

Robotics and Autonomous Systems

Thesis Defense

Joint Learning of Reward Machines and Policies for Multi-Agent Reinforcement Learning in Non-Cooperative Stochastic Games

School for Engineering of Matter, Transport and Energy

Hyohun Kim

Advisor: Dr. Zhe Xu

Abstract

Multi-agent reinforcement learning (MARL) plays a pivotal role in artificial intelligence by facilitating the learning process in complex environments inhabited by multiple entities. This thesis explores the integration of learning high-level knowledge through reward machines (RMs) with MARL to effectively manage non-Markovian reward functions in non-cooperative stochastic games. Reward machines offer a sophisticated way to model the temporal structure of rewards, thereby providing an enhanced representation of agent decision-making processes. A novel algorithm JIRP-SG is introduced, enabling agents to concurrently learn RMs and optimize their best response policies while navigating the intricate temporal dependencies present in non-cooperative settings. This approach employs automata learning to iteratively acquire RMs and utilizes the Lemke-Howson method to update the Q-functions, aiming for a Nash equilibrium. It is demonstrated that the method introduced reliably converges to accurately encode the reward functions and achieve the optimal best response policy for each agent over time. The effectiveness of the proposed approach is validated through case studies, including a Pacman Game scenario and a Factory Assembly scenario, illustrating its superior performance compared to baseline methods. Additionally, the impact of batch size on learning performance is examined, revealing that a diligent agent employing smaller batches can surpass the performance of an agent using larger batches, which fails to summarize experiences as effectively.

March 25, 2024; 10:30 AM; ECG G214;
Zoom Link: <https://asu.zoom.us/j/5025932499>