

EDUCATION

- **Carnegie Mellon University** Pittsburgh, PA
PhD in Computer Science. Advisors: Ruben Martins and Claire Le Goues 2021 - 2025
 - Thesis: Integrating static analysis with large language models for software evolution
 - Research: large language models, software verification, software security
- **Carnegie Mellon University** Pittsburgh, PA
MSc in Computer Science. Advisors: Ruben Martins, Vincent Hellendoorn and Claire Le Goues 2023
- **Queen's University** Kingston, ON
BEng, Computer Engineering and Mathematics Dual Degree 2016-2020
 - Thesis: Modeling the loss function of generative adversarial networks (GANs) with Rényi information measures

EXPERIENCE

- **Amazon AWS** Seattle, WA
Applied Scientist II (AGI) Jan 2025 - Present
 - LLM agentic-coding team: trained 360B and 450B parameter LLMs (e.g., Nova Pro, Nova Reasoning) for mathematical super intelligence and coding correctness with SFT, RL, and RAG.
 - Lead the RL effort to achieve > 13% benchmark lift on agentic math and coding tasks with verified, synthetic data (e.g., Lean specifications and Lean proofs).
 - Managed multiple external data collection projects for SFT and rubrics based RL on verified code, combined with software verification for data augmentation.
- **OpenAI** Remote
Research Resident November 2024 - May 2025
 - Finetuning, benchmarking, and evaluating GPT model capabilities as part of the Preparedness team.
- **Amazon AWS** Portland, OR, Arlington, VA
Applied Scientist Intern Summers 2023, 2024
 - RAG and RL on Mathlib4 library for LLM Lean proofs generation. Achieved 15% lift on IMO and AIME etc. benchmarks.
 - Finetuned a 16B parameter LLM for Neural Machine Translation (NMT) on the Rust programming language, resulting in a patent filing for the AWS Amazon Automated Reasoning Group.
- **Microsoft Research** Redmond, WA
Research Intern Summer 2022
 - RL and formal methods (FM) research as part of the MSR RiSE group.
 - Implemented the AlphaZero reinforcement learning model for synthesizing Nvidia collective communication algorithms. Sped up Monte Carlo Tree Search (MCTS) algorithm by 70% with DGX-1 GPU optimization.
- **Advanced Micro Devices (AMD)** Toronto, ON
Software Engineer May 2020 - Sept 2021
 - High Dynamic Range (HDR) and Freesync features for GPU driver development.

SELECT PUBLICATIONS

- **Revisiting Unnaturalness for Automated Program Repair with Large Language Models**
Aidan Z.H. Yang, Sophia Kolak, Vincent Hellendoorn, Ruben Martins, Claire Le Goues
IEEE International Conference on Software Engineering (ICSE), 2025
- **Large Language Models for Test-Free Fault Localization**
Aidan Z.H. Yang, Ruben Martins, Claire Le Goues, Vincent Hellendoorn
IEEE International Conference on Software Engineering (ICSE), 2024
- **SOAR: A Synthesis Approach for Data Science API Refactoring**
Ansong Ni, Daniel Ramos, Aidan Z.H. Yang, Ines Lynce, Vasco Manquinho, Ruben Martins, Claire Le Goues
IEEE International Conference on Software Engineering (ICSE), pp. 112-124, 2021

PATENTS

- Verified Equivalent Rust Transpilation with Large Language Models, filed with Amazon AWS

TEACHING

- Lead TA for “Research and Innovation in Computer Science” at CMU, Fall 2024
- Lead TA for “Large Language Models and Applications” at CMU, Fall 2023
- Lab TA for “Data Structures and Algorithms” at Queen’s University, Winter 2018

SERVICE AND AWARDS

- Tenure track hiring committee member at Carnegie Mellon University, 2024-2025
- Program committee member (technical paper reviewer) for ASE 2024
- Software engineering PhD graduate assembly representative, 2021-2024
- Finalist at SPLASH Graduate Student Research Competition, 2020
- SIGSOFT CAPS Travel Award for MSR 2019, SPLASH 2019, ICSE 2019, and ICSE 2020