Charles Jin

Research Interests

My current research agenda is developing rigorous scientific methods for characterizing the cognitive and linguistic abilities of LLMs (or their lack thereof). More broadly, I am interested in learning systems whose knowledge is symbolic in nature.

Contact

32 Vassar St, Bldg 32-G730
Cambridge, MA 02139
ccj@csail.mit.edu
charlesjin.com

Education

Massachusetts Institute of Technology, September 2019 - August 2024 (expected).
Ph.D. Student in Computer Science.
Thesis: On the Acquisition of Formal Semantics in Statistical Models of Language.
Advisor: Martin Rinard.

Yale University, Aug 2012 - May 2016.
Combined B.S./M.S. in Computer Science.
B.S. in Mathematics, *with distinction.*
GPA: 3.96/4.00.

Awards and Honors

Schulz Prize, 2016. Awarded to a Silliman College senior for academic excellence in the physical sciences or mathematics.
summa cum laude, Yale, 2016.
Phi Beta Kappa, Yale, 2015.
Moulton Ely Grant, 2014. Small grants that provide support for students in entrepreneurial endeavors.
Sherwood E. Silliman Fellowship, 2013. Covered a 2-week collaboration at Case Western Reserve University.
Yale College First-Year Summer Research Fellowship in the Sciences & Engineering, 2013. Awarded to approximately 70 students per year. Funded a summer of research on project "Image Segmentation of Dense Capillary Meshes."

Preprints

Charles Jin, Zhang-Wei Hong, Farid Arthaud, Idan Orzech, and Martin Rinard. **"Decentralized Inference via Capability Type Structures in Cooperative Multi-Agent Systems"**. arXiv:2304.13957. 2023.

Charles Jin and Martin Rinard. **"Manifold Regularization for Locally Stable Deep Neural Networks"**. arXiv:2003.04286. 2020.

Refereed Publications

Charles Jin and Martin Rinard. **"Latent Causal Probing: A Formal Perspective on Probing with Causal Models of Data"**. 1st Conference on Language Modeling (**COLM 2024**). To appear.

Charles Jin and Martin Rinard. **"Emerging Representations of Formal Semantics in Language Models Trained on Programs"**. Forty-first International Conference on Machine Learning (**ICML 2024**). Vienna, Austria. 2024.

Charles Jin, Melinda Sun, and Martin Rinard, **"Incompatibility Clustering as a Defense Against Backdoor Poisoning Attacks"**. The Eleventh International Conference on Learning Representations (**ICLR 2023**). Kigali, Rwanda. 2023.

Charles Jin, Phitchaya Mangpo Phothilimthana, and Sudip Roy, **"Neural Architecture Search using Property Guided Synthesis"**. Proceedings of the ACM on Programming Languages, Volume 6, Issue OOPSLA2, 1150-1179 (**OOPSLA 2022**). Auckland, New Zealand. 2022.

Charles Jin and Martin Rinard. **"Towards Context-Agnostic Learning Using Synthetic Data"**. Advances in Neural Information Processing Systems 34 (**NeurIPS 2021**). Virtual. 2021.

Limor Appelbaum, Alexandra Berg, Jose Cambronero, Thurston Dang, **Charles Jin**, Lori Zhang, Steven Kundrot, Matvey Palchuk, Laura Evans, Irving Kaplan, and Martin Rinard. **"Development of a pancreatic cancer prediction model using a multinational medical records database"**. Journal of Clinical Oncology (**JCO**) 39:3_suppl, 394-394. 2021.

Muthu Baskaran, **Charles Jin**, Benoit Meister, and Jonathan Springer. **"Automatic Mapping and Optimization to Kokkos with Polyhedral Compilation"**. 2020 IEEE High Performance Extreme Computing Conference (**HPEC20**). Waltham, MA, USA. 2020.

Charles Jin, Muthu Baskaran, Benoit Meister, and Jonathan Springer. **"Automatic Parallelization to Asynchronous Task-Based Runtimes Through a Generic Runtime Layer"**. 2019 IEEE High Performance Extreme Computing Conference (**HPEC19**). Waltham, MA, USA. 2019.

Charles Jin, Muthu Baskaran, and Benoit Meister. **"POSTER: Automatic Parallelization Targeting Asynchronous Task-Based Runtimes"**. 2019 28th International Conference on Parallel Architectures and Compilation Techniques (**PACT19**), 465-466. Seattle, WA, USA. 2019.

Charles Jin and Muthu Baskaran. **"Analysis of Explicit vs. Implicit Tasking in OpenMP Using Kripke"**. 2018 IEEE/ACM 4th International Workshop on Extreme Scale Programming Models and Middleware (**ESPM2**), 62-70, held in conjunction with **SC18**. Dallas, TX, USA. 2018.

Teaching

Undergraduate Science and Quantitative Reasoning tutor at Yale University, Spring 2015 - Spring 2016.

Industry Experience

Google, Student Researcher, September 2021 - June 2022.
Google, Research Intern, June 2021 - September 2021.

- Devised an efficient (psuedo-polynomial time) program synthesis method for performing neural architecture search.
- Implemented technique within a large-scale evolutionary framework.
- Accepted at OOPSLA 2022 as "Neural Architecture Search using Property Guided Synthesis".

Reservoir Labs, Research Engineer, June 2018 - August 2019.

- Implemented new backends for a polyhedral optimizing compiler for targeting task-based runtimes (Legion and OpenMP).
- Designed a new lightweight runtime layer to enable automatic extraction of dynamic task-based parallelism. Extended compiler backend to support heterogeneous dynamic task-based parallelism using GPUs (CUDA).
- Evaluated performance of parallel programming models targeting exascale systems with heterogeneous architectures (e.g., OpenMP, Legion, Charm++, Kokkos, OCR).
- Contribute to reports and papers, including grant proposals and reports.