

# WANPENG ZHANG (张万鹏)

School of Computer Science, Peking University, Beijing, China

✉ zawnpn@gmail.com 🔗 zhangwp.com

## Intro

---

I am a Ph.D. Candidate in the School of Computer Science at Peking University, advised by Professor Zongqing Lu.

My research interests are: *Reinforcement Learning, Generative Modeling, and Multimodal LLMs.*

## Education

---

### Peking University

Sep. 2022 – Now

*Ph.D. Candidate, School of Computer Science*

*Supervisor: Zongqing Lu*

### Tsinghua University

Sep. 2019 – Jun. 2022

*Master of Science Degree, Department of Computer Science and Technology*

*Supervisor: Xi Xiao*

### Nankai University

Sep. 2015 – Jun. 2019

*Bachelor of Science Degree, School of Mathematical Sciences*

## Experience

---

### Beijing Academy of Artificial Intelligence (BAAI)

May. 2024 – Now

*Research Intern*

*Beijing, China*

### Tencent AI Lab

Jun. 2020 – Jul. 2021

*Research Intern*

*Shenzhen, China*

## Representative Work

---

**Tackling Non-Stationarity in Reinforcement Learning via Causal-Origin Representation.** (ICML 2024)

*Wanpeng Zhang, Yilin Li, Boyu Yang, Zongqing Lu.*

- COREP tackles non-stationarity problems in RL by employing a dual-GAT structure to learn a stable graph representation based on the causal origins of non-stationarity.

**AdaRefiner: Refining Decisions of Language Models with Adaptive Feedback.** (NAACL 2024)

*Wanpeng Zhang, Zongqing Lu.*

- By actively receiving and incorporating feedback from the RL agent, AdaRefiner uses this information to finetune the LLM. This process enhances the LLM's understanding of downstream tasks and the agent's status, thereby improving the guidance it provides to the agent.

**Entity Divider with Language Grounding in Multi-Agent Reinforcement Learning.** (ICML 2023)

*Ziluo Ding\*, Wanpeng Zhang\*, Junpeng Yue, Xiangjun Wang, Tiejun Huang, Zongqing Lu.*

- EnDi is designed for improving policy generalization in multi-agent RL by enabling agents to independently learn and execute subgoals at the entity level, which is enhanced with opponent modeling to manage subgoal conflicts and coordinate strategies.

## Publication

---

- **Model-Based Opponent Modeling.** X Yu, J Jiang, **W Zhang**, H Jiang, Z Lu. (NeurIPS 2022)
- **iGrow: A Smart Agriculture Solution to Autonomous Greenhouse Control.** X Cao, Y Yao, L Li, **W Zhang**, Z An, Z Zhang, S Guo, L Xiao, X Cao, D Luo. (AAAI 2022)
- **Efficient and Stable Information Directed Exploration for Continuous Reinforcement Learning.** M Chen, X Xiao, **W Zhang**, X Gao. (ICASSP 2022)
- **Robust Model-based Reinforcement Learning for Autonomous Greenhouse Control.** **W Zhang**, X Cao, Y Yao, Z An, X Xiao, D Luo. (ACML 2021)
- **Sample Efficient Reinforcement Learning via Model-Ensemble Exploration and Exploitation.** Y Yao, L Xiao, Z An, **W Zhang**, D Luo. (ICRA 2021)
- **A Simulator-based Planning Framework for Optimizing Autonomous Greenhouse Control Strategy.** Z An, X Cao, Y Yao, **W Zhang**, L Li, Y Wang, S Guo, D Luo. (ICAPS 2021)
- **Self-Paced Probabilistic Principal Component Analysis for Data with Outliers.** B Zhao, X Xiao, **W Zhang**, B Zhang, G Gan, S Xia. (ICASSP 2020)

## Patent

---

**Method, device and equipment for determining parameters and storage medium.**  
(CN112527104A)

*Wanpeng Zhang, Dijun Luo, Xi Xiao.*

## Award

---

- **Rhino-bird Elite Training Program of Tencent AI Lab.** (Jul. 2021)
- **MCM/ICM Meritorious Winner.** (Jan. 2017)
- **Second Prize of National University Students Mathematical Contest in Modeling.** (Jan. 2016)

## Service

---

### Conference Reviewer

- ICML 2022, 2023, 2024
- NeurIPS 2022, 2023
- ICLR 2024