

64c Approximate Dynamic Programming Based Strategy for Markov Decision Problems in Process Control and Scheduling

Jay H. Lee

Most interesting problems in process control and scheduling can be formulated as a Markov Decision Problem (MDP). This includes real-time decision problems (e.g., feedback control and information-based rescheduling) that involve significant amounts of stochastic uncertainty. Optimal policy for MDPs can be derived by solving an associated stochastic dynamic programming (DP) problem. However, the computational complexity of stochastic dynamic programming is such that it is not a feasible approach for most practical problems. This is the reason why in practice one resorts to the popular approach of solving deterministic optimal control problem at each time step with feedback update (as in MPC), an approach which can be highly suboptimal. The framework of Approximate Dynamic Programming (ADP) offers some promising venue for pursuing the stochastic DP approach. In ADP, the computation is made feasible by pursuing a solution within a significantly restricted subset of the state space. The quality and computational complexity of the solution strongly depends on the choice of this subset. In ADP, this “working region” of the state space is identified by performing stochastic simulations of the closed-loop system with one or more known suboptimal policies. By solving the dynamic program within the state space defined by these simulations, one finds the best interpolated state trajectories for each encountered situation. This way, one can view the Bellman iteration in dynamic programming as a way to blend the simulated suboptimal policies in an optimal manner. In this presentation, we will describe how we can get around the ‘curse-of-dimensionality’ associated with the traditional solution approach to stochastic dynamic programming problems. We will then bring forth some key issues in applying the ADP approach to process control and scheduling problems. These include the choice of function approximator for the cost-to-go approximation and the restriction of the solution to the “working region” so that unreasonable extrapolations of the cost-to-go data are avoided. We will also identify some key situations where such an approach could offer a significant advantage over the existing approach. A number of examples drawn from process control and scheduling will be presented to make the case.