Ming-Ho Yee

mhyee.com • github.com/mhyee • mh@mhyee.com

Experience

Research Scientist — **Menlo Park CA**
Meta — *Aug 2024–present*

- Worked on the runtime software for the Meta Training and Inference Accelerator, Meta's custom AI silicon.
- Investigated and demonstrated functional readiness for enabling the next generation PyTorch runtime on MTIA.
- Maintained and refactored a testing/benchmarking tool to improve modularity, allowing internal users to easily extend the tool with custom functionality.
- Implemented APIs for the new Rust runtime, and validated functionality on a variety of test suites and benchmarks.

Research Assistant and Teaching Assistant — **Boston MA**
Northeastern University — *Sep 2016–Apr 2024*

- Built TypeWeaver, the first machine-learning-based tool to migrate files from JavaScript to TypeScript, with a 69% success rate (as measured by type checking).
- Fine-tuned and evaluated a code LLM to generate type definitions for TypeScript, allowing 47% of files (with missing type definitions) to type check (22% absolute improvement).
- Built an interpreter in OCaml for a subset of R to model static and dynamic program analysis.
- Sped up test suite by 15% by optimizing dominance graph construction in Ř, a just-in-time compiler for R.
- Co-chaired, organized, and led over 40 student volunteers at ECOOP/ISSTA 2018, an international conference for programming languages and software engineering with over 600 attendees.
- Mentored undergraduate, master's, and Ph.D. students by providing feedback and advice on project planning, software development, and written and oral communication.
- Teaching assistant for "Fundamentals II, Introduction to Class-based Program Design" (CS 2510) and "Fundamentals of Software Engineering" (CS 4530): helped design assignments, held office hours, and graded exams.

Researcher Intern — **Cambridge UK**
Microsoft Research — *Sep–Dec 2019*

- Implemented and evaluated different memory management strategies for Project Verona, a research programming language for concurrent ownership.

Research Assistant and Teaching Assistant — **Waterloo ON**
University of Waterloo — *Sep 2014–Aug 2016*

- Designed, implemented, and benchmarked the interpreter and compiler for the functional sub-language of Flix.
- Experimented with different code generation techniques for Flix, such as using Scala macros.
- Teaching assistant for "Foundations of Sequential Programs" (CS 241, CS 241E) and "Compiler Construction" (CS 444): held office hours and provided feedback on assignments and exams for the first offering of CS 241E.

Software Development Engineering Intern — **Redmond WA**
Microsoft — *Sep–Dec 2013; May–Jul 2014*

- Implemented user-defined literals, a C++11 feature, in the Microsoft Visual C++ compiler, bringing the compiler closer to C++ standards conformance.
- Prototyped concepts lite in the Microsoft Visual C++ compiler, a feature that was eventually added to the C++ standard and shipped in Visual Studio 2019.

Software Development Engineering Intern — **Redmond WA**
Microsoft — *Jan–Apr 2013*

Undergraduate Research Assistant — **Waterloo ON**
University of Waterloo — *Apr–Dec 2012*

Web Developer Intern — **Toronto ON**
Eqentia — *Jan–Apr; Aug–Dec 2011*

Software Development Research Intern — **Markham ON**
Genesys Telecommunications Laboratories — *May–Aug 2010*

Software Developer Intern — **London ON**
Robarts Research Institute — *Jun–Aug 2008*

Publications

MH Yee and A Guha (2023). Do Machine Learning Models Produce TypeScript Types That Type Check?, *ECOOP*.

L von Werra, H de Vries, et al. (2023). StarCoder: may the source be with you!, *TMLR*.

F Cassano, **MH Yee**, N Shinn, A Guha, S Holtzen (2023). Type Prediction With Program Decomposition and Fill-in-the-Type Training, *technical report*.

F Cassano et al. (2023). MultiPL-E: A Scalable and Polyglot Approach to Benchmarking Neural Code Generation, *TSE*.

O Flückiger, G Chari, **MH Yee**, J Ječmen, J Hain, J Vitek (2020). Contextual Dispatch for Function Specialization, *OOPSLA*.

O Flückiger, G Chari, J Ječmen, **MH Yee**, J Hain, J Vitek (2019). R Melts Brains: An IR for First-Class Environments and Lazy Effectful Arguments, *DLS*.

MH Yee, A Badouraly, O Lhoták, F Tip, J Vitek (2019). Precise Dataflow Analysis of Event-Driven Applications, *technical report*.

O Flückiger, G Scherer, **MH Yee**, A Goel, A Ahmed, J Vitek (2018). Correctness of Speculative Optimizations with Dynamic Deoptimization, *POPL*.

M Madsen, **MH Yee**, O Lhoták (2016). From Datalog to Flix: A Declarative Language for Fixed Points on Lattices, *PLDI*.

M Safa, **MH Yee**, D Rayside, C T Haas (2016). Optimizing Contractor Selection for Construction Packages in Capital Projects, *ASCE J. Comput. Civ. Eng.*

E Zulkoski, C Kleynhans, **MH Yee**, D Rayside, K Czarnecki (2014). Optimizing Alloy for Multi-objective Software Product Line Configuration, *ABZ*.

R Bartha, **MH Yee**, R Rupsingh, M Smith, M Borrie (2009). Altered macromolecule signal in the hippocampus in alzheimer patients measured by 1H magnetic resonance spectroscopy, *Alzheimer's & Dementia*.

Technical Skills

- Implementation experience: interpreters, just-in-time compilers, memory management, program analysis.
- Languages: C, C++, Rust, Java, Scala, OCaml, Python, Ruby, JavaScript, TypeScript, R.
- Compilation targets: ARM, JVM, LLVM, MIPS, x86.

Education

Doctor of Philosophy in Computer Science — **Boston MA**
Northeastern University — *Apr 2024*
- Thesis: Predicting TypeScript Type Annotations and Definitions with Machine Learning
- Advisor: Arjun Guha

Master of Mathematics in Computer Science — **Waterloo ON**
University of Waterloo — *Jun 2017*
- Thesis: Implementing a Functional Language for Flix (completed Sep 2016)
- Advisor: Ondřej Lhoták

Bachelor of Software Engineering — **Waterloo ON**
University of Waterloo — *Jun 2014*
- With Distinction — Dean's Honours List