Module Based Reinforcement Learning for a Real Robot

Zsolt Kalmár, Csaba Szepesvári, and András Lőrincz

The behavior of reinforcement learning (RL) algorithms is best understood in completely observable, finite state- and action-space, discrete-time controlled Markov chains. In contrary, robot-learning domains are inherently infinite both in time and space, and are partially observable. Here we suggest to break up the problem into overlapping subtasks and design controllers for each of the subtasks. Then, operating conditions are attached to the controllers (together the controllers and their operating conditions are called modules), and possible additional features are designed to facilitate observability. A new discrete time-counter is introduced at the "module-level" that clicks only when a change in the value of one of the features is observed. The interactions of the controllers are analyzed (several tools are proposed for doing this), and RL is proposed to learn an optimal switching strategy for switching between the modules. The approach was tried out on a real-life robot. Several RL algorithms were compared by ANOVA and it was found that the model-based approach worked significantly better than the model-free approach. The learned switching strategy performed equally well as a handcrafted version. Moreover, the learned strategy seemed to exploit certain properties of the environment which could not have been seen in advance, supporting the view that adaptive algorithms are advantageous to non-adaptive ones in complex environments.