Bottom-Up Multi-Agent Reinforcement Learning for Complexity and Uncertainty

Name: Takumi Aotani
Laboratory's name: Intelligent System Control
Supervisor's name: Kenji Sugimoto

Abstract: Multi-agent systems (MASs) are expected to be applied to various real-world problems where a single agent cannot accomplish given tasks. Due to the inherent complexity and uncertainty in the real-world MASs, the manual design of group behaviors of agents is intractable. Multi-agent reinforcement learning (MARL), which is a framework for multiple agents in the same environment to learn their policies adaptively, would be a promising methodology. The conventional MARL methods, however, target a limited class of MASs assuming a common task and a centralized system.

In this study, I propose "bottom-up MARL" as an autonomous distributed framework in which agents have their own tasks. Compared to the conventional MARL, the agents' own tasks increase the complexity of group behaviors, and the autonomous distributed manner increases the uncertainty that threatens safe learning. Hence, this study addresses these issues by developing i) a reward shaping algorithm for the group behaviors, and ii) a meta-optimization method for the bias-variance trade-off in the model learning. These are elemental techniques for learning the group behaviors safely, using the predictive models of the rewards and the dynamics, and the reshaped reward, in model-based reinforcement learning.

i) In order to learn the group behaviors, the rewards of all agents are shared and predicted. Each agent obtains a reshaped reward for the group behaviors based on the prediction for its own state. The proposed algorithm consists of four components: reward prediction, promotion of exploration, classification of interests, and reward shaping. The effectiveness of the proposed method is verified by the simulations and the experiments using real robots.

ii) For long-term safety assurance in uncertain environments, avoiding the predictions with large errors by the dynamics model, is necessary. In this study, we formulate the bias-variance trade-off as a multi-objective optimization problem and develop a meta-optimization method to adjust the trade-off simultaneously with model training. The effectiveness of the proposed method is verified by the simulations.