Leon (Liang) **Lu**

Email lianglu@cs.cmu.edu
Phone (412) 616-6391

LINKS

GitHub github.com/chaosarium
LinkedIn linkedin.com/in/leonlianglu
Website chaosarium.xyz

EDUCATION

Carnegie Mellon University – School of Computer Science Aug. 2022 – Expected Dec. 2025
B.S. in Computer Science (with planned concentrations in Language Technologies and Computer Systems) Pittsburgh, PA
Cumulative GPA: 4.0/4.0
Relevant Coursework: Computer Systems, Algorithms and Data Structures (Sequential and Parallel), Imperative Computation, Functional Programming, Intermediate Python | Current: Advanced NLP (graduate level), Compiler Design, Database Systems
Selected Class Projects: Compiler (subset of C to x86-64), Retrieval-augmented generation, Dynamic memory allocator, Proxy server

EXPERIENCE

Neural Protolanguage Reconstruction, LLab at Language Technology Institute, Carnegie Mellon University
Researcher | May – Aug. 2023; **Research Group Member** | Nov. 2022 – current
- Pioneering **novel neural techniques** to improve **reconstructions** of **protolanguage** word forms, mentored by Professor Mortensen
- Achieved **3% and 1% accuracy improvement** on **Middle Chinese and Latin** reconstruction respectively compared to state-of-the-art methods (**Python**, **PyTorch**, 10k lines, 4,000+ experiments on WandB)
- **First author** of one accepted publication in LREC-Coling 2024; first author of one other pending publication

Mohimani Lab, Computational Biology Department, Carnegie Mellon University
Software Developer Research Assistant | Jan. – May 2023
- Developed command line tools with Professor Mohimani's research team to process **terabyte-scale** genomics and metabolomics **data** for antibiotics discovery (**Rust**, 80k lines, fully **documented** and **unit-tested** code base)
- **Optimized** the **large-scale** tandem mass spectrometry data **clustering algorithm** Networking+ and **improved accuracy** of the **memory-efficient** mass spectrometry search tool MASST+, achieving performance **2 orders of magnitude faster** than existing tools
- One publication in **Nature Biotechnology**: Fast mass spectrometry search and clustering of untargeted metabolomics data (DOI doi.org/10.1038/s41587-023-01985-4)

School of Computer Science, Carnegie Mellon University
Teaching Assistant, 15-150 Principles of Functional Programming | Aug. 2023 – current
- Designing and grading homework assignments on **functional programming** in the **Standard ML** programming language
- Teaching labs to 15+ students and holding weekly office hours to help students understand course content and appreciate functional programming

Thousand Sunny Technology Ltd.
Software Developer Intern | Jun. – Jul. 2022 | Shenzhen, China
- Built blockchain testing infrastructure automation system with **Electron** desktop application and **React** front-end
- Developed **internal dashboard** with **REST API**, **Redis**, and **JavaScript** to control and monitor hundreds of nodes

PROJECTS [more on online portfolio]

Influx Jul. 2023 – current
- **Designing** and **building** an open-source, NLP-enhanced, content-based **language learning app**
- **Developing** algorithms and data structures for text processing (Rust), multi-lingual tokenization and lemmatization integrations (Python), back-end with **REST API** (**Rust**, Axum, SurrealDB), and front-end user interfaces (**Svelte**, **TypeScript**)

C0VM.TS Dec. 2022 – Jan. 2023
- Contributed to a browser-based **editor**, **bytecode interpreter**, and **debugger** for CMU's c0 programming language (c0 is a safe subset of C) (**TypeScript**, **React**)

NCPA room reservation system 2019 – 2022
- Implemented and maintained online booking systems for music practice rooms and art studios on campus (**JavaScript**, **Express.js**)
- Managed Git repository, performed code review, and incorporated code contributions

HONORS & AWARDS

Honorable Mention, 2024 Outstanding Undergraduate Researcher Award, Computing Research Association Dec. 2023

Dean's List, High Honors, Carnegie Mellon University Jan. 2024, May 2023, Dec. 2022

SKILLS

Programming Python, Rust, C, JavaScript, TypeScript, SQL, Standard ML, OCaml, x86-64 assembly, C++, HTML, CSS, Java
Technologies **Data and ML:** PyTorch, NumPy, Pandas, NLP, Deep Learning, Neural Networks, Research; **Full-Stack:** Node.js, Axum, Express.js, Flask, Svelte, React, Electron, MongoDB, REST API; **General:** Algorithms, Git, Regex, Linux, LaTeX