Statement of Purpose

Wentse Chen

My research interests lie in multi-agent reinforcement learning and robotics, with the goal of robots being able to collaborate, explore and learn like human beings. During my undergraduate years, I had the opportunity to work in the TSAIL group at Tsinghua University and visit the Hybrid Robotics group at UC Berkeley. These opportunities inspired me to apply to your M.S. program in the Robotics Institute with a focus on robotics and decision-making.

Tikick[1]: My research starts with cooperative multi-agent RL. To complete Google research football (GRF) games by controlling all players in the team simultaneously, we developed a distributed learning system as well as a novel offline reinforcement learning algorithm to learn a powerful cooperative multi-agent AI from the fixed single-agent dataset. Our work is the first learning-based multi-agent AI that can take over GRF games. We also show that our work outperforms the state-of-the-art single-agent RL-based algorithm. Although this project takes us one step closer to a strong multi-agent AI system, it focuses only on the fixed environment and still has a lot of space for improvement. Then I try to further strengthen the multi-agent AI system utilizing best-response methods. However, the meta-learner will stop improving after a period of self-play training due to the lack of population diversity. Therefore, I turn to solve the problem of improving the policy diversity of RL.

DGPO[2]: In this project, we proposed an on-policy framework for discovering multiple strategies for the same task in a single training process. We introduced diversity metrics based on mutual information and formalized two constrained optimization problems to efficiently discover a set of optimal strategies. Our approach outperforms previous algorithms in terms of efficiency, efficacy, and diversity. I believe that diversity has a great impact on multi-agent systems. For cooperation tasks, diversity improves the robustness and generality of the algorithm. That is, when multi-agent systems learn a set of diverse policies, they are better able to adapt to environmental perturbations. Moreover, they adapt well even when there are agents in the system that use different policies, which is helpful for human-robot cooperative tasks. For competitive tasks, diversity allows for better convergence of the self-play training algorithm. This is also the first time I have led a project, from motivation, algorithm proposal to experimentation.

TiZero[3]: At the same time, I try to approach the problem from another perspective. In this project, we develop an adaptive curriculum for learning, a novel self-play strategy, and an

---

[1] **Chen Wentse*** ,Huang Shiyu*, Zhang Longfei, Li Ziyang, Zhu Fengming, Ye Deheng, Chen Ting, and Zhu Jun. TiKick: Towards Playing Multi-agent Football Full Games from Single-agent Demonstrations. Offline Reinforcement Learning Workshop at Neural Information Processing Systems, (2021).

[2] **Chen Wentse**, Huang Shiyu, Chiang Yuan, Chen Ting, and Zhu Jun. DGPO: Discovering Multiple Strategies with Diversity-Guided Policy Optimization. The 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS). Under review. (2022).

[3] Lin Fanqi*, Huang Shiyu*, Pearce Tim, **Chen Wentse**, and Tu Wei-Wei. TiZero: Mastering Multi-Agent Football with Curriculum Learning and Self-Play. The 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS). Under review. (2022)

objective that optimizes the policies of multiple agents jointly to train multi-agents AI systems to play GRF. We demonstrate that our method outperforms previous systems by a large margin. In this project, I have a deeper understanding of large-scale multi-agent RL.

Collaborative Quadruped robots: Nevertheless, I got concerned that reinforcement learning is limited to playing video games. Then I turned my direction closer to robotics and found a match between my interests with Prof. Koushil Sreenath. Under his guidance, I led a project on collaborative quadrupedal robots. Specifically, we proposed a decentralized, hierarchical controller to manipulate multiple quadruped robots to navigate a cable-towed load through cluttered spaces. This project also determines my future research topic for developing large-scale, decentralized, RL-based algorithms for navigation or control tasks of robots.

I believe that decision-making is a promising topic with a myriad of unexplored research directions. Firstly, large-scale, long-horizon sequential decision-making in stochastic or partially observed environments is still challenging, especially in multi-agent scenarios. Developing a concise, efficient, and general decision-making algorithm based on reinforcement learning and game theory will benefit practical problems, i.e., robotics and self-driving cars. Secondly, the policy diversity of RL is an underestimated research direction. It allows machines to learn like human beings. Diving into how policy diversity affects the robustness and generality of RL is a promising direction. Thirdly, it is still difficult to combine model-based control and model-free reinforcement learning to construct a robust hybrid robotics system. Trading off between the robustness and the efficacy of the system is still a problem to be solved. Finally, applying cutting-edge decision-making algorithms on solving practical problems will solve many difficulties. For example, decentralized MARL algorithms allow robots do online inference. On the other hand, discovering problems in real-life scenarios can deepen our understanding of the algorithm, such as continuous control. I hope to explore these areas during my graduate studies.

I am applying to the MSR program in CMU because of its impactful research groups. There are several professors whose projects are particularly appealing to me. Prof. Jiaoyang Li work on multi-agent/robot systems is fascinating to me, which is the most related direction to my current research. Prof. Changliu Liu's has wonderful works on applying reinforcement learning to robotics control. I also admire Prof. Deepak Pathak's work on locomotion or navigation task of legged robots. After pursuing my degree in your program, I plan to continue my research in the Ph.D. career, and this M.S. experience can definitely pave the way towards it. Overall, I believe CMU's collaborative environment and rich resources can provide the best guidance to my academic career.