

NATIONAL UNIVERSITY OF SINGAPORE

School of Computing

PH.D DEFENCE - PUBLIC SEMINAR

Title: Efficient hierarchical reinforcement learning through core task abstraction and context reasoning

Speaker: Mr Li Zhuoru

Date/Time: 11 January 2016, Monday, 04:00 PM to 05:30 PM

Venue: MR6, AS6-05-10

Supervisor : Dr Leong Tze Yun, Associate Professor, School of Computing

Abstract:

In this work, we propose a new approach to hierarchical reinforcement learning (HRL) and task decomposition. We target a class of reinforcement learning (RL) problems that use factored representation and contain multiple goals. We model each problem as a Markov decision process (MDP), and attempt to decompose the root problem into multiple smaller task MDPs by focusing on each task's relevant features. There are two main issues that we are trying to address. First, how to efficiently learn the transition dynamics for each task MDP, as well as for the root problem. Second, how the agent can select the right task to execute, or adjust its behavior in the presence of multiple active tasks. We propose context sensitive reinforcement learning (CSRL), which consists of two levels, (i) an efficient reinforcement learning algorithm based on core task abstraction (CTA) at the lower level, and (ii) a task reasoning mechanism at the hierarchical execution level.

We propose CTA methods for efficient learning of task dynamics by exploiting the contextual independences of the state variables conditional on the action executed. We analyse the sample complexity of the CTA methods, and show that our method can guarantee task level optimality. We demonstrate that our approach can achieve global optimum in problems where other HRL approaches fail to do so. Comparing to other state-of-the-art methods, the CTA methods have the following advantages. First, our methods guarantee task level optimality, which is stronger than recursive optimality in many other HRL methods. Second, our method can perform better in problems with unexpected changes in goals and can solve a class of new problems: those with unknown subtask goals. Third, the CTA methods require minimum exploration.

We propose two hierarchical execution mechanism: task as options (CSRL-SMDP) and dynamic adaptation (CSRL-Dynamic). For the former, we study the theoretical properties of hierarchical options learning, in which the tasks are modelled as options and learned using the CTA algorithms. Despite the learned model being non-stationary due to improving task

policies, we can derive the sample complexity. We propose CSRL-SMDP-SIMU, showing that by using simulation on the model learned by CTA methods, our framework can perform better than using SMDP-RMAX. For dynamic adaptation, we focus on the problems where multiple tasks can be active at the same time. In these scenarios, the agent should not just blindly execute the locally optimal policy, but it should also consider the other active tasks, and make necessary adaptations. The hierarchical execution mechanism has the following advantages. First, it requires a few samples to learn the effect of executing a task in a state, allowing the agent to learn a good policy in very few episodes. Second, CSRL-Dynamic supports multiple tasks being active simultaneously. Third, our framework can scale well in problems with over one million states, and to the best of our knowledge, we are not aware of any other HRL methods capable of handling problems of this scale.

We test our framework on a set of common benchmark problems in HRL literature and some of their extensions. We also experiment with robotic domains in complex virtual environments involving assistive care scenarios. We demonstrate the promising performance of our framework in all these challenging settings.