CELINE LEE

celine.y.lee@gmail.com ◇ cl923@cornell.edu
celine-lee.github.io

EDUCATION

Cornell University *2021 - Present*
PhD student in Computer Science *(Minor in Computational Linguistics)*
Advisor: Alexander (Sasha) Rush

University of Pennsylvania *2016 - 2020*
M.S.E. in Embedded Systems
B.S.E. in Computer Science, Electrical Engineering

PUBLICATIONS

Sampling Language from Latent System 2 Reasoning
(NeurIPS 2024 - System 2 Reasoning At Scale Workshop)
· C. Lee, A. Sultan, T. Naseem, A. Rush, R. Astudillo

The Counterfeit Conundrum: Can Code Language Models Grasp the Nuances of Their Incorrect Generations?
(ACL 2024 - Findings)
· A. Gu, C. Lee*, W. Li*, N. Jain*, T. Olausson*, K. Sen, A. Solar-Lezama

Guess & Sketch: Language Model Guided Transpilation
(ICLR 2024)
· C. Lee, A. Mahmoud, M. Kurek, S. Campanoni, D. Brooks, S. Chong, G.Y. Wei, A. Rush

Mixture of Soft Prompts for Controllable Data Generation
(EMNLP 2023 - Findings)
· D. Chen, C. Lee, Y. Lu, D. Rosati, Z. Yu

Batched Vectorized Earley Parsing
2023 Mid-Atlantic Student Colloquium on Speech, Language and Learning (MASC-SLL 2023)
· C. Lee, A. M. Rush

A Survey on Semantic Parsing for Machine Programming
2021 KDD Workshop on Programming Language Processing (KDD 2021 - PLP Workshop)
· C. Lee, J. Gottschlich, D. Roth

PATENTS

"Methods, apparatus, and articles of manufacture to generate command lists to be offloaded to accelerator circuitry"
(Pending) USPTO Application No. 17/559556

"Methods and apparatus to determine refined context for software bug detection and correction"
(Pending) USPTO Application No. 17/554918

"Methods and apparatus to train models for program synthesis"
(Pending) USPTO Application No. 17/551170

"Methods and apparatus to improve data quality for artificial intelligence"
(Pending) USPTO Application No. 17/540050

HONORS AND AWARDS

University of Pennsylvania 2020 Fred Ketterer Memorial Award
University of Pennsylvania 2020 Walter Korn Award
University of Pennsylvania Dean's List

RESEARCH, WORK EXPERIENCE

IBM Thomas J. Watson Research Center Summer 2024
Research Intern

· Research project in code program-mediated latent modeling for synthetic data generation.

Harvard University Sept 2023 - 2024
SEAS Researcher

· Research methods for incorporating machine learning techniques into code processing, particularly of low-level code such as assembly or compiler intermediate representations. Supervised 3 interns.

Merly.ai Jan 2022 - Aug 2022
Research Scientist & Software Engineer

· Led qualitative & quantitative data analysis and paper writing for MP-CodeCheck, our automatic code anomaly detection system.

Intel Labs Feb - Dec 2021
PI: Prof. Justin Gottschlich

· Developed a research proof-of-concept for a machine-learned data structure for C code.
· Ideated, wrote, and submitted four patent applications in machine learning, program synthesis, and heterogeneous systems.

Cognitive Computation Group, University of Pennsylvania Jan 2020 - Feb 2021
PI: Prof. Dan Roth

· Research and development of nominal, multi-lingual, and joint-predicate semantic role labeler for natural language processing of complex sentence structures.

VMware May - August 2019
Product Development Intern, Member of Technical Staff

· Built out a desired state controller Golang service that sits on VMwares management server, to manage the bringup, duplication, and reconfiguration process of VMware virtualized datacenters.

UPenn Autonomous Vehicles Group Jan - Oct 2019
PI: Prof. Rahul Mangharam

· Build image tracking program for traffic video feed, to develop autonomous driving algorithms.

UTC Aerospace Systems - ISR and Space Systems May - Aug 2018
Electrical Engineering Intern

· Developed software to determine space compliance of electrical board design by parsing specifications and BOM to derate electrical components.
· Developed user-friendly auto-sequencing interface for engineers to communicate with FPGA during testing. (Awarded Excellence Award.)

Singh Nanotechnology Center Jan 2017 - Jan 2018
PI: Prof. Mark G. Allen

· Enabled greater control of the microstructure electrodeposition process by reprogramming the electroplating robot in the UPenn MicroSensors and MicroActuators (MSMA)

TEACHING

Teaching Assistant: Practical Applications of Machine Learning	Cornell University
Head Teaching Assistant: Break Through Tech AI	Cornell University
Teaching Assistant: CS 5781 Machine Learning Engineering	Cornell University
Mentor (4 high school students)	The Polygence Program
Teaching Assistant: CIS 519 Applied Machine Learning	University of Pennsylvania
Teaching Assistant: ESE 215 Intro. Circuit Theory	University of Pennsylvania
Teaching Assistant: ESE 111 Intro. Electrical, Systems Engineering	University of Pennsylvania

SERVICE

Cornell Computer Science Visit Days Czar

Cornell Computer Science Admissions Committee

Cornell Tech PhD Student Board

Cornell Computer Science Graduate Organization