Chen-Xiao Gao

typoverflow@gmail.com | typoverflow | Home Page

Education

Nanjing University, Nanjing, China 09/2022 – 06/2025 (expected)
Master of Science in Computer Science

Nanjing University, Nanjing, China 09/2018 – 06/2022
Bachelor of Science in Computer Science, GPA: 4.59/5.0, Rank: 5/77

Awards & Honors

2022,2023 Graduate Academic Scholarship (top **20%**)

2021 Outstanding Student of Nanjing University (top **10%**)

2021 National Scholarship (top **0.1%**)

2020 The First-Class Chow Tai Fook Scholarship (top **0.3%**)

2020 People's Scholarship (top **30%**)

Publications

(* denotes equal contribution)

- Dmitry Shribak*, Chen-Xiao Gao*, Yitong Li, Chenjun Xiao, Bo Dai. Diffusion Spectral Representation for Reinforcement Learning. NeurIPS'24. [webpage, paper link]
- Rui Kong, Chenyang Wu, Chen-Xiao Gao, Yang Yu, Zongzhang Zhang. Efficient and Stable Offline-to-Online Reinforcement Learning via Continual Policy Revitalization. IJCAI'24. [paper link]
- Chengxing Jia*, Chen-Xiao Gao*, Hao Yin, Fuxiang Zhang, Xiong-Hui Chen, Tian Xu, Lei Yuan, Zongzhang Zhang, Yang Yu, Zhi-Hua Zhou. Policy Rehearsing: Training Generalizable Policies for Reinforcement Learning. ICLR'24. [paper link]
- Chen-Xiao Gao, Chenyang Wu, Mingjun Cao, Rui Kong, Zongzhang Zhang, Yang Yu. ACT: Empowering Decision Transformer with Dynamic Programming via Advantage Conditioning. AAAI'24. [paper link]
- Renzhe Zhou, Chen-Xiao Gao, Zongzhang Zhang, Yang Yu. Generalizable Task Representation Learning for Offline Meta-Reinforcement Learning with Data Limitations. AAAI'24. [paper link]
- Chengxing Jia, Fuxiang Zhang, Yi-Chen Li, Chen-Xiao Gao, Xu-Hui Liu, Lei Yuan, Zongzhang Zhang, Yang Yu. Disentangling Policy from Offline Task Representation Learning via Adversarial Data Augmentation. AAMAS'24. [paper link]
- Lei Song*, Chen-Xiao Gao*, Ke Xue, Chenyang Wu, Dong Li, Jianye Hao, Zongzhang Zhang, Chao Qian. Reinforced In-Context Black-Box Pptimization. ArXiv preprint. 2024.
- Chen-Xiao Gao, Shengjun Fang, Chenjun Xiao, Yang Yu, Zongzhang Zhang. Hindsight Preference Learning for Offline Preference-based Reinforcement Learning. Arxiv preprint. 2024.
- Chen-Xiao Gao, Chenyang Wu, Mingjun Cao, Chenjun Xiao, Yang Yu, Zongzhang Zhang. Behavior-Regularized Diffusion Policy Optimization for Offline Reinforcement Learning. ArXiv preprint. 2025.

Research Experiences

The Chinese University of Hong Kong, Shenzhen, China 02/2024 – 09/2024
Research Assistant, Advisor: Prof. Chenjun Xiao and Prof. Bo Dai

- Diffusion Spectral Representation for Reinforcement Learning
  Developed Diff-SR (accepted at NeurIPS'24) to extract **spectral representations** that effectively capture the latent structure of the transition function. These representations are sufficiently expressive to represent the value functions of any policy, thereby facilitating efficient planning and reinforcement learning.

Nanjing University, Nanjing, China 09/2022 – Present
Research Assistant, Advisor: Prof. Zongzhang Zhang

- Offline Model-Based Adaptable Policy Learning
  Investigated the idea of **rehearsal** in offline reinforcement learning, which aims to generate diverse yet eligible dynamics models using an extremely limited amount of offline data and optimize a contextual policy with the generated models. By recognizing the context, the policy can effectively generalize to the specific environment during the online phase. Accepted by ICLR'24.

- Offline Meta-Reinforcement Learning with Data Limitations
  Data limitations in offline scenarios may cause reinforcement learning agents to learn spurious task representations. Developed two algorithms, ReDA (accepted at AAMAS'24) and Gentle (accepted at AAAI'24), to mitigate this issue by employing adversarial data augmentation and reconstruction objectives respectively, enhancing the robustness and performance in offline meta-RL settings.
- Sequence Model Meets Decision Making
  Applied sequence modeling techniques to decision-making problems such as continuous control and black-box optimization. For continuous control, we identified potential failure modes of the Decision Transformer and proposed ACT (accepted by AAAI'24) to leverage advantage conditioning to achieve robust control. For black-box optimization, we propose distilling and reinforcing existing black-box optimization algorithms by fitting regret-augmented learning histories of the behavior algorithms, enabling the sequence model to function as a universal optimizer (currently in submission).

ByteDance, Beijing, China 07/2021 – 11/2021
Research Intern in Game Development

- Deployed reinforcement learning in game scenarios, combining reinforcement learning and self-supervised imitation learning to achieve stable and progressive performance that surpassed baseline algorithms.
- Explored the incorporation of human knowledge into deep reinforcement learning algorithms by transforming predefined decision trees to initialize agents in challenging exploration scenarios.

Projects

Large-Scale Reinforcement Learning for MOBA Games 02/2024 – 09/2024
Team Project (Team Leader)

- Utilized Tencent's platform to design human-level AI agents for the game *Honor of Kings*, tackling complex decision-making problems with large state and action spaces.
- Developed efficient distributed RL algorithms based on asynchronous PPO, incorporating competition and game theory to address multi-agent challenges. Built a robust data analysis pipeline to facilitate the iteration of algorithm development.
- Ranked **1st** in the 1v1 leaderboard and **5th** in the 3v3 leaderboard among over 40 teams nationwide.

Open-Sourced Reinforcement Learning Algorithm Framework 07/2022 – Present
Individual Project

- As the main developer of `Unstable-Baselines`, `OfflineRL-Lib` and `WiseRL`, implemented over 15 popular online, offline, and preference-based reinforcement learning algorithms with flexible and research-friendly module design.
- Received over 200 stars on GitHub.

Skills

- **Proficient**: Python, PyTorch, TensorFlow, Linux, LaTeX, Git
- **Familiar**: C/C++, Jax, MATLAB, HTML, etc.
- **Language**: TOFEL iBT 109