Donald Pinckney

donald_pinckney@icloud.com · https://donaldpinckney.com · Google Scholar · github.com/donald-pinckney

Skills

Languages: Rust, Python, JavaScript/TypeScript, SQL, C/C++, Swift, Haskell, Go, OCaml.
Tools: Docker, PostgreSQL, AWS, serverless computing (Lambda), CI/CD, React, CouchDB, Redis.
Specialties: building agentic systems, code-assistant agents, enterprise AI integrations, developer tooling / program analysis, dependency / package management, serverless computing, compilers, formal verification, distributed systems, high-performance computing

Experience

Gitar (venture-backed startup, gitar.ai), *Software Engineer* Apr 2024 – Present

- Building Jimy, the first AI coding agent which leverages deep static analysis and source code-level optimization algorithms to **scale agentic AI to massive enterprise codebases while minimizing hallucinations**.
- Leading the AI research team on designing and implementing techniques for **bridging core compiler algorithms with generative AI** in Jimy.
- Architected and implemented static analysis technologies for **automated code refactoring solutions** for enterprise clients, including multiple **Fortune 500 companies**.
- Spearheaded the design and implementation of a **novel testing strategy** based on automated test generation and transpilation to **catch bugs early** during the critical development phase. Found many deep compiler analysis bugs, thus prevented sending **incorrect automated pull requests** to clients.
- Oversaw client onboarding, identified their **software requirements**, and designed solutions that met their needs while also generalizing to other clients.

Northeastern University, Programming Research Lab, *PhD* Sep 2020 – Nov 2024

- Built a new generation of **intelligent package managers** for JavaScript and Python based on combining **foundational constraint solving algorithms** with **generative AI** to automatically fix common developer issues, such as solving runtime errors, patching security vulnerabilities and reducing code size.
- Supervised and guided an undergraduate student in building a **distributed system** using **relational databases** and **container orchestration** to archive every **NPM** package (over 36 million, 20+ TB) with low-latency (< 1 min) within a large (50,000 CPU core) **high-performance computing** (HPC) cluster.
- Developed a novel methodology (MultiPL-E) to standardize the **evaluation of large language model (LLM) code generation** across 19 programming languages, which is used extensively by researchers at **Hugging Face**, ServiceNow, IBM Research and SAP.

Draper Laboratory, *Research Scientist Intern* Feb 2023 – May 2023

- Contributed to enhancing national security by participating in a **DARPA**-funded defensive **cybersecurity** research program (AMP) automatically verifying correctness of **binary security patches**.

Uber, Programming Systems Group, *Programming Systems Research Intern* May 2020 – Dec 2020

- Designed **dynamic analysis**-based tooling informed by **natural language processing** of crash logs that was used in a company-wide effort to repair over 75% of flaky tests, significantly reducing **CI** backlogs.

University of Massachusetts Amherst, *MS* Sep 2018 – May 2020

- Pioneered the study of **formal semantics** for **serverless computing** (FaaS), laying a theoretical foundation for cloud providers to develop new FaaS abstractions, such as Microsoft Azure's Durable Functions.
- Reduced code size by 23% and sped up programs by 15% for multithreaded **WebAssembly** by extending a **JIT compiler** (Wasmtime) with stack capture instructions in assembly.

Apple Inc., *macOS Frameworks Team, Intern* Jun 2016 – Aug 2016

- Created a new user-interface feature simplifying tab navigation in the native macOS UI framework (AppKit), and perfected reliability of it across first-party and third-party apps so it could ship in **macOS High Sierra**.
- Presented the feature before a distinguished panel, including Apple's Senior Vice President Craig Federighi, earning recognition as one of the **top 10 intern projects** from a pool of hundreds of competitors.

Education

Northeastern University, PhD in Computer Science Sep 2020 – Nov 2024
Focused on Programming Languages, ***GPA: 4.00****, Advised by Drs. Arjun Guha and Jonathan Bell*

University of Massachusetts Amherst, MS in Computer Science Sep 2018 – May 2020
Focused on Programming Languages, ***GPA: 3.87****, Advised by Drs. Arjun Guha and Yuriy Brun*

University of California Davis, BS in Computer Science and Mathematics Sep 2014 – Jun 2018
Double Major in Computer Science & Engineering and Mathematics, ***GPA: 3.94***

Leadership and Involvement

Northeastern University CS2500, *Head Teaching Assistant* Jul 2021 – Dec 2021

- Developed homework assignments that guided students through a learning experience focusing on the fundamental principles of **datatype design** and **functional programming**.
- Managed a team of 100+ TAs to efficiently grade over 1000 assignments weekly while orchestrating engaging office hours to foster student participation and learning.

Citrus Circuits FIRST Robotics Team, *Team Mentor* Jul 2014 – May 2015

- Coached the team to win the **FRC World Championship** for the first time in 2015, and spearheaded the development of advanced statistical methods for optimizing robot draft selection.
- Created an innovative **iOS development-based curriculum** to onboard new students interested in programming to the robotics team, which substantially increased student enrollment and engagement.

Publications

ICSE 2023 *Flexible and Optimal Dependency Management via Max-SMT.* **Donald Pinckney**, Federico Cassano, Arjun Guha, Jonathan Bell, Massimiliano Culpo, Todd Gamblin. [paper] [talk] [github] [install]

MSR 2023 *A Large Scale Analysis of Semantic Versioning in NPM.* **Donald Pinckney**, Federico Cassano, Arjun Guha, Jonathan Bell. [paper]

ESEC/FSE 2023 Demo Track *npm-follower: A Complete Dataset Tracking the NPM Ecosystem.* **Donald Pinckney**, Federico Cassano, Arjun Guha, Jonathan Bell. [paper] [talk] [dataset] [github]

TSE 2023 *MultiPL-E: A Scalable and Polyglot Approach to Benchmarking Neural Code Generation.* Federico Cassano, John Gouwar, Daniel Nguyen, Sydney Nguyen, Luna Phipps-Costin, **Donald Pinckney**, Ming-Ho Yee, Yangtian Zi, Carolyn Jane Anderson, Molly Q Feldman, Arjun Guha, Michael Greenberg, Abhinav Jangda. [paper] [talk] [github] [website]

DLS 2020 *Wasm/k: Delimited Continuations for WebAssembly.* **Donald Pinckney**, Yuriy Brun, Arjun Guha. [paper] [talk] [github] [website]

OOPSLA 2019, Distinguished Paper Award *Formal Foundations of Serverless Computing.* Abhinav Jangda, **Donald Pinckney**, Yuriy Brun, Arjun Guha. [paper] [talk] [website]