

Daniel Zou

☎ (703) 939-0901 | ✉ dlzou@berkeley.edu | 📄 dlzou | 🌐 dlzou | 🐙 dlzou.github.io

EDUCATION

University of California, Berkeley

B.Sc. Electrical Engineering and Computer Sciences, with Honors; GPA: 3.94/4.00

Berkeley, CA

Aug 2019 – May 2023

Academy of Science and Technology

High School Diploma; Rank: 2/712

The Woodlands, TX

Aug 2015 – May 2019

RESEARCH EXPERIENCE

Berkeley Sky Computing Lab

Undergraduate Research Assistant

Berkeley, CA

Sep 2022 – Present

- Contributed to Alpa, a system for automatically parallelizing the training and serving of large-scale neural networks.
- Implemented swapping between GPU and main memories in the Alpa runtime to enable deployment of large models on clusters with limited GPU memory resources.
- Supervised by Lianmin Zheng and Professor Joesph Gonzalez.

Berkeley RISE Lab

Undergraduate Research Assistant

Berkeley, CA

Mar 2021 – Present

- Contributed to NumS, a high performance distributed numerical library for Python with a NumPy-like interface, with support for Ray and Dask backends.
- Studied and innovated on techniques like operator fusion and cost estimation to dynamically schedule sparse-aware tensor operations on multi-node systems.
- Implemented distributed algorithms such as quickselect for block-partitioned arrays, then analyzed the scaling properties of said algorithms through benchmarks on an EC2 cluster.
- Supervised by Melih Elibol, PhD and Professor Ion Stoica.

WORK EXPERIENCE

NVIDIA Corporation

Autonomous Vehicles Software Infrastructure Intern

Santa Clara, CA

May 2022 – Aug 2022

- Designed and built on-call automation and metrics collection tooling from the ground up, for use by NVIDIA's AV infrastructure organization.
- Architected a concurrent, event-driven service with WebSockets, producer-consumer queue, in-memory cache, SQL database, and a querying command interface.
- Worked with multiple stakeholder representatives to revise features and improve usability of the tooling during development and beta testing. Presented the product to 50+ engineers and managers, then engaged with discussions of expanding product scope to other organizations.
- Deployed for use by 300+ engineers and business partners, and growing.

ORGANIZATIONS

Computer Science Undergraduate Association

VP of Technology & Root Staff

Berkeley, CA

Sep 2020 – Aug 2022

- Administered a GPU cluster, a web hosting server, and other free computing services used by 400+ members.
- Maintained a full stack web application written in Django.
- Configured a multi-instance Postfix mail server.
- Served as the leader of CSUA Root Staff for a semester upon election to VP of Technology. Acted as liaison with campus IT staff to resolve issues regarding networking services.

Computer Science Mentors

Junior Mentor

Berkeley, CA

Jan 2021 – May 2021

- Taught CS 61C: Great Ideas in Computer Architecture, which covers C, RISC-V assembly, CPU logic design, memory hierarchy, and parallel programming.
- Mentored a group of student every week by giving short lectures and practice problem walk-throughs.

FIRST Robotics Team 1477 (Texas Torque)

The Woodlands, TX

Lead Programmer

Jun 2018 – May 2019

- Recruited a team of 6 programmers and oversaw the development of the robot control system.
- Designed new autonomous control framework that enables scheduling actions for multiple subsystems in parallel. Built a custom object tracking CV pipeline that runs on a coprocessor.
- Reached the quarterfinal of the 2019 FRC World Championship as a team.

SELECTED PROJECTS

PintOS

- An instructional OS written in C and x86. Completed in CS 162: Operation Systems and Systems Programming over the course of a semester.
- Implemented basic user processes, kernel threads, concurrency primitives, and concurrent filesystem.

NP-Hard Graph Problem Solver | [GitHub](#)

- An open-ended project to find optimal solutions for a set of inputs to a given NP-hard graph problem. Completed in CS 170: Efficient Algorithms and Intractable Problems.
- Designed a greedy algorithm as a baseline and a simulated annealing algorithm to explore the space of potentially better solutions. Wrote scripts to merge best solutions over multiple runs.

NumC: Toy Reimplementation of Core NumPy Functionality

- Completed in CS 61C: Great Ideas in Computer Architecture.
- Lightweight numerical library for Python, with underlying data structures written in C for performance.
- Support efficient matrix slicing and common operations like add, abs, matmul, matpow. Optimized matmul with threading, loop tiling, and SIMD intrinsics. Optimized matpow with dynamic programming.

Deep NLP Chatbot | [Blog](#)

- Self-studied RNNs and its variants, as well as the TensorFlow ML framework.
- Implemented an attention-based seq2seq model that uses a beam search decoder for evaluation.

AWARDS & ACHIEVEMENTS

National Merit Finalist: Selected among high school seniors scoring in the top 0.5% of the PSAT.

American Invitational Mathematics Exam Qualifier: Scored in the top 5% of students taking AMC 10 & 12.

SKILLS

Programming: Python, C, Go, Java, OCaml, Lua, SQL, RISC-V, x86

Libraries: SciPy, PyTorch, JAX, Ray, Spark, Django

Technologies: Linux, Git, Jira, Docker, AWS EC2

Other: native English and Mandarin Chinese, technical and blog writing

SELECTED COURSEWORK

CS 61C: Great Ideas in Computer Architecture

CS 161: Computer Security

CS 162: Operating Systems and Systems Programming

CS 164: Programming Languages and Compilers

CS 170: Efficient Algorithms and Intractable Problems

CS 182: Deep Neural Networks

CS 186: Introduction to Database Systems

CS 189: Introduction to Machine Learning

ENGIN 185: The Art of STEM Communication

Please refer to my transcript for a full record of my technical and non-technical coursework.