Ketaki Joshi

ketaki.joshi@yale.edu
+1 203-906-4461
https://joshi-ketaki.github.io/

Experience

- **Yale University** New Haven, USA
  Ph.D candidate (Advisor: Abhishek Bhattacharjee) August 2019 - Present
  - Mitigating catastrophic forgetting using principles of Episodic Memory (Ongoing Research).
    - Developed a cognitively inspired context-based memory technique to mitigate catastrophic forgetting in LSTMs for continual learning applications such as memory prefetching.
    - Achieved speedup by factor of 1.6 compared to existing regularization technique and reduced external storage by factor of 16 than existing replay implementation in initial prototype evaluations.
    - Investigating against existing ML-based prefetching solutions.
  - Prefetching using principles of Complementary Learning System (CLS).
    - Investigated use of CLS inspired techniques to improve prefetching in GPUs.
    - Identified and investigating existence of natural replay in GPUs. This would potentially remove the need to implement explicit replay techniques to avoid catastrophic forgetting.
  - Enhance cognitive systems research using algorithmic principles.
    - Developed a similarity detection technique to identify computational similarity between cognitive models.
    - This tool helps neuroscientists reuse, build and understand cognitive models.
  - Developed asynchronous system calls in CertiKOS which were two orders of magnitude faster than existing synchronous system calls.

  Teaching Fellow- Introduction to Systems Programming Spring '21, Spring '22, Fall '22
  - Deployed a toy-compiler to introduce students to local, global, peephole optimizations, register allocation and assembly code generation.

- **NVIDIA** Santa Clara, USA
  Unified Virtual Memory Intern(Mentor: Guilherme Cox) June 2022 - August 2022
  - Deployed an access-aware eviction algorithm that enabled support for irregular access patterns along with traditional streaming patterns.
  - Achieved a performance improvement of two orders of magnitude in the best case and same as the existing algorithm in the worst case.

  Architecture Research Intern(Mentor: Daniel Lustig) June 2021 - October 2021
  - Delivered a driver shim to execute CUDA programs transparently under the application's hood on CPU SIMD units along with GPUs.
  - Achieved speedup of a factor of 1.5 compared to pure CPU-SIMD execution and a factor of 1.8 compared to a pure GPU execution.

- **NVIDIA** Pune, India
  GPU Compiler Developer January 2017 - August 2019
  - Delivered compiler frontend and backend interface design, support and assembly generation for the deep learning matrix operations. The instructions were exposed in CUDA 10.0 and CUDA 10.1. (MMA etc.)
  - Delivered the entire assembly generation and decoding for Turing architecture.

- Led analysis of key benchmarks to identify opportunities for using newly introduced uniform register file within the compiler.
- Led the design and implementation of a framework to auto generate assembly and decoding of assembly instructions for the compiler.

Tools Developer June 2015 - July 2016

- Developed a no-reference image analysis tool to detect artefacts in images rendered across different GPU architectures. Eradicated existing manual analysis and achieved 98% accuracy.
  Submitted to Nvidia internal conference and filed an Invention Submission Form

- **Shoreline IoT** Pune, India
  Member of Technical Staff September 2016 - January 2017
  - Led prototype development of a device for remote maintenance of IoT systems.

- **Indian Institute of Technology, Bombay** Mumbai, India
  Undergraduate Thesis Intern (Advisor: Uday Khedker) May 2014 - June 2015
  - Developed a custom compiler optimizer generator from given local and global data flow equations.

- **Indian Institute of Tropical Meteorology** Pune, India
  Undergraduate Research Intern (Mentor: Narendra Karamarkar) February 2014 - June 2014
  - Deployed a N-SAT solver as a tool for use in weather prediction analysis modules.

Education

- **Yale University** New Haven, USA
  Ph.D Computer Science (Advisor: Abhishek Bhattacharjee) August 2019 - Present

  M.Phil in Computer Science August 2019 - 2021
  – **Thesis:** "Single Source Code, Hardware Agnostic Heterogeneous Systems."

  Masters in Computer Science August 2019 - 2020
  – **Thesis:** "Detecting Computational Clones in Brain Models."

- **University of Pune** Pune, India
  Bachelor's in Computer Engineering 2011 - 2015
  Institute Rank : 1/200, University Rank : 5/9000
  – **Thesis:** "OptGen : A Custom Compiler Optimization Generator."

Skills

- **Programming:** C, Python, x86 Assembly, NVIDIA PTX Assembly, CUDA, C++, MATLAB, Octave.
- **ML frameworks:** Pytorch, Raytune.
- **Source Control:** Git, PerForce.
- **Writing:** Latex, Word.

Publications

- **Ketaki Joshi**, Raghavendra Pradyumna Pothukuchi, Andre Wibisono, Abhishek Bhattacharjee "Mitigating Catastrophic Forgetting in Long Short-Term Memory Networks.", arXiv:2305.17244 [cs.LG].
- Wu Michael, **Joshi Ketaki**, Sheinberg Andrew Cox Guilherme, Khandelwal Anurag, Pothukuchi Raghavendra Pradyumna, Bhattacharjee Abhishek, "Prefetching Using Principles of Hippocampal-Neocortical Interaction.", HOTOS '23.

- J. Veselý, R. P. Pothukuchi, **K. Joshi**, S. Gupta, J. D. Cohen and A. Bhattacharjee, "Distill: Domain-Specific Compilation for Cognitive Models.", 2022 IEEE/ACM International Symposium on Code Generation and Optimization (CGO), 2022, pp. 301-312, doi: 10.1109/CGO53902.2022.9741278.
- J. Veselý, R. P. Pothukuchi, **K. Joshi**, S. Gupta, J. D. Cohen and A. Bhattacharjee, "Cognac: Domain-Specific Compilation for Cognitive Models."

Talks

- **NVIDIA - Internship Talk** — Santa Clara, USA
  "Access Guided Eviction for Unified Virtual Memory." — August 2022
- **Yale University - Area Exam Talk** — New Haven, USA
  "Single Source, Hardware Agnostic Heterogenous Systems." — December 2021
- **NVIDIA Research - Internship Talk** — Santa Clara, USA
  "CUDA Task launcher for GPU and CPU SIMD units." — October 2021
- **ACM-W: Cummins College of Engineering for Women, University of Pune** — Pune, India
  "A Custom Compiler Optimization Pass Generator." — September 2015