JIFENG WU

Ithaca, NY • jifengwu2k@gmail.com • github.com/jifengwu2k • jifengwu2k.github.io

EDUCATION

UNIVERSITY OF BRITISH COLUMBIA **Vancouver, Canada**

MSc. in Computer Science **November 2024**

- GPA 4.00, TOEFL: Reading 30, Listening 30, Speaking 27, Writing 25
- Coursework: Software Engineering, Automated Testing, Type Systems, Deep Learning with Structure

WUHAN UNIVERSITY **Wuhan, China**

BEng. in Software Engineering, Outstanding Graduate **June 2022**

- GPA 3.93, GRE: Verbal Reasoning 165, Quantitative Reasoning 170, Analytical Writing 4.5
- Coursework: Operating Systems, Database Systems, Software Engineering, Embedded Software Design, Computer Systems: A Programmer's Perspective, Compilation Technology and Application

EXPERIENCE

UNIVERSITY OF BRITISH COLUMBIA **Vancouver, Canada**

QuAC: Quick Attribute-Centric Type Inference for Python **January 2023 - August 2024**

- Implemented QuAC, a novel Python type inference tool using attribute sets and information retrieval.
- QuAC demonstrated high coverage and accuracy, outperforming non-LLM baselines, especially in predicting container type parameters and non-built-in types, while significantly reducing run times.
- Compared to LLM-based methods, QuAC is nearly two orders of magnitude faster while achieving greater consistency in its container type parameter predictions.
- Accepted to *Proceedings of the ACM on Programming Languages, Issue OOPSLA*, 2024
- GitHub: github.com/jifengwu2k/quac

UNIVERSITY OF BRITISH COLUMBIA **Vancouver, Canada**

Impact of Synthetic Data on Image Captioning Models **October 2023 - December 2023**

- Image classification models trained on real data augmented with data from "in-the-wild" generative models can achieve high accuracy and robustness. However, for *image captioning models*, we found that combining real and generated data during training can hurt their performance and robustness.
- Nevertheless, image captioning models trained on synthetic images and fine-tuned with real data outperform those trained solely on real data, highlighting the potential of using synthetic images for pretraining.
- GitHub: github.com/jifengwu2k/The-Impact-of-Synthetic-Data-on-Image-Captioning-Models

WUHAN UNIVERSITY **Wuhan, China**

Effective Stack Wear Leveling for NVM **August 2021 - August 2022**

- Proposed Loop2Recursion, a stack wear leveling technique implemented as an LLVM pass for increasing the lifespan of NVM with limited write durability by converting wear-heavy loops into recursive functions.
- Loop2Recursion outperforms state-of-the-art methods by significantly improving stack wear leveling and reducing performance overhead.
- Published in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* 2023
- GitHub: github.com/jifengwu2k/loop2recursion

WUHAN UNIVERSITY **Wuhan, China**

Community Detection Using Social Networks and Trajectories **September 2019 - August 2021**

- Explored using trajectories to identify spatiotemporally cohesive user groups in large social networks.
- Developed a linear-time spatiotemporal trajectory similarity algorithm and community detection algorithm.
- Comprehensive evaluations on two datasets demonstrated the effectiveness and efficiency of these algorithms.
- GitHub: github.com/jifengwu2k/community-detection-using-social-relations-and-trajectories

OPEN-SOURCE CONTRIBUTIONS

INTERACTIVE DEPENDENCY RESOLVER **PyPI** (**Python 2**+)

- A tool to interactively resolve Python wheel dependencies for debugging "dependency hell" scenarios.
- PyPI: pypi.org/project/interactive-dependency-resolver

PYTHON C-API WRAPPER **GitHub** (**Python 2**+, **C++11**)

- Lightweight C++11 wrapper for Python's C API using RAII for safe reference management.
- GitHub: github.com/jifengwu2k/python-api-wrapper

IPYTHON KERNEL EXECUTOR **PyPI** (**Python 2**+)

- Isolated, sandboxed Python execution environment with comprehensive output capture capabilities.
- PyPI: pypi.org/project/ipython-kernel-executor

PROCLAUNCH **PyPI** (**Python 2**+)

- Spawn Windows/Unix processes through *ctypes* and native APIs, without *subprocess* or shell interpretation.
- PyPI: pypi.org/project/proclaunch

RESUMABLE FILE SERVER **PyPI** (**Python 2**+)

- Drop-in replacement for *http.server* with support for HTTP Range requests for efficient large file transfers.
- PyPI: pypi.org/project/resumable-file-server

UNICODE STRING TO IDENTIFIER **PyPI** (**Python 2**+)

- Text normalization tool that converts arbitrary input to valid Python identifiers while preserving semantics.
- PyPI: pypi.org/project/unicode-string-to-identifier

PUT-BACK ITERATOR **PyPI** (**Python 2**+)

- Iterator that allows putting items back after consuming them for lookahead or backtracking.
- PyPI: pypi.org/project/put-back-iterator

SKILLS

- Languages: Python (Expert), C/C++ (Advanced), Bash.
- Frameworks: PyTorch, NetworkX, PyQt/PySide, LLVM, Intel Pin, Git, Docker.
- Domains: Compilers, Type Systems, Systems Programming, Developer Tooling.

AWARDS

- 3x Outstanding Student Scholarship (Top 5%); 2x Advanced Individual in Social Work
- China Software Cup (Second Prize, 2020); Service Outsourcing Innovation and Entrepreneurship Competition for Chinese College Students (Third Prize, 2021)