# First-Order Markov Decision Processes

Matthew Greig
mgreig@purdue.edu
Electrical and Computer Engineering
Purdue University

Markov Decision Processes (MDPs) [7] have developed lately as a standard method for representing uncertainty in decision-theoretic planning. Traditional MDP solution techniques have the drawback that they require an explicit state space, limiting their applicability to real-world problems due to the large number of world states occurring in such problems. Recent work addresses this drawback via compactly specifying the state-space in factored form as the set of possible assignments to a set of state-variables [1]. Such Factored MDPs (FMDPs) allow for easily representing exponentially large state spaces. The algorithms for planning using MDPs, however, still run in time polynomial in the size of the state-space or exponential in the number of state-variables, and methods for taking advantage of the structure of the state-space in planning with FMDPs have been proposed in the literature [1] [3].

Many complex systems are naturally represented using some form of relational representation. We propose to investigate methods of incorporating relational representations into the state-space of MDPs to allow their application to complex problems for which it may be impractical to use less expressive representations. The idea of using more expressive representations in MDPs is of course not new to the stochastic planning community—recent work by Boutilier et al. identifies limitations in the applicability of FMDPs to problems with complex state-spaces and proposes First-Order MDPs (FOMDP) based on the situation calculus [2]. This work still has some limitations—needing a separate description for each possible outcome of an action, and requiring all intermediate value functions computed to be piecewise constant. We are currently investigating alternatives to this approach that will avoid the first of these limitations by using Bayesian network (BN) techniques to represent action effects.

In FMDPs the action effects are compactly represented using a Two-Stage Temporal Bayesian Network [4] or Dynamic Bayesian Network (DBN) by exploiting the fact that the effects of an action on a state-variable often only depend on a small number of other state-variables. Recent work has developed Probabilistic Relational Models (PRMs) for representing probability distributions over relational domains, building on the ideas of BNs [5][6]. PRMs offer an alternative approach to representing action effects in relational MDPs. Our current interest lies in developing a method for representing action effects using a PRM, in much the same manner that DBNs are used. Of prime importance when developing this representation of the transition matrix is to ensure that the structure of the compact representation can also be exploited by solution techniques so as to keep planning tractable.

[1] Boutilier, Craig; Dearden, Richard; and Goldszmidt, Moises. Decision-theoretic planning: Structural assumptions and computational leverage. *Journal of Artificial Intelligence Research*, 11, 1999
[2] Boutiler, Craig; Reiter, Ray; and Price, Bob. Symbolic Dynamic Programming for First-Order MDPs. In *Proceedings IJCAI-01*. 2001
[3] Dean, Thomas and Givan, Robert. Model minimization in Markov decision processes. In *Proceedings AAAI-97*. 1997
[4] Dean, Thomas and Kanazawa, Keiji. A Model for Reasoning about Persistence and Causation. *Computational Intelligence*, pages 143-150, 1989.
[5] Koller, Daphne and Pfeffer, Avi. Probabilistic frame-based systems. In *Proceedings AAAI-98*. 1998
[6] Poole, David. Probabilistic Horn abduction and Bayesian networks. *Artificial Intelligence*, 64, pages 81-129. 1993
[7] Puterman, M. L. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, Wiley, New York, NY, 1994