Bayesian Model-Based Reinforcement Learning

Nikos Vlassis (Luxembourg Centre for Systems Biomedicine)

March 25, 2011

Reinforcement Learning (RL) refers to a class of methods addressing the problem of sequential decision making under uncertainty. This problem can often be formalized as a Markov decision process (MDP), in which case optimal decision making corresponds to computing an optimal policy of this MDP. When the environment model (stochastic dynamics of the MDP) is known, an MDP optimal policy can be computed efficiently by dynamic programming or linear programming. On the other hand, when the environment model is unknown, the agent must learn its policy online by 'exploring' the environment. But what is the best way to explore an unknown environment?

In this talk I'll describe Bayesian Model-Based RL, an approach that in principle allows us solve the RL exploration problem optimally. The main idea is to form beliefs over the space of all possible MDP models, and then compute an optimal policy over that belief space. However, searching for policies over an infinitely dimensional belief space of models is a daunting task, making Bayesian RL a notoriously difficult problem to solve. In this talk I'll present some recent theoretical results and discuss some open research questions to this problem.