

# Alexander Shypula

*“We choose to go to the Moon in this decade and do other things,  
not because they are easy, but because they are hard” - JFK*

## Personal info

---

**Name:** Alexander Shypula

**Address:** 2626 Tunnel Boulevard 458, Pittsburgh PA, 15203

**Phone number:** +1 (908) 745 1361

## Education

---

**Carnegie Mellon University**

Graduated 2021

*Master of Science in Artificial Intelligence and Innovation*

*Language Technologies Institute, School of Computer Science*

*GPA: 4.020 (4.0+ grading system)*

- Member of Neulab (Associate Professor Graham Neubig’s Lab).
- Relevant coursework: Program Analysis (PhD), Neural Networks for NLP (PhD), Multilingual NLP (PhD), Introduction to Deep Learning (PhD), Introduction to Machine Learning (PhD), Deep Reinforcement Learning and Control (PhD).

**New York University**

Graduated 2017

*Bachelor of Science (Statistics and Finance)*

*GPA: 3.702 (4.0 grading system, Cum Laude)*

*Note: a leave of absence was taken for one semester to pursue a Fulbright-Hays scholarship for immersive study in Xi’an, China*

## Major projects

---

**Structure-Aware Program Translation**

Sept 2021 - Present

*Advisors: Yoon Kim (MIT CSAIL), Jie Chen (MIT-IBM Lab)*

- Focus on improving the compositional generalization of neural sequence models when applied to program translation: main emphasis on developing tree-to-tree attention mechanisms by leveraging supervised attention, latent variable models, and program parse tree structure (e.g. abstract-syntax trees and concrete-syntax trees).

**Learning to Superoptimize Real-world Programs**

Jan 2020 - Sept 2021

*Advisors: Graham Neubig (CMU), Edward Schwartz (CMU SEI), Claire Le Goues (CMU)*

- Ambitious research project addressing how neural sequence models can be used to superoptimize programs with insights relevant for tuning large language models.
- Focused on programs at the x86-64 level, for which software existed for generating testcases, instrumenting testcases, and proving equivalence with SMT solvers.
- Demonstrated that standard policy gradient algorithms (REINFORCE) are sub-optimal when compared to bootstrapping a pre-trained model to discover superoptimized programs and training on such superoptimized programs.
- Ultimately superoptimized over the GCC -O3 baseline for 6.2% of all test set programs, outperforming standard policy gradient algorithms by 5x. Additionally, witnessed that the model consistently learned to “fuzz” the STROKE toolkit by discovering and converging on SMT solver exploits.

- Engineering involved: Mining, compiling, and disassembling open source C code to create a parallel corpus of 1.15 million functions in assembly (unoptimized and optimized pairs); engineering a distributed actor-learner architecture for reinforcement learning to overcome RL environment latency (due to program execution and SMT solver verification); dealing with quirky bugs in a SMT solver toolkit; reading x86 assembly to manually verify superoptimized program equivalence.

## Employment

---

**Massachusetts Instituted of Techology (CSAIL)** 2021 - Present  
*Visiting Researcher / Research Specialist*

- Work on structure-aware program translation with Yoon Kim (MIT CSAIL) and Jie Chen (MIT-IBM Watson AI lab)

**IBM** 2017-2019  
*Data Scientist*

Highlighted Experiences

- Work with members from IBM Watson and IBM Research speech researcher Dr. Hong-Kwang (Jeff) Kuo to build a speech recognition pipeline to compete for a major fast food company's automated drive-thru transformation project. Project overseen by IBM's Chief Strategy Officer.

## Awards / Recognition / Leadership

---

Received

- 2015: **Fulbright-Hays Scholarship** for immersive Chinese study in Xi'an China
- 2018: **IBM Security and Mobile Invention Development Team Board Member**. Among the top 10 reviewers of patents to prioritize patents for prior art search and the patent filing process.
- 2018: **Consultant of the Year Award**. Highest form of recognition for entry-level consultants; awarded to 10 consultants out of over 800 that year.
- 2019-2021: **CMU MSAIL Graduate Student Assembly Representative**. Created a research reading group for student enrichment and responsible for an \$18K budget for student events.

Pending

- National Defense Science and Engineering Graduate (NDSEG) Fellowship program

## Languages

---

*Formal Languages* Python, Java, C, OCaml, x86-64 assembly

*Natural Languages* English (native)  
 Chinese (Mandarin, Intermediate)  
 French (Intermediate)