Zihan Zhang

GitHub: SubSir tiancaizhangdaxian@sjtu.edu.cn

EDUCATION

Shanghai Jiao Tong University, Shanghai, China 09/2023 – 06/2027 (Expected)
B.Eng. in Computer Science & Technology
Member of ACM Honors Class, top 5% students
GPA (All): 3.91/4.3, Ranking: 8/27
Selected courses:

- Compiler Design and Implementation 94/100
- Computer Systems 92/100
- Principle and Practice of Computer Algorithms 93/100
- Modern Algebra 97/100
- Linear Algebra 97/100

RESEARCH INTERESTS

My research interests lie in efficient machine learning, with a particular focus on model compression. My current projects specifically explore quantization techniques for large language models. In the long term, my goal is to develop and deploy lightweight models that can deliver high performance at minimal cost, making advanced AI systems more accessible, sustainable, and widely applicable across real-world scenarios.

RESEARCH EXPERIENCE

Shanghai Jiao Tong University, EPCC Lab 06/2023 –
Advisors: Prof. Jingwen Leng

- Algorithm-Hardware Co-design for large language model (under review at ASPLOS 2026).

PERSONAL PROJECTS

rCore

A RISC-V operating system kernel implemented in Rust to explore core OS concepts. Features include Sv39 virtual memory with page tables, a two-tier heap allocator (Buddy + Slab), and a POSIX-like Virtual File System (VFS) with block-based storage. Implements process scheduling, user/kernel isolation, and hardware abstraction with MMU support. Optimized via block caching, efficient address translation, and memory protection mechanisms.

Mx Compiler & JITC

Developed a full compiler for Mx* (a C-like educational language) in Python using ANTLR, targeting RISC-V via LLVM IR. Supports OOP features (classes, inheritance, dynamic allocation), f-string interpolation, and robust semantic analysis with type checking and scope management. Complemented by JITC, a Just-In-Time compiler that parses and executes simplified LLVM IR using ANTLR-generated lexer/parser, supporting 28 IR constructs and both listener/visitor traversal patterns.

RISCV-CPU

A synthesizable RISC-V CPU in Verilog featuring out-of-order execution via Tomasulo's algorithm, including reservation stations, load/store buffers, and a reorder buffer. Integrates a memory hierarchy with instruction cache, 128KB RAM, and UART-based I/O for host communication. Successfully deployed on a Basys3 FPGA with support for precise exceptions, branch prediction recovery, and full test infrastructure.

Raft Consensus Implementation

Implemented the Raft consensus algorithm as part of MIT 6.824, achieving fault-tolerant replicated state machines with leader election, log replication, and safety guarantees. Designed to handle network partitions, server crashes, and log consistency under concurrent operations. Repository is private per course policy.

TEACHING EXPERIENCE

Data Structure (02/2025 – 06/2025)
Programming (09/2024 – 02/2025)

HONORS

2024–2025 Zhiyuan Honorary Scholarship
2023–2024 Zhiyuan Honorary Scholarship

TECHNICAL SKILLS

Programming: Python, C/C++, CUDA, Go, Verilog
Tools: Git, Markdown, LaTeX