Jianing Ye

jianing.y@wustl.edu · heavycrab.github.io

Research Interests

Multi-Agent Coordination, LLM Agents, Reinforcement Learning, Robotics

Education

Washington University in St. Louis (WashU), St. Louis, US 2024 – Present

PhD student in Computer Science. Advisor: Chongjie Zhang.
Affiliation: Computer Science & Engineering, McKelvey School of Engineering.
Research Interests: Reinforcement Learning, Multi-agent Coordination, LLM Agents, Robotics

Tsinghua University (THU), Beijing, China 2021 – 2024

Master's student in Computer Science and Technology. Advisor: Chongjie Zhang, Yi Wu.
Affiliation: Institute for Interdisciplinary Information Sciences (IIIS).
Research Interests: Reinforcement Learning, Multi-agent Coordination

Peking University (PKU), Beijing, China 2017 – 2021

B.S. and *B.S. (Hons)* in Computer Science and Technology.
Affiliation: Turing Class (CS area), School of Electronics Engineering and Computer Science (EECS).

Shaoxing NO.1 High School, Zhejiang, China 2014 – 2017

Publications

(* indicates equal contribution)

1. Buzhao Liu*, **Jianing Ye***, Xiu Yuan*, Yangtong Feng, Chengsong Huang, Tyler Dinh, Yevgeniy Vorobeychik, Ning Zhang, Chongjie Zhang. RoboAuto: Autonomous Robot Program Search through Multi-Level Physical Understanding. *arXiv preprint arXiv: 2610.36533*, 2026.
2. **Jianing Ye**, Yang Zhang, Xiu Yuan, Buzhao Liu, David Park, Qihan Liu, Chongjie Zhang. MATCHA: Bridging the Gap between Sample and Wall-Clock Efficiency in Model-Based Multi-Agent Reinforcement Learning. *Under review*, 2026.
3. Dingyang Chen, **Jianing Ye**, Zhenyu Zhang, Xiaolong Kuang, Xinyang Shen, Ozalp Ozer, Chongjie Zhang, Qi Zhang. Correlated Policy Optimization in Multi-Agent Subteams. *International Conference on Learning Representations (ICLR)*, 2026.
4. Yang Zhang, Xinran Li, **Jianing Ye**, Shuang Qiu, Delin Qu, Xiu Li, Chongjie Zhang, Chenjia Bai. Revisiting multi-agent world modeling from a diffusion-inspired perspective. *Advances in Neural Information Processing Systems (NeurIPS)*, 2025.
5. Hao Hu*, Yiqin Yang*, **Jianing Ye**, Chengjie Wu, Ziqing Mai, Yujing Hu, Tangjie Lv, Changjie Fan, Qianchuan Zhao, Chongjie Zhang. Bayesian Design Principles for Offline-to-Online Reinforcement Learning. *International Conference on Machine Learning (ICML)*, 2024.
6. Qihan Liu*, **Jianing Ye***, Xiaoteng Ma*, Jun Yang, Bin Liang, Chongjie Zhang. Efficient Multi-agent Reinforcement Learning by Planning. *International Conference on Learning Representations (ICLR)*, 2024.
7. Hao Hu*, Yiqin Yang*, **Jianing Ye**, Ziqing Mai, Chongjie Zhang. Unsupervised Behavior Extraction via Random Intent Priors. *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.
8. **Jianing Ye**, Chenghao Li, Yongqiang Dou, Jianhao Wang, Guangwen Yang, Chongjie Zhang. Towards Global Optimality in Cooperative MARL with the Transformation And Distillation Framework. *arXiv preprint* arXiv:2207.11143v4, 2022.
9. Hao Hu, **Jianing Ye**, Zhizhou Ren, Guangxiang Zhu, and Chongjie Zhang. Generalizable Episodic Memory for Deep Reinforcement Learning. *International Conference on Machine Learning (ICML)*, 2021.

10. Jianhao Wang*, Zhizhou Ren*, Beining Han, **Jianing Ye** and Chongjie Zhang. Towards Understanding Linear Value Decomposition in Cooperative Multi-Agent Q-Learning. *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.

Experience

CSE, Washington University in St. Louis 2024.9 – Present

PhD student Advisor: Chongjie Zhang

Research Outline:

- **LLM Agents for Robot Program Synthesis**
  - Co-led RoboAuto, an LLM-agent framework that autonomously searches for robot programs (Code-as-Policy) grounded in multi-level physical understanding at the object, scene, and interaction levels.
  - Built a hierarchical knowledge base of object-centric physical knowledge learned from demonstration videos, physical scene grounding via point-cloud measurement, and multi-stage planning in which the agent designs its own experiments to resolve physical uncertainty.
  - Achieved state-of-the-art success rates across five benchmark families (e.g., 73.4% vs. 30.8% over ASPIRE on LIBERO-PRO Long), and outperformed ASPIRE and $\pi_{0.5}$ on long-horizon real-robot tasks with a WidowX arm.
- **Efficient Model-Based Multi-Agent Reinforcement Learning**
  - Led MATCHA, a system–algorithm co-design that makes search-based model-based MARL efficient on modern accelerators, achieving up to **16.3**× end-to-end wall-clock speedup over MAZero while matching or exceeding the sample efficiency of strong model-based baselines on SMAX, Overcooked, and MPE.
  - Designed Path-Guided Frontier Search (PGFS), which replaces sequential MCTS tree traversal with one-shot vectorized sampling over the search frontier, reducing worst-case search complexity from $O(N^2)$ to $O(N \log N)$ with a provable policy improvement guarantee.
  - Built a device-centric, end-to-end vectorized training framework that keeps data collection, planning, reanalysis, and learning on the GPU, and scales to multiple GPUs without lowering the update-to-data ratio.
  - Collaborated on diffusion-inspired multi-agent world modeling (NeurIPS 2025) and correlated policy optimization for multi-agent subteams (ICLR 2026).

IIIS, Tsinghua University 2021.9 – 2024.6

Master student Advisor: Chongjie Zhang, Yi Wu

Research Outline:

- **Effective and Efficient Multi-agent Reinforcement Learning**
  - Analyzed the convergence properties for Linear Value Decomposition and IGM Value Decomposition under the framework of Factorized Multi-Agent Fitted Q-Iteration (FMA-FQI).
  - Analyzed the suboptimality of existing Multi-Agent Policy Gradient and Value-Decomposition MARL algorithms when using gradient descent optimizers.
  - Proposed the Transformation And Distillation (TAD) framework to achieve global optimality in cooperative MARL, and designed a novel cooperative MARL algorithm, TAD-PPO, which outperforms the state-of-the-art algorithms in various cooperative MARL benchmarks.
  - Proposed the Model-based Multi-agent Reinforcement Learning algorithm MAZero, which learns a multi-agent world model and does MCTS planning to achieve the SOTA sample efficiency in cooperative MARL.
- **Offline-to-Online Reinforcement Learning**

- Proposed UBER, an unsupervised approach to extract useful behaviors from offline, reward-free datasets via diversified rewards.
- Proposed a Bayesian design principle for offline-to-online fine-tuning to avoid the inefficiency or unstability caused by overpessimistic or overoptimistic fine-tuning policy.

- **Episodic Memory in Deep RL**
  - Presented a novel framework, GEM, to build generalizable episodic memory and improve sample efficiency. GEM consists of a novel value planning scheme for the continuous domain and a twin back-propagation process to reduce overestimation along the trajectories.

Teaching Assistant Courses:

- **Research Practice**
  Instructor: Chongjie Zhang *2021.9 – 2022.1*
- **Diploma Project (Thesis)**
  Instructor: Chongjie Zhang *2022.2 – 2022.6*

Turing Class, Peking University 2017 – 2021

Undergraduate student

The ICPC International Collegiate Programming Contest: I attended the ICPC contest for several years under the supervision of Guojie Luo. I took part in 5 regional contests, 2 continent finals and all got gold medals. In 2019, I attended world finals in Portugal on behalf of Peking University and got the 13th place.

Teaching Assistant Courses:

- **Introduction to the Theory of Computation**
  Instructor: Xiaoming Sun *2020.2 – 2020.6*

TCS book reading seminar: I organized a weekly seminar of theoretical computer science. We read books like **Computational Complexity: A Modern Approach** in this seminar.

Institute of Computing Technology Chinese Academy of Science 2019.8 – 2020.6

Research Intern Mentor: Xiaoming Sun

Analysing the minimum representation of some special quantum gates.

Beijing Megvii Co., Ltd. 2017.3.1 – 2017.8.1

Research Intern Mentor: Haoqiang Fan

Liveness test algorithm design, 3D Facial hair reconstruction, problem design for MegCup contest.

- Built and programmed a mechanical arm for data automatic acquisition under various environment.
- Designed an algorithm to extract fine 2D mask for eyebrows.
- Set the final problem for MegCup2017.

Skills

- algorithm design, data structure, multi-agent reinforcement learning

Honors and Awards

Excellent Graduate of Peking University	2021
National Scholarship	2018
Peking University Three Good Students Model	2018
Peking University Freshman Scholarship	2017
Peking University Third Class Scholarship	2019
13th place, Award on ICPC World Finals	2019
Gold Medal, Award on ACM-ICPC East Asia Continent Final (EC-Final)	2018
Gold Medal, Award on ACM-ICPC Xuzhou regional	2018
Gold Medal, Award on ACM-ICPC Qingdao regional	2018
Gold Medal, Award on ACM-ICPC East Asia Continent Final (EC-Final)	2017
Gold Medal, Award on ACM-ICPC Xi'an regional	2017
Gold Medal, Award on ACM-ICPC Qingdao regional	2017
Gold Medal, Award on ACM-ICPC Yinchuan regional	2021
Gold Medal, National Training Team, Award on NOI	2016
Best School Team, Award on SNACKDOWN	2017
2nd place, Award on SenseTime Ace Coder Challenge	2018