Christos Vasileiou

: chrivasileiou2@gmail.com | : christos-vasileiou | : chrivasileiou | : github.io

Summary

Ph.D. researcher in Computer Engineering at UC San Diego working at the intersection of **LLM reinforcement learning and chip design/test**, building **machine learning applications in hardware manufacturing, design and testing.** Enthusiastic and research-driven engineer, specializing in designing, developing, training, and refining complex GenAI and deep learning solutions for digital twin technologies, with a proven ability to tackle complex problems with innovative approaches. **Highly motivated to advance expertise and make impactful contributions in the fields of applications of AI and Deep Learning.**

Education

PhD in **Computer Engineering** — Sep 2025 - Aug 2027
University of California at San Diego, UCSD

Master's in **Computer Engineering** — Jan 2021 - May 2023
University of Texas at Dallas, UTD

Integrated Master's (B.S. & M.S.) in **Computer Engineering** — Sep 2014 - Aug 2020
University of Thessaly at Volos, Greece

Selected Research

LLM Reinforcement Learning for Chip Design and Test

- Built an EDA-grounded LLM/RL environment spanning RTL Verilog, Synopsys Design Compiler synthesis, Formality equivalence checking, ASAP7 gate-level netlists, TetraMAX ATPG/fault simulation, SFT and GDPO training and evaluation.
- Constructed a simulator-verified corpus of approximately 1.09M fault-conditioned records across 2,870 synthesized netlists, combining commercial ATPG labels with an independent Liberty-aware gate-level fault simulator.
- Trained with Group reward-Decoupled Normalization Policy Optimization (GDPO), a GRPO variant that normalizes each reward component separately, using simulator-driven rewards that parse model-generated actions, execute proposed test vectors against the circuit, and use actual fault activation, propagation, output observability, and prediction consistency as reinforcement signals.
- Implemented a custom dual-adapter GDPO trainer that maintains a frozen SFT/reference LoRA alongside a trainable policy adapter on a quantized backbone, avoiding lossy adapter merging while supporting multi-turn tool interaction.
- Scaled SFT/GDPO experiments with distributed GPU training and dedicated vLLM serving, and developed verifier-guided inference strategies including best-of-N, PUCT/MCTS, and evolutionary search.
- https://github.com/christos-vasileiou/atpgllm/tree/develop
- https://github.com/christos-vasileiou/asap7-language-of-test

Professional Experience

TCAD Software Research Intern — May 2025 - August 2025
Intel Corp.

- Improved transistor–level placer with a routing–aware recommendation system, achieving a 12% increase in routable placements over baseline.
- Grounded the approach in Bayes' theorem by modeling the recommender as a bipartite Graph Neural Network (GNN) and training with Bayesian Personalized Ranking (BPR) loss.

TCAD Software Research Intern — May 2024 - August 2024
Intel Corp.

- Developed GPU-based neural models for semiconductor thermal modeling, improving predictive accuracy by approximately **10%** while reducing model parameters by **50%**.
- Improved yield-modeling data infrastructure by mapping manufacturing diagnosis results to physical **layout defect segments**, enabling ML-driven analysis of test and defect data.

Pathfinding Test Developer Intern — May 2022 - August 2022
Intel Corp.

- Employed Machine Learning techniques and statistical analysis for the detection and identification of manufacturing defects within Integrated Circuits (ICs).

– Innovated machine learning methodologies to discern equivalence relationships, significantly minimizing the required test points for stimuli in manufacturing processes and thereby substantially decreasing computation time.

Asynchronous Hardware Designer Intern June 2018 - September 2018
Circuit And Systems Laboratory

– Implementation of an Asynchronous Blade/Sharp (Resilient) handshake on FPGA as an intern at the CAS (Circuits and Systems) Lab.

– A project related to clockless control circuit implementation of a resilient circuit, capable of detecting Flip-Flop setup errors and mitigating the timing.

– Received the Sterling medal (1st place in 2018 internships).

Publications

Language of Test **Manuscript in preparation**

– Built an EDA-grounded stack for an AI-powered application of LLM/RL environment spanning RTL Verilog, Synopsys Design Compiler synthesis, Formality equivalence checking, ASAP7 gate-level netlists, TetraMAX ATPG/fault simulation, and SFT/GDPO training and evaluation.

– Developed a fault-simulation-driven generative test synthesis methodology combining Supervised Fine-Tuning (SFT) and Group reward-Decoupled Normalization Policy Optimization (GDPO), a variant of Group Relative Policy Optimization (GRPO), infrastructure for tool-using LLMs, including simulator-verified rewards, structured-output parsing, optimized netlist caching, evaluation metrics, scalable batch EDA execution, with inference-time search to improve targeted stuck-at fault detection.

Teaching Llamas to Test: A Language-Based Approach **ITC '25**

– Pioneered an LLM-based ATPG paradigm that "learns the language of test," enabling a fine-tuned 7 B-parameter Llama-2 trained on synthetic combinational netlists to generate stuck-at-fault vectors that detect $> 80\%$ of target faults on unseen gate-level netlists (pass@4)

– Devised a two-stage pipeline domain-specific supervised fine-tuning followed by GRPO reinforcement learning with fault-simulation rewards

– Lifted single-shot fault-detection success from 41.8% to 61.6%, demonstrating the feasibility of data-efficient, explainable generative test synthesis.

Unveiling the Mask **ICCAD '25**

– Introduced the first wafer-level mask-set attestation framework that mines standard inline and e-test data; lightweight SVC models with NSGA-II feature selection achieved 94%+ lot-level accuracy (100% with four-layer changes) across 8K GF12LP FinFET wafers, flagging malicious or replica masks without foundry trust.

– Proposed one-class and multi-class attestation modes that exploit systematic process-variation "fingerprints," while offering domain-expert interpretability of critical predictors (e.g., IDB, Ieff), paving a practical path toward secure, ML-assisted semiconductor supply chains.

Efficient CNN-Based Super Resolution Algorithms for Mmwave Mobile Radar Imaging **ICIP '22**

– An innovative super resolution approach to emerging modes of near-field synthetic aperture radar (SAR) imaging.

– Developed a Generative Adversarial Network (GAN) by removing distortion and enhancing clarity, shape & texture of data by 50% compared to SOTA techniques while reducing time consumption by $\sim$1000x.

– Developed for embedded systems and mobile devices reducing by 90% number of parameters while enhancing the performance.

Technical Skills

- **Languages:** Python, C/C++, Rust, Bash, Git/GitHub, Slurm, Tcl, CUDA
- **RL and ML systems:** GRPO, GDPO, SFT, PyTorch, Tensorflow, TRL, Hugging Face, DeepSpeed, Accelerate, Distributed Training, vLLM, Weights&Bias
- **Familiarity with Concepts:** CNNs, NLP Transformers, Diffusion/Vision Transformers
- **EDA / Hardware:** Verilog, Yosys, Synopsys Design Compiler, Formality, TetraMAX, ATPG, Fault Simulation, RTL-to-GDS flow, HLS, Static Timing Analysis
- **Tools:** Cursor, VSCode, Git/Github, Google Colab, Jupyter Notebook, NotebookLM, OneNote