Options and Recognizers in Reinforcement Learning

Melanie Coggan

Reinforcement Learning (RL) involves learning how to behave – how to map situations to actions – without the assistance of a teacher providing examples or guidance. The learner, or *agent*, receives rewards from the environment and determines the best course of action to be the one that will maximize its expected reward over time. Although several algorithm have been proposed that will converge to optimal behaviour in small environments, these are inadequate for the larger, more realistic tasks we would like our agents to be able to perform. We present two notions aimed at improving the performance of agents in RL. The first, options, allow the agent to abtract away a series of one-step actions into a macro-action. The agent, which now only needs to consider actions at a higher level, can thus learn faster or in larger environments. However, the use of options can lead to convergence problems and suboptimal behaviour, issues that have not really been dealt with. We offer a simple technique for improving the performance of options in such cases. The second notion, recognizers, are used in conjunction with a new algorithm for learning. They distort behaviour by imposing a filter on the permited actions, and have the potential to work on much larger environments We show experimentally that this algorithm reaches the optimal behaviour policy, and does so faster than previously considered algorithms.