

# DONALD PINCKNEY

donald\_pinckney@icloud.com · <https://donaldpinckney.com> · [Google Scholar](#)

## EDUCATION

---

### Northeastern University

Sep 2020 – May 2024

*PhD in Computer Science, **GPA: 4.00**, Advised by Dr. Arjun Guha and Dr. Jonathan Bell*

- *Primary research goals:* Building DevOps tooling to automate **dependency management**, and integrating **large-language models** with **formal verification** to enable trustworthy AI
- *Relevant courses:* Foundations of Distributed Systems

### University of Massachusetts Amherst

Sep 2018 – May 2020

*MS in Computer Science, **GPA: 3.87**, Advised by Dr. Arjun Guha and Dr. Yuriy Brun*

- *Primary research goal:* Increasing performance and security of the emerging Web technologies of **WebAssembly** and **Serverless Computing** by improving compilers and program analysis techniques
- *Relevant courses:* Programming Languages, Compilers, Algebraic Topology, Machine Learning, Neural Networks

### University of California Davis

Sep 2014 – Jun 2018

*Double Major in Computer Science & Engineering and Mathematics, **GPA: 3.94***

## EXPERIENCE

---

### Northeastern University, Programming Research Lab, *PhD Candidate*

Sep 2020 – May 2024

- *Automating Dependency Repair*
  - Reducing time developers waste on dependency management by building an **automated dependency repair** tool for the Python / Pip ecosystem.
  - Combining **fine-tuning large language models** (LLMs) with **formal verification techniques** to harness the complex problem-solving skills of LLMs while preventing hallucinations.
  - Running large-scale experiments using **container orchestration** on a **high-performance computing** (HPC) cluster to statistically infer which versions of packages introduce incompatibilities.
- *Optimizing Package Management*
  - Proposed a foundational theory of dependency management (PacSolve) based on **constraint solving** that enables novel user-customizable optimizations, and implemented the theory using the Z3 Max-SMT solver.
  - Used PacSolve to improve the security and efficiency of package management in the JavaScript ecosystem by building a fork of the **NPM CLI** (MaxNPM) which reduces security vulnerabilities by 33%, decreases code size by 21%, and upgrades versions by 14%.
  - Presented MaxNPM at **ICSE 2023**, the leading software engineering conference.
- *Big-Data Analyses of the NPM Ecosystem*
  - Supervised and guided an undergraduate student in building a **distributed system in Rust** to scrape and archive every NPM package (over 36 million) as they are published with low-latency (< 1 min) within a large (50,000 CPU core) **high-performance computing** (HPC) cluster.
  - Performed massively parallel data analyses using **container orchestration** to study **semantic versioning** and **malware** across NPM, publicly released the dataset, and presented it at **MSR 2023** and **ESEC/FSE 2023**, leading software engineering conferences.
- *Code LLM Evaluation*
  - Collaborated on developing 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.

### University of Massachusetts Amherst, *Graduate Researcher*

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](#). Presented and won **distinguished paper award** at **OOPSLA 2019**, a leading programming languages conference.
- 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. Presented at **DLS 2020**, the leading dynamic language symposium.

### 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**) aimed at modularly verifying correctness of **binary security patches**, using the Lean theorem prover and Z3.

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

- Categorized and quantified the root causes of flaky tests across all of Uber by using **natural language processing** to cluster crash logs, thus informing where to target interventions.
- Built **dynamic analysis**-based tooling, informed by the crash log clustering, that was used in a company-wide effort to repair over 75% of flaky tests, significantly reducing **CI** backlogs.
- Created an effective tool prototype to debug the trickiest flaky tests based on concurrency bugs, by using **constraint solving** and **record-replay** techniques.

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

- Implemented and shipped a **new user-interface feature** for **macOS High Sierra** and onwards, catering to tens of millions of Mac users, significantly simplifying tab navigation.
- 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.

## SKILLS

---

**Languages:** Python, Rust, JavaScript/TypeScript, SQL, C/C++, Java, Swift, Haskell, OCaml, Agda, Idris, Lean, Coq, Racket, R, x86 Assembly, Objective-C, Matlab

**Tools:** Z3, Docker, Singularity, Kubernetes, Slurm, Spack, Redis, PostgreSQL, AWS, Google Cloud Platform, ASM, Emscripten, GitHub Actions CI/CD, version control, build systems, debuggers, profilers

**Specialties:** dependency/package management, serverless computing, formal verification, compilers, security, large language models, distributed systems, high-performance computing, systems design, data mining, mobile development

## 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 Culpò, 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. [\[paper\]](#) [\[talk\]](#) [\[website\]](#)