Wei Fu

Email: fuwth17@gmail.com Phone: +86 18801030958 Location: Beijing, CN

Website: https://garrett4wade.github.io/

EDUCATION

Ph.D. in Institute for Interdisciplinary Information Sciences (IIIS) Tsinghua University
Beijing, China | 2021.9 – 2026.7

Advisor: Yi Wu

Research: RL, MARL, Distributed Systems

B.E. in Department of Electronic Engineering Tsinghua University
Beijing, China | 2017.9 – 2021.7

Advisor: Yi Wu

Thesis: Distributed Systems for Multi-Agent Games

High School Yaohua High School
Tianjin, China | 2014.9 – 2017.7

PUBLICATIONS

Jiaxuan Gao, Wei Fu, Minyang Xie, Shusheng Xu, Chuyi He, Zhiyu Mei, Banghua Zhu, Yi Wu. "Beyond Ten Turns: Unlocking Long-Horizon Agentic Search with Large-Scale Asynchronous RL". Arxiv Preprint, 2025.

Wei Fu, Jiaxuan Gao, Xujie Shen, Chen Zhu, Zhiyu Mei, Chuyi He, Shusheng Xu, Guo Wei, Jun Mei, Jiashu Wang, Tongkai Yang, Binhang Yuan, Yi Wu. "AReaL: A Large-Scale Asynchronous Reinforcement Learning System for Language Reasoning". Arxiv Preprint, 2025.

Jiaxuan Gao, Shu Yan, Qixin Tan, Lu Yang, Shusheng Xu, Wei Fu, Zhiyu Mei, Kaifeng Lyu, Yi Wu. "How Far Are We from Optimal Reasoning Efficiency?". Arxiv Preprint, 2025.

Zhiyu Mei*, Wei Fu*, Kaiwei Li, Guangju Wang, Huanchen Zhang, Yi Wu. "ReaL: Efficient RLHF Training of Large Language Models with Parameter Reallocation". MLSys 2025, 2025. ((*: Equal Contribution)).

Jiaxuan Gao, Shusheng Xu, Wenjie Ye, Weilin Liu, Chuyi He, Wei Fu, Zhiyu Mei, Guangju Wang, Yi Wu. "On Designing Effective RL Reward at Training Time for LLM Reasoning". Arxiv Preprint, 2024.

Shusheng Xu, Wei Fu, Jiaxuan Gao, Wenjie Ye, Weilin Liu, Zhiyu Mei, Guangju Wang, Chao Yu, Yi Wu. "Is DPO Superior to PPO for LLM Alignment? A Comprehensive Study". ICML 2024, 2024. (Oral, 1.5%).

Zhiyu Mei*, Wei Fu*, Guangju Wang, Huanchen Zhang, Yi Wu. "SRL: Scaling Distributed Reinforcement Learning to Over Ten Thousand Cores". ICLR 2024, 2024. ((*: Equal Contribution) ES-FoMo Workshop Oral).

Yunfei Li, Jinhan Li, Wei Fu, Yi Wu. "Learning Agile Bipedal Motions on a Quadrupedal Robot". ICRA 2024, 2024. (EXPO Best Paper Final List).

Wei Fu, Weihua Du, Jingwei Li, Sunli Chen, Jingzhao Zhang, Yi Wu. "Iteratively Learn Diverse Strategies with State Distance Information". NeurIPS 2023, 2023.

Wei Fu, Chao Yu, Zelai Xu, Jiaqi Yang, Yi Wu. "Revisiting Some Common Practices in Cooperative Multi-Agent Reinforcement Learning". ICML 2022, 2022. (Spotlight Talk).

Zihan Zhou*, Wei Fu*, Bingliang Zhang, Yi Wu. "Continuously Discovering Novel Strategies via Reward-Switching Policy Optimization". ICLR 2022, 2022. ((*: Equal Contribution)).

Wei Fu, Chao Yu, Yunfei Li, Yi Wu. "Unlocking the Potential of MAPPO with Asynchronous Optimization". CICAI 2021, 2021. (Oral).

EXPERIENCE

Research Intern Ant Research, RL Lab
2025.06 – 2025.08

We built **ASearcher**, a search agent trained with **AReaL**. Our agent exhibits extreme long-horizon search, with tool calls exceeding 40 turns and output tokens exceeding 150k during training time. With a simple agent design and no external LLMs, ASearcher achieves Avg@4 scores of 51.1 on xBench and 58.7 on GAIA, surpassing existing open-source 32B agents.

Research Intern Ant Research, RL Lab
2024.12 – 2025.06

I lead the **AReaL** project, an asynchronous RL system for reasoning and agentic LLM training. It delivers the simplest solution for customizing agentic workflows, while reducing overall training time by 2.5× compared to the best open-source synchronous systems.

Research Intern Shanghai Qi Zhi Institute
2023.5 – 2024.10

We designed **ReaL**, an efficient distributed system for LLM RLHF. It exhibits an average 26% improvement over heuristic approaches based on Megatron-LM, and at least 2× higher throughput than DeepSpeed-Chat.

Research Intern Shanghai Qi Zhi Institute
2021.10 – 2023.5

We implemented **SRL**, a general-purpose game-centric RL system that can scale to over 10k CPU cores and 96 GPUs. We were the first to reproduce OpenAI's results in the Hide-And-Seek game.

Research Intern ByteDance
2020.6 – 2020.9

We implemented a Ray-based system (not RLlib) to train an agent to play Unity-based FPS games. While the baseline required several days for training, our system could finish training within approximately 8 hours.

AWARDS & HONORS

ByteDance Scholarship 2024

Scholarships in Tsinghua University 2017 – 2024

SERVICE

Reviewer NeurIPS/ICML/ICLR 2022 – 2024