Xiaoyu Wen

wenxiaoyu@mail.nwpu.edu.cn ◇ (+86) 18768515132 ◇ Github

Education

Northwestern Polytechnical University Xi'an, China
Master of Engineering in Artificial Intelligence. GPA: 87.65/100. Advised by Dr. Zhen Wang Expected April 2025
Role: Class Monitor, Research Institute Volunteers Department Leader

Harbin Engineering University Harbin, China
Bachelor of Engineering in Computer Science and Technology. GPA: 87.50/100, Rank: 3/34 09/2018 - 06/2022
Major Courses: Linear Algebra(98), Algorithms Design and Analysis(95), Artifical Intelligence(excellent)

Professional Experience

Tencent AI Platform Dept. Shenzhen, China
Role: Machine Learning Engineer 05/2024 - Present
Brief introduction: Research on designing anthropomorphic game AI bots via Reinforcement Learning.

- Primarily focus on designing effective navigation algorithms, including reward shaping and algorithm innovation.
- Fine-tuning a game voice Q&A assistant using RLHF, such as DPO and SimPO.

Shanghai AI Lab Shanghai, China
Role: Member of the RL group; Advised by Dr. Chenjia Bai 10/2023 - 04/2023
Brief introduction: Research on offline reinforcement learning and cross-domain policy adaptation.

- Primarily focus on the transition from offline policy to online learning and cross-dynamics generalization issues, investigating cross-domain policy generalization and cross-task reuse in reinforcement learning.
- Address the shortcomings of existing offline reinforcement learning algorithms, such as low data efficiency and poor adaptability to real-world applications.
- Planning to explore the application of diffusion models in reinforcement learning, aiming to quantify cross-domain dynamic differences and enhance data efficiency.

Research Interest

I am broadly interested in research on **Reinforcement Learning**, and my current research mainly focus on:

- Generalization & Adaptation (Cross-dynamics/Rewards; Robustness; Meta RL; Sim-to-Real transfer)
- Learning with Prior Knowledge (Offline RL; Offline-to-Online RL; Learning from Demonstration)
- Learning with Generative models (Diffuser; Decison-Diffuser; SynthER)
- Human-in-the-Loop (Task Specification with Preferences/Symbolic Representation/Language Instruction)

Research Proposal

- **Cross-Domain Diffusion-Based Policy Adaptation**
  Cross-domain offline reinforcement learning leverages source domain data with diverse transition dynamics to alleviate the data requirement for the target domain. However, simply merging the data of two domains leads to performance degradation due to the dynamics mismatch. Previous methods have attempted to address this problem via dynamics discrepancy, value differences, and representation mismatch. And It has not been studied in the diffusion-based policy (which is more robust and also can be studied in some fancy experiment beyond D4RL).

- **Human-in-the-Loop Learning beyond Preferences or Demonstrations**
  How to integrate human knowledge into the RL system is a long-standing problem. Existing researches primarily focus on utilizing human demonstrations or human preferences for learning. However, perception, reasoning, and planning, which are essential compositions of human intelligence, are mostly overlooked by current machine learning systems. A general Human-in-the-Loop system embedding these factors would be cutting-edge technology.

- **Practical Application of Generative Flow Networks**
  GFlownet, as a deep learning approach to optical flow estimation, integrates cutting-edge technologies from computer vision, image processing, and machine learning. It not only accurately predicts pixel displacements in image sequences but also enables precise motion tracking and analysis in dynamic scenarios, holding critical implications for real-time video processing and motion capture. And I am deeply intrigued by it and particularly inclined towards exploring practical applications in this area.

Publications

- Contrastive Representation for Data Filtering in Cross-Domain Offline Reinforcement Learning [arXiv]
  Xiaoyu Wen, Chenjia Bai, Kang Xu, Xudong Yu, Yang Zhang, Xuelong Li, Zhen Wang.
  International Conference on Machine Learning (**ICML**), 2024
- Towards Robust Offline-to-Online Reinforcement Learning via Uncertainty and Smoothness [Link]
  Xiaoyu Wen*, Xudong Yu*, Rui Yang, Haoyuan Chen, Chenjia Bai, Zhen Wang.
  Journal of Artificial Intelligence Research (under review), 2023
- Project paper: Electric Vehicle Charging Guidance Strategy Considering Waiting Time at Target Charging Stations
  Botao Wang, **Xiaoyu Wen**, Zhaoheng Cao, Jialong Sun, Chao Gao, Peican Zhu, Zhen Wang.
  International Conference on Information, Communication and Networks (**ICICN**), EI, 2023
- Patent: A Reinforcement Learning Method and Device for Transferring from Offline policy to Online Learning
- Patent: A Cross-Domain Offline Reinforcement Learning Method and Device Based on Contrastive Representation
- Software Copyright: Data Mining-based Traffic Product Precision Marketing Software

Honors and Awards

Graduate Innovation and Entrepreneurship Fund **(Key Project)**	2023
Outstanding Graduate Student	2023
A-class Excellent Scholarship, Northwestern Polytechnical University **(top 5%)**	2022
A-class Excellent Scholarship, Harbin Engineering University **(top 5%)**	2021
Mathematical Contest in Modeling (MCM), **Meritorious Winner**	2021
Outstanding Class Leader	2021
Merit Student	2020
B-class Excellent Scholarship, Harbin Engineering University **(top 10%)**	2020
B-class Excellent Scholarship, Harbin Engineering University **(top 10%)**	2019
National College Students Cryptography Competition Semi-Finals	2019

Technical Skills

Programming Languages: C++ == Python > Go
Tools: git for code management; common operations in Linux
Deep Learning: Pytorch and Tensorflow for conducting machine experiments and designing neural networks

Miscellaneous

- Licenses & certifications: Driver's license, National Basketball Second Level Referee
- Languages: English - Fluent, Mandarin - Native speaker
- Hobbies: Basketball, Soccer, Photography