先端科学技術研究科　修士論文要旨

所属研究室 （主指導教員）	数理情報学 （池田　和司　（教授））		
学籍番号	2311030	提出日	令和 7年 1月 14日
学生氏名	市原　有生希		
論文題目	Towards Robust Reinforcement Learning Algorithms for Nominal Environment		

要旨

One of the principal obstacles in real-world applications of reinforcement learning is the discrepancy between simulated and actual real environments. Consequently, a policy trained in a simulated setting might fail to produce the intended actions in the real world, owing to factors such as noise, modeling inaccuracies, and varying environmental conditions.

To mitigate this issue, the Robust Markov Decision Processes (RMDPs) framework concentrates on designing algorithms resilient to model discrepancies. In a RMDP, one considers a family of possible transition probabilities and reward functions, and selects the worst-case transition probabilities and reward functions within this set for policy optimization. Recent studies suggest that accounting for the entropy and divergence of the policy can capture the worst-case scenario for a given reward function. Certain challenges persist despite the introduction of various algorithms for handling transition probabilities. In particular, the support of the distribution can be inconsistent, and states that do not transition in the real environment may still be assigned nonzero transition probabilities.

In this work, we add a divergence about a soft-optimal policy, and we replace the worst-case transition probabilities with a KL divergence term relative to the nominal environment's transition probabilities. From this operation can be solved the RMDPs 's challenges.