

# 郝晓田

xiaotianhao@tju.edu.cn | +86-151-2293-8076



## 个人简介

研究方向为 RL (强化学习) 与 MAS (多智能体系统), 关注其在 RLHF、广告投放、组合优化、游戏 AI 等领域应用, 具有非常强的探索精神和动手实现能力, 擅长因地制宜地解决企业实际问题, 有丰富的团队协作和落地实践经验。曾获华为诺亚方舟实验室优秀实习生 (Top 3/200), 个人主页: tjuHaoXiaotian。

## 实习经历

### 华为 (2022.1-2023.7) 诺亚方舟实验室-决策推理研究团队-Research Intern

① 盘古大模型 RLHF 优化项目: ① 负责 Reward Model 设计与训练, 基于 ChatGLM LoRA, DeepSpeed 搭建训练框架, 针对产品线标注数据集, 输出测试精度 89% 的模型。② 提出基于 Stochastic Beam Search 实现指数空间高效无放回 top-k 采样, 提升样本质量 & 多样性, 提高 RAFT 效率。③ 整理 RLHF 技术博客。

② 高维优化项目: 针对多智能体系统中状态-动作空间高维难以优化问题, 提出结合置换不变性与置换同变性的归纳偏置设计算法, 在星际争霸等基准测试环境中取得 SOTA 性能, 研究成果发表在 ICLR 2023, 开源代码库 pymarl3。

### 华为 (2020.6-2021.6) 诺亚方舟实验室-企业智能团队-Algorithm Engineer

① 自研大规模线性规划求解器项目: 提出基于最大堆的 pricing 加速算法, 预估变量价值构造最大堆实现高效 top-1 (入基) 检索。在华为供应链多工厂排产等核心业务落地, 千万变量线性规划求解时间提速一倍以上, 已集成在华为云天筹 AI 求解器对外发布, 该技术被评为华为潜高等级专利 (公开号: CN115496247A)。

② DPDP 优化项目: 针对复杂 DPDP 问题, 提出基于 RL 的分层搜索算法, 上层负责动态问题切分, 下层负责静态问题求解, 在松山湖制造投递运输系统落地测试, 订单超时减少 42%, 车辆行驶距离减少 23%, 研究成果发表在 NeurIPS 2021。

### 阿里巴巴 (2018.10-2019.10) 阿里妈妈-精准广告团队-Research Intern

① 多渠道联合的广告序列化投放项目: 广告精排阶段, 研究广告不同投放组合对用户未来行为及价值的影响, 建模为动态背包问题 (物品价值和成本受投放策略的影响), 采用 Bilevel 优化, 并证明最优性。线上 A/B test 广告主 ROI +25%, 双十一期间 20% 流量, ROI +11%, 研究成果发表在 ICML-2020。

② 手淘 Banner 钻石展位, 多智能体锦囊优化项目: 广告召回阶段, 广告主之间的博弈优化, learning 算法加速求解大规模 b-matching 问题, 大盘 RPM +19% (离线测试), 研究成果发表在 IJCAI-2020。

## 发表论文 | 一作

- Multiagent Gumbel MuZero: Efficient Planning in Combinatorial Action Spaces (arXiv 2023).  
作者: Xiaotian Hao, Jianye Hao, Chenjun Xiao, Kai Li, Dong Li.
- Boosting Multiagent Reinforcement Learning via Permutation Invariant and Permutation Equivariant Networks (ICLR 2023).  
作者: Xiaotian Hao, Jianye Hao, Hangyu Mao, Weixun Wang, Dong Li, Yan Zheng.
- A Hierarchical Reinforcement Learning Based Optimization Framework for Large-scale Dynamic Pickup and Delivery Problems (NeurIPS 2021).  
作者: Yi Ma, Xiaotian Hao, Jianye Hao, Jiawen Lu, Mingxuan Yuan, Zhaopeng Meng.
- Dynamic Knapsack Optimization Towards Efficient Multi-Channel Sequential Advertising. (ICML 2020).  
作者: Xiaotian Hao, Zhaoqing Peng, Yi Ma, Junqi Jin, Jianye Hao, Rongquan Bai, Chuan Yu, Han Li, Jian Xu, Kun Gai.
- Learning to Accelerate Heuristic Searching for Large-Scale Maximum Weighted b-Matching Problems in Online Advertising (IJCAI 2020).  
作者: Xiaotian Hao, Junqi Jin, Jianye Hao, Jin Li, Weixun Wang, Han Li, Jian Xu, Kun Gai.
- Independent Generative Adversarial Self-Imitation Learning in Cooperative Multiagent Systems (AAMAS 2019).  
作者: Xiaotian Hao, Weixun Wang, Jianye Hao, Yaodong Yang.

## 教育背景

### 天津大学 (2020-2024)

博士: 软件工程, 导师: 郝建业  
预计 2024 年毕业

### 天津大学 (2017-2020)

硕士: 软件工程, 导师: 郝建业

### 天津大学 (2013-2017)

本科: 软件工程

Rank: 3 / 110, 保研本校

## 学术链接

### Google 学术

XiaotianHao, Citation 290, h-index 9  
i10-index 9

### GitHub

tjuHaoXiaotian, 开源 MARL 研究领域 SOTA 算法 pymarl3, MultiagentRL 和 MA-MuZero 代码库, 关注开源进展

### 知乎

郝晓田, 获 1.3k 收藏, 专注 RL 领域知识输出, 代表文章: LLaMA2 技术细节, 值分解强化学习算法回顾等。

## 比赛经历

NeurIPS 2020 MineRL 国际竞赛冠军

NeurIPS 2022 SMARTS 最佳参赛奖

先知·兵圣 2018 人机对抗大赛一等奖  
互联网+创新创业大赛天津赛区银奖

## 其他奖项

华为优秀实习生 (Top 3/200)

研究生国家奖学金 (Top 2/92)

## 发表论文 | 合作

- **Learning to Collaborate by Grouping: a Consensus-oriented Strategy for Multiagent RL (ECAI 2023).**  
作者: Jingqing Ruan, [Xiaotian Hao](#), Dong Li, Hangyu Mao.
- **PORTAL: Automatic Curricula Generation for Multiagent Reinforcement Learning (AAMAS 2023).**  
作者: Jizhou Wu, Tianpei Yang, [Xiaotian Hao](#), Jianye Hao, Yan Zheng, Weixun Wang, Matthew E Taylor.
- **Transformer-based Working Memory for Multiagent RL with Action Parsing (NeurIPS 2022).**  
作者: Yaodong Yang, Guangyong Chen, Weixun Wang, [Xiaotian Hao](#), Jianye HAO, Pheng-Ann Heng
- **PMIC: Improving Multiagent RL with Progressive Mutual Information Collaboration (ICML 2022).**  
作者: Pengyi Li, Hongyao Tang, Tianpei Yang, [Xiaotian Hao](#), Tong Sang, Yan Zheng, Jianye Hao, Matthew E Taylor, Wenyuan Tao, Zhen Wang, Fazl Barez
- **SEIHAI: A Sample-efficient Hierarchical AI for the MineRL Competition (DAI 2021).**  
作者: Hangyu Mao, Chao Wang, [Xiaotian Hao](#), Yihuan Mao, Yiming Lu, Chengjie Wu, Jianye Hao, Dong Li, Pingzhong Tang
- **Large Scale Deep Reinforcement Learning in War-games (BIBM 2020).**  
作者: Hanchao Wang, Hongyao Tang, Jianye Hao, [Xiaotian Hao](#), Yue Fu, Yi Ma
- **Combining Sequence and Network Information to Enhance Protein-Protein Interaction Prediction (BIBM 2020).**  
作者: Leilei Liu, Xianglei Zhu, Yi Ma, Haiyin Piao, Yaodong Yang, [Xiaotian Hao](#), Yue Fu, Li Wang, Jiajie Peng
- **From Few to More: Large-scale Dynamic Multiagent Curriculum Learning (AAAI 2020)**  
作者: Weixun Wang, Tianpei Yang, Yong Liu, Jianye Hao, [Xiaotian Hao](#), Yujing Hu, Yingfeng Chen, Changjie Fan, Yang Gao
- **Action Semantics Network: Considering the Effects of Actions in Multiagent Systems (ICLR 2020)**  
作者: Weixun Wang, Tianpei Yang, Yong Liu, Jianye Hao, [Xiaotian Hao](#), Yujing Hu, Yingfeng Chen, Changjie Fan, Yang Gao
- **一种多智能体合作式环境下基于带权估计的策略优化算法 (计算机学报 2020)**  
作者: Yan Zheng, Jianye Hao, Zongzhang Zhang, Zhaopeng Meng, [Xiaotian Hao](#)
- **Learning Adaptive Display Exposure for Real-Time Advertising (CIKM 2019).**  
作者: Weixun Wang, Junqi Jin, Jianye Hao, Weinan Zhang, Jun Wang, [Xiaotian Hao](#), Yixi Wang, Han Li, Jian Xu, Kun Gai

## 预发表论文

- **Multiagent Gumbel MuZero: Efficient Planning in Combinatorial Action Spaces.**  
作者: [Xiaotian Hao](#), Jianye Hao, Chenjun Xiao, Kai Li, Dong Li.
- **Hybrid CtrlFormer: Learning Adaptive Search Space Partition for Hybrid Action Control via MCTS.**  
作者: Jiashun Liu, [Xiaotian Hao](#), Jianye Hao, Yan Zheng, Yujing Hu, Tangjie Lv, Changjie Fan, Zhipeng Hu
- **Mitigating Stochastic Communication Latency in Real-Time Control through Multi-Step Reinforcement Learning.**  
作者: Jiashun Liu, [Xiaotian Hao](#), Jianye Hao, Yan Zheng, Yujing Hu