



DISSERTATION DEFENSE



VIVEK VEERIAH

On Addressing the Problem of Discovery for Reinforcement Learning

Tuesday, April 12, 2022

2:00pm – 4:00pm

[Virtual](#)

ABSTRACT: Reinforcement Learning (RL) studies the problem of an agent learning to interact with an environment to maximize its reward. Recent breakthroughs in Go, Atari, and Robotics are a result of combining advances from Deep Learning with RL algorithms, which has moved the field from hand-designing features to learning features directly from raw data. However, RL practitioners still need to define what predictive knowledge an agent needs to learn to maximize its performance, instead of the agent autonomously discovering them. Developing the ability to discover knowledge could allow the agents to learn as efficiently as humans. In this thesis, we propose methods that enable an agent to discover some forms of knowledge directly from experience.

First part of this thesis explores an approach to efficiently learn multiple chunks of knowledge, in the form of optimal behaviors to achieve goals in an environment, from a single trajectory of experience. Second part of the thesis explores approaches to discover useful knowledge of various forms to maximize the agent's task performance. It consists of three works: the first one introduces an architecture and an associated meta-gradient algorithm to discover predictive questions about the agent's experience to drive representation learning. Second work introduces another meta-gradient approach to discover temporal-abstractions in the form of options that enables a hierarchical RL agent to learn faster on new, unseen tasks. The third one presents an algorithm to select a small number of affordances in the form of actions or options from a continuous space to improve the performance of a model-based planning agent on hierarchical tasks. Final part of this thesis studies the problem of learning many option-conditional predictions in a complex environment. We demonstrate the extensive thought experiment from Ring (2021), which conjectured that high-level, abstract knowledge could be represented as layered predictions of agent's sensorimotor stream. We also introduce an approach to discover those option-conditional predictions which were previously hand-defined and demonstrate the feasibility of this approach on simple partially-observable environments.

Overall, the main contributions of this thesis are new learning algorithms to enable agents to identify and acquire useful predictive knowledge directly from experience.

CHAIR: Prof. Satinder Singh Baveja