Decision-theoretic planning with non-Markovian rewards.


Summary: A decision process in which rewards depend on history rather than merely on the current state is called a decision process with non-Markovian rewards (NMRDP). In decision-theoretic planning, where many desirable behaviours are more naturally expressed as properties of execution sequences rather than as properties of states, NMRDPs form a more natural model than the commonly adopted fully Markovian Decision Process (MDP) model. While the more tractable solution methods developed for MDPs do not directly apply in the presence of non-Markovian rewards, a number of solution methods for NMRDPs have been proposed in the literature. These all exploit a compact specification of the non-Markovian reward function in temporal logic, to automatically translate the NMRDP into an equivalent MDP which is solved using efficient MDP solution methods. This paper presents NMRDPP (Non-Markovian Reward Decision Process Planner), a software platform for the development and experimentation of methods for decision-theoretic planning with non-Markovian rewards. The current version of NMRDPP implements, under a single interface, a family of methods based on existing as well as new approaches which we describe in detail. These include dynamic programming, heuristic search, and structured methods. Using NMRDPP, we compare the methods and identify certain problem features that affect their performance. NMRDPP’s treatment of non-Markovian rewards is inspired by the treatment of domain-specific search control knowledge in the TLPlan planner, which it incorporates as a special case. In the First International Probabilistic Planning Competition, NMRDPP was able to compete and perform well in both the domain-independent and hand-coded tracks, using search control knowledge in the latter.

Keywords: Markovian Decision Process; Non-Markovian Reward Decision Process Planner

http://www.jair.org/papers/paper1676.html