

PU WANG

GitHub ◊ Phone: (+86) 131-9455-8889 ◊ Email: puwang0508@gmail.com

EDUCATION

College of Computer Science and Technology, Zhejiang University

Hangzhou, China

B.E. in Artificial Intelligence (Turing Class), with an honors degree from Chu Kochen Honors College

Sept. 2023 – Jun. 2027 (expected)

GPA: 3.87/4.0 (88.78/100)

Related courses: Introduction to Theoretical Computer Science (91), Introduction to Natural Language Processing (92), Machine Learning (96), Computer Systems I/II/III (89/94/97), Advanced Data Structure and Algorithm Analysis (91), Mathematical Analysis I/II (95/89)

RESEARCH INTERESTS

My research interests lie in the theory and algorithms of **sequential decision-making**, with a growing focus on **large language models** and **language agents**. I am interested in the algorithmic foundations of how language models and agents *reason, plan, optimize behavior, and make decisions* in interactive environments. More broadly, I aim to connect ideas from reinforcement learning, imitation learning and preference learning to modern problems in *LLM post-training, RLHF, agentic AI, and alignment*.

PUBLICATIONS & SUBMITTED MANUSCRIPTS

Tree-Guided Identify-Then-Exploit: A Unified Framework of Best Arm Identification and Regret Minimization for Dueling Bandits.

Pu Wang, Yao-Xiang Ding

Preprint: arXiv:2606.01799.

RESEARCH EXPERIENCE

Research Intern at State Key Lab of CAD&CG

Mar 2025 – Present

Advisor: Prof. Yao-Xiang Ding

Zhejiang University

• **TG-ITE: A Unified Framework for Dueling Bandits**

Mar 2025 – May 2026

- Proposed **Tree-Guided Identify-Then-Exploit (TG-ITE)**, the first unified framework for N -armed dueling bandits jointly handling best arm identification (BAI), weak regret, and strong regret under only the Condorcet-winner assumption.
- Built on a shared identification primitive TREEASCENT that exploits a tree-based tournament decomposition to locate the Condorcet winner, paired with objective-specific exploitation strategies tailored to each goal, the framework achieves bounds that closely match the state-of-the-art.

• **Imitation Learning from Partially Observable Experts**

Jun 2025 – Dec 2025

- Formulated a new IL setting in which a full-state learner mimics M heterogeneous experts, each observing only a partial mask of the state—modeled as expert-specific masked POMDP views of a shared underlying MDP.
- Proved a **negative identifiability result**: expert-mask coverage alone is insufficient—without further structure, neither the latent causal mask nor the full-information optimal policy can be recovered from naïve demonstration pooling.
- Proposed **POIL**, a GAIL-based framework alternating visibility-aware discrimination and active querying to recover the causal mask; preliminary experiments on MuJoCo locomotion (Walker2d, HalfCheetah).

PROJECTS

Five-Stage Pipelined RISC-V CPU & xv6-Inspired OS

Mar 2024 – May 2025

Course Project for Computer Systems I-III

Zhejiang University

- **Hardware Architecture:** Designed and implemented a five-stage pipelined RISC-V CPU in SystemVerilog, featuring BTB/BHT-based branch prediction, a 2 KiB cache, and an MMU supporting Sv39 virtual memory.
- **Kernel Development:** Brought up a minimalist xv6-inspired operating system on the custom CPU, implementing page-table-based virtual memory, trap/interrupt handling and user-to-privileged mode transitions, and a FAT32-style mini file system.

Needle: A Compiler from C to RISC-V

Mar 2025 – Jun 2025

Course Project for Principles of Compilers

Zhejiang University

- **Frontend & IR:** Implemented an end-to-end compiler for a C subset using *Flex* and *Bison*, featuring AST construction, semantic analysis, and Intermediate Representation (IR) generation.
- **Code Generation:** Developed a RISC-V code generator handling calling conventions and control flow, verifying correctness through simulation tools.

ACHIEVEMENTS

Zhejiang University Scholarship – Second Prize

2024, 2025

Second Class Scholarship for Elite Students in Basic Sciences

2025

SKILLS

Programming Languages	Python, C/C++, SystemVerilog
Machine Learning Tools	PyTorch, Gymnasium, Stable-Baselines3
Systems & Infra	Linux, Git, Shell Scripting, Docker, Slurm
Writing & Typesetting	L ^A T _E X, Typst, TikZ

I maintain an open technical notebook (400,000+ words) on computer science, mathematics, and AI, as a long-term knowledge-building practice.