Aarush Goradia

aarush.goradia@princeton.edu | linkedin.com/in/aarushgoradia

EDUCATION

Princeton University | Princeton, NJ | May 2027
B.S.E. in Electrical and Computer Engineering; Minors in Computer Science and English | GPA: 3.70
Relevant Coursework: Advanced Computer Architecture, VLSI, Contemporary Logic Design, Advanced Circuit Design, Programming Systems, Parallel Computing, Machine Learning, Natural Language Processing

EXPERIENCE

Digital Verification Intern | Qualcomm | San Diego, CA | Jun 2026 - Aug 2026
- Architected SystemVerilog Assertions (SVA) and executed formal verification in Cadence JasperGold to mathematically prove corner-case correctness for proprietary Network-on-Chip (NoC) power-management and clock-gating features.
- Built UVM verification components and scoreboard checkers in SystemVerilog to validate dynamic power-state transitions and low-power protocol handshake logic across NoC interconnect nodes.
- Built Python and multi-agent LLM tools for NoC latency profiling, CDC checks, and bug triage, adopted across NoC, LLCC, and Memory Controller teams — cutting debug cycles from weeks to hours.
- Received full-time return offer.

Technology Development Intern | Advanced Micro Foundry | Singapore | Jun 2025 - Jul 2025
- Automated GDS fabrication checks and integrated KLayout DRC into tape-out workflows, reducing preliminary review from hours to minutes.
- Rewrote core GDS analysis routines in C++ for 15-20x faster execution than prior Python implementations.

RESEARCH & PROJECTS

Senior Thesis | Princeton University | Advisor: David Wentzlaff | Sep 2026 - Present
- Building a three-node off-the-shelf system for local inference on trillion-parameter sparse mixture-of-experts models, storing the full model in NVMe NAND and staging active experts through DRAM into GPU memory.
- Built a bandwidth model of the NAND, DRAM, PCIe, and inter-node data path, showing sequential decode is bandwidth-bound at ~1.5 tokens/s on the hardware.
- Modifying llama.cpp (C/C++) with expert caching, prefetching, and compute–I/O overlap, targeting over 10 tokens/s; evaluating token latency, throughput, cache-hit rate, and storage/network bandwidth against out-of-core baselines.

Researcher | Malik Research Group, Princeton University | Jul 2025 - Present
- Architected an RTL-SystemVerilog Assertion dataset using GPT-5 generation and JasperGold formal verification to filter synthesized and real assertion pairs for correctness.
- Fine-tuned Qwen2.5-Coder-7B-Instruct with 4-bit QLoRA; the proof-filtered model achieved a 50.2% formal-verification yield versus 41.4% for GPT-5 with approximately 200x fewer parameters.
- Co-first-authored the Veri2 paper (under submission) and evaluated four data-quality conditions on a held-out 578-design test set, showing proof-filtered data outperformed larger syntax-filtered and unfiltered datasets.

Processor Design & Branch Prediction | Nov 2024 - May 2025
- Integrated bimodal, gShare, and TAGE-lite branch predictors into a five-stage in-order RISC-V CPU and evaluated them on seven benchmarks; gShare cut nested-loop mispredictions from 90.6% to 3.8% versus predict-not-taken.
- Analyzed how decode-stage prediction's one-cycle taken-branch redirect penalty limited IPC gains despite higher prediction accuracy.

Researcher | Sengupta Lab, Princeton University | Aug 2024 - May 2025
- Automated transformer layout simulation in Cadence Virtuoso/EMX and trained a neural network to predict transformer geometry from lumped parameters with 90% accuracy.

SKILLS

Hardware: SystemVerilog, Verilog, RTL design, functional verification, testbenches, computer architecture, FPGA prototyping
Programming: Python, C++, C, Java, Linux shell
Tools: PyTorch, NumPy, Hugging Face Transformers/PEFT, llama.cpp, JasperGold, AMD Vivado, Cadence Virtuoso/EMX, KLayout
Interests: PC building, robotics, strategy games, film, and literature