

# Ming-Ho Yee

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

---

## EXPERIENCE

### Ph.D. Candidate

Northeastern University

Boston MA

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 large language model for code 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 the relationship between static and dynamic program analysis.
- Sped up test suite by 15% by optimizing dominance graph construction in  $\tilde{R}$ , 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

Microsoft Research

Cambridge UK

Sep – Dec 2019

- Explored, implemented, and tested different memory management strategies for [Project Verona](#).

### MMath Candidate

University of Waterloo

Waterloo ON

Sep 2014 – Aug 2016

- Designed and led the implementation of the original interpreter and compiler for the functional sub-language of [Flix](#).
- Experimented with different code generation techniques for Flix, such as Scala macros and generating Scala code.
- 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

Microsoft

Redmond WA

May – Jul 2014

- Prototyped [concepts lite](#) in the Microsoft Visual C++ (MSVC) compiler, a feature that was eventually added to C++20.

### Software Development Engineering Intern

Microsoft

Redmond WA

Sep – Dec 2013

- Implemented [user-defined literals](#) in the Microsoft Visual C++ (MSVC) compiler, a C++11 feature that was missing from MSVC.

### Software Development Engineering Intern

Microsoft

Redmond WA

Jan – Apr 2013

- Developed a heap memory collection tool for debugging .NET applications.
- Designed and conducted performance tests for the memory collection tool.

### Undergraduate Research Assistant

University of Waterloo

Waterloo ON

May – Dec 2012

- Analyzed software dependencies using the LLVM infrastructure.
- Investigated automatic test case generation from stack traces.

### Developer

Engagio (formerly Eqentia)

Toronto ON

Sep – Dec 2011

### Developer

Eqentia

Toronto ON

Jan – Apr 2011

### Software Development Research Intern

Genesys Telecommunications Laboratories

Markham ON

May – Aug 2010

### Junior Developer

Robarts Research Institute

London ON

Jul – 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](#), *preprint*.
- 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++, Java, Scala, OCaml, Python, Ruby, JavaScript, TypeScript, R.
- Compilation targets: ARM, JVM, LLVM, MIPS, x86.

## EDUCATION

### Doctor of Philosophy in Computer Science

*Northeastern University*

- Thesis: [Predicting TypeScript Type Annotations and Definitions with Machine Learning](#)
- Advisor: [Arjun Guha](#)

**Boston MA**

*Apr 2024*

### Master of Mathematics in Computer Science

*University of Waterloo*

- Thesis: [Implementing a Functional Language for Flix](#) (completed Aug 2016)
- Advisor: [Ondřej Lhoták](#)

**Waterloo ON**

*Jun 2017*

### Bachelor of Software Engineering

*University of Waterloo*

- With Distinction — Dean's Honours List

**Waterloo ON**

*Jun 2014*