

张之夏

22376220@buaa.edu.cn | zhixiazhang67@gmail.com | +86 18291752379

教育背景

北京航空航天大学

计算机科学与技术

北京

2022 – 至今

- **GPA:** 3.78/4.00 (加权平均: 90.2)
- **荣誉:** 校级优秀学生, 校级优秀学生干部, 校级学术奖学金一等奖, 校级学科竞赛将学金一等奖, 校级科技创新二等奖学金, 校级社会服务奖学金二等奖
- **主修课程:** 工科大学物理 I (96) 工科大学物理 II (100) 线性代数 (99) 工科数学分析 II (93) 概率统计 (98) 数据结构 (97) 计算机组成原理 (94)

发表论文&专利

- [1] **Zhang, Z.**, et al. HAPCO: Heterogeneous Agent Collaborative Policy Optimization. (Submitted to ICML 2026).
- [2] Huang, Z., ..., **Zhang, Z.**, et al. (Second Student Author) Does Your Reasoning Model Implicitly Know When to Stop Thinking? (Submitted to ICML 2026).
- [3] Xu, R., Liu, Z., **Zhang, Z.**, et al. Adversarial Policy Transfer in Mixed Cooperative-Competitive Games. (Submitted to ICML 2026).
- [4] Xia, H., Li, S., Xu, R., **Zhang, Z.**, et al. Bayesian Robust Financial Trading with Adversarial Synthetic Market Data. (Accepted by KDD 2026).
- [5] 刘祥龙、李思辰、徐睿、刘志晨、张之夏、郭晔阳. 一种基于智能群体博弈的可转移鲁棒性评估方法. (中国专利: CN121118964A, 2025 年 12 月发布)

研究经历

科研实习, 北京航空航天大学 (班义琨, KTL 实验室)

2025.7 – 至今

- **HACPO: Heterogeneous Agent Collaborative Policy Optimization (一作):** 提出了一种新的多智能体强化学习训练范式, 能够共享 rollout, 最大化 rollout 利用率, 并促进异构智能体之间的知识转移。实现了算法创新, 弥补了由于异构性造成的能力差异和策略分布不一致。
- **LLM Reasoning Efficiency Optimization (学生二作):** 研究了长链推理过程中的冗余, 提出了基于强化学习的机制, 将“停止推理时机”的隐式“信心”转移到推理过程中。
- **MARL & RLVR Integration:** 将可验证奖励的强化学习 (RLVR) 与多智能体强化学习 (MARL) 相结合的技术探索, 在 VeRL 框架内设计并开发了相关算法。

科研实习, 北京航空航天大学 (刘祥龙)

2024.10 – 2025.10

- **Transferable Adversarial Policy Framework:** 出了一个可转移的对抗策略框架, 在未知的合作-竞争环境中 zero-shot 攻击。该框架利用大语言模型 (LLMs) 在训练过程中提取欺骗性策略, 并在测试过程中通过贝叶斯推理推测攻击权重。
- **Transferable Robustness Evaluation for Intelligent Swarms:** 开发了一种新型评估框架, 用于评估智能群体游戏中的智能体鲁棒性, 能够在不同任务、策略类型和智能体配置间泛化。通过利用大语言模型 (LLMs) 提取群体互动数据中的行为模式, 并采用多场景迭代训练方法, 提高了评估效率和系统适应性。(中国专利: CN121118964A)
- **Bayesian Robust Financial Trading:** 开发了一种贝叶斯稳健交易框架, 通过生成多样的市场条件来增强训练。该方法应用贝叶斯神经虚拟自我博弈 (Bayesian NFSP) 优化交易策略, 使智能体能够在宏观经济不确定性下保持稳定的盈利性和风险控制。

专业技能 & 研究兴趣

- 编程语言: Python, Java, C
- 工具 & 语言: PyTorch, LaTeX, English (IELTS: 7.0, GRE: 319)
- 研究兴趣: Multi-Agent Reinforcement Learning (MARL), Reinforcement Learning from Verifiable Rewards (RLVR), Vision Language Action (VLA) models