Research Statement

Yuming Chen

My research interests lie at the intersection of **reinforcement learning**, **game theory**, and **multi-agent systems**, particularly in the interactions between agents and humans. My long-term goal is to develop agents that can coordinate, collaborate, and even guide humans in real-world environments.

Experience

Learning Temporal Information from Low-Quality Data

Learning from low-quality data is critical for agents operating in real-world environments, where "Data is more costly". During my master's research, I tackled the challenge of recovering hand movement sequences from a single blurry image in the field of Computer Vision. Most work in Hand Pose Estimation focuses on sharp images, but a great number of in-the-wild hand images are blurry, leading to poor estimation performance.

To address this, we proposed a Multi-Hypothesis model that generates multiple plausible estimations from a single image, capturing meaningful temporal movement sequences from ambiguous patterns. This approach leverages a **Reward Model** to select the best estimations. The reward model is similar to the one in Reinforcement Learning from Human Feedback (RLHF). In our task, however, there is a ground-truth reward to indicate the quality of estimations. It is different from giving a partial order for text pairs based on incomplete human preference. The reward model fails to learn a good ranking for all estimations with the one-to-one comparison used in RLHF. We tackled the *Reward Ambiguity* problem by proposing a novel loss function that improves alignment with the target by selecting the top-n estimations.

Our findings have potential applications in Inverse Reinforcement Learning and Imitation Learning, where efficiently using existing data is crucial. And utilising low-quality data enlarge the available dataset. This work has been submitted to the Winter Conference on Applications of Computer Vision (WACV) 2025.

Learning with Opponent Awareness

Prior to my postgraduate studies, I explored the cooperation problem in **Multi-Agent Reinforcement Learning** (MARL). In many multi-agent tasks, optimizing the total reward of all agents is essential. For instance, in a self-driving car system, all agents aim to avoid traffic jams, or in Multi-Agent MuJoCo, agents strive for task completion by cooperation. However, agents do not always face the same partners, making policies that ignore the others' behaviors potentially suboptimal and unstable in real-world scenarios, where exists different type of partners.

To address this, we trained policies conditioned on both personal observations and the opponent's policy. We introduced **Contrastive Learning** to learn stable representations of the opponent's policy, which can extract invariant information during interactions. Similar idea is also introduced in robotics, e.g., Reusable Representations in Robotic Manipulation (R3M). This work[1], accepted at the International Conference on Neural Information Processing (ICONIP) 2023, demonstrates the benefits of opponent-aware strategies in multi-agent systems. It is also helpful to the interaction with humans.

Additional Research

Beyond these projects, I have engaged in other research work related to robotics and game theory. For instance, I explored learning grasping policies for dexterous hands through **Third-Person Imitation Learning** using video data of human grasping. Additionally, I developed a robust Unmanned Aerial Vehicle (UAV) control policy through Reward Shaping and Policy Space Response Oracles (PSRO) in simulation.

In game theory, I studied the decomposition of Markov Games, aiming to split multi-agent tasks into fully cooperative and fully competitive components. These experiences improve my robotics skills and theoretical skills.

Future Research

I am excited to further explore human-aware robotic agents using MARL. MARL research not only involves controlling multiple agents but also finding optimal policies for agents that consider others. It is, in my opinion, a crucial aspect in enabling agents to have human-aware. Your work *Mapless Navigation among Dynamics with Social-Safety Awareness: A Reinforcement Learning Approach from 2D Laser Scans* greatly inspired me, especially the elegant design of reward and observation. I believe incorporating opponent modelling, a concept from MARL, by predicting pedestrian velocities could make robots more robust against a diverse set of pedestrians, such as varying moving speeds.

Moreover, I am interested in counterfactual reasoning, which has been widely studied in both MARL and game theory. Your work *Offline Learning of Counterfactual Predictions for Real-World Robotic Reinforcement Learning* taught me a novel way to make counterfactual predictions based on the behaviour policy. It may further accelerate the learning process by applying different counterfactual prediction methods, such as maintaining state values or using regret as the counterfactual value. counterfactual baselines. I believe maintaining state values or sampling from the target policy could further accelerate learning.

Looking ahead, I am keen to investigate embodied policy learning from offline third-person datasets and explore the application of learning algorithms to embodied large language models (LLMs) agents. I am also eager to enhance my practical programming skills for robotic arms, dexterous hands, and other robotic platforms.

In conclusion, I am eager to contribute to the development of intelligent, human-aware robotic agents that can interact and cooperate with humans in dynamic, real-world environments. My experience in reinforcement learning and computer vision has provided me with the technical foundation to pursue this research. I am excited about the opportunity to work under your guidance and further explore the problems related to embodied robotic agents. I look forward to the possibility of contributing to and growing within your research group.

References

[1] Yuming Chen and Yuanheng Zhu. Policy representation opponent shaping via contrastive learning. In Biao Luo, Long Cheng, Zheng-Guang Wu, Hongyi Li, and Chaojie Li, editors, *Neural Information Processing*, pages 124–135, Singapore, 2024. Springer Nature Singapore.