

Experience feedback about asynchronous policy iteration and observable time MDPs

Emmanuel Rachelson (University of Crete) 29/05/2009.

Representing time-dependency and temporal interactions in stochastic decision processes raises many questions, both from the modeling and the resolution points of view. In this talk, I will try to provide some feedback from my personal experience on these two topics. Several different options in the MDP (and related) literature have been adopted to model temporal stochastic decision processes. By focusing on the problems of time-dependency and concurrency, I will explain why Generalized Semi-Markov Decision Processes (GSMDPs) are a natural way of modeling temporal problems. In particular, we will point out an interesting link with Partially Observable MDPs which will emphasize the complexity of their resolution. Then, from the resolution point of view, I will introduce a methodology based on Asynchronous Policy Iteration and direct utility estimation designed for observable-time GSMDPs. Based on the experience feedback of this work, we will emphasize ideas concerning local value functions and policies and relate them to some recent advances in machine learning.