Transfer Learning through Abstraction Using Learning Vector Quantization

Ahmad Afif Mohd Faudzi, Hirotaka Takano and Junichi Murata

***Abstract*—Reinforcement learning (RL) enables an agent to find a solution to a problem by interacting with the environment. However, the learning process always starts from scratch and possibly takes a long time. Here, knowledge transfer between tasks is considered. In this paper, we argue that an abstraction can improve the transfer learning. Modified learning vector quantization (LVQ) that can manipulate its network weights is proposed to perform an abstraction that is expected to provide a simple representation of the transferred knowledge for human interpretation, an adaptation that is expected to train the agent to adapt to new environments and a precaution that is expected to provide a better prior information. At first, the abstraction is performed by extracting an abstract policy out of a learned policy which is acquired through conventional RL method, Q-learning. The abstract policy then is used in a new task as prior information. Here, the adaptation or policy learning as well as new task's abstract policy generating are performed using only a single operation. Finally, as a precaution of future tasks, a common abstract policy that extracts the similarities of past tasks' experiences is introduced. Our simulations show that the representation of acquired abstract policy is interpretable, that the modified LVQ successfully performs policy learning as well as generates abstract policy and that the application of generalized common abstract policy produces better results by more effectively guiding the agent when learning a new task.**

***Index Terms*—Transfer learning, Abstraction, Learning Vector Quantization, Reinforcement learning**