan University*

School of Computer Science *September 2020 - June 2021*

Introduced freshmen students of Class 10, Grade 2020 to university life and Computer Science, actively answering their questions.

Skills

Domains	Data Science (Data Analysis, Data Visualization, Machine Learning/Deep Learning, Bioinformatics, Graph Data Mining, Constraint Programming), Scripting, Systems Programming, GUI Programming, DevOps.
Languages	Python, C++, Unix command line (Bash/Zsh).
Frameworks	NumPy, Pandas, Matplotlib, scikit-learn, PyTorch, NetworkX, SymPy, Z3; Python C-API, PyBind11, LLVM, Intel Pin; PyQt/PySide; Version Control Systems (Git/GitHub); Debuggers (gdb/lldb/pdb); Containerization (Docker).
Software	Adobe software (Photoshop, Illustrator, Premiere), Office software (Word, PowerPoint, Excel).
Systems	Linux, macOS, Windows.
Soft Skills	Interdisciplinary Vision and Learning Capabilities; Mentoring and Tutoring; Teamwork and Collaboration; Literature Review, Requirements Analysis, System Design, Experiment Design; Exploratory Data Analysis; Code Comprehension and Debugging; Oral and Written Communication and Presentation.