# Envelope-based Planning in Relational MDPs

Natalia H. Gardiol and Leslie Pack Kaelbling

**Introduction:** A mobile robot acting in the world is faced with a large amount of sensory data and uncertainty in its action outcomes. Indeed, almost all interesting sequential decision-making domains involve large state spaces and large, stochastic action sets. We investigate a way to act intelligently as quickly as possible in domains where finding a complete policy would take a hopelessly long time. This approach, Relational Envelope-based Planning (REBP) [4] tackles large, noisy problems along two axes. First, describing a domain as a relational MDP (instead of as an atomic or propositionally-factored MDP) allows problem structure and dynamics to be captured compactly with a small set of probabilistic, relational rules. Second, an envelope-based approach to planning lets an agent begin acting quickly within a restricted part of the full state space and to judiciously expand its envelope as resources permit.

**The case for both probability and logical structure:** Quickly generating generating usable plans when the world abounds with uncertainty is an important and difficult enterprise. Consider the classic blocks world domain: the number of ways to make a stack of a certain height grows exponentially with the number of blocks on the table; and if the outcomes of actions are uncertain, the task becomes even more daunting. We want planning techniques that can deal with large state spaces and large, stochastic action sets since most compelling, realistic domains have these characteristics. We are investigating a method for planning in very large domains by using expressive rules to restrict attention to high-utility subsets of the state space.

Much of the work in traditional planning techniques centers on propositional, deterministic domains. See Weld's survey [9] for an overview of the extensive work in this area. Efforts to extend classical planning approaches into stochastic domains include mainly techniques that work with fully-ground state spaces [10, 1]. Conversely, efforts to move beyond propositional STRIPS-based planning involve work in mainly deterministic domains [5, 8].

But the world is not deterministic: for an agent to act robustly, it must handle uncertain dynamics as well as large state and action spaces. Markov decision theory provides techniques for dealing with uncertain outcomes in atomic-state contexts, and much work has been done in using structured representations to solve very large MDPs and some POMDPs [7, 2, 6]. While these techniques have moved MDP techniques from atomic-state representations to factored ones, they still operate in fully-ground state spaces.

In order to describe large stochastic domains compactly, we need relational structures that can represent uncertainty in the dynamics. Relational representations allow the structure of the domain to be expressed in terms of object *properties* rather than object identities and thus yield a much more compact representation of a domain than the equivalent propositional version can. Efficient solutions for probabilistic, first-order MDPs are difficult to come by, however.

**Planning with an envelope in relational domains:** Our approach strikes a balance along two axes: between fully ground and purely logical representations, and between straight-line plans and full MDP policies. We represent world dynamics by a compact set of relational rules, and we extend the envelope method of Dean et al. [3] to use these structured dynamics. An example rule, showing the dynamics of the action *move(A,B)* is shown below. Each rule schema contains the action name, precondition, and a set of probabilistic effects (or outcomes)

$$\begin{array}{ll}
\multicolumn{2}{l}{\mathsf{move}(A,B)} \\
pre: & (\mathsf{clear}(B,\mathsf{t}), \mathsf{hold}(\mathsf{nil}), \mathsf{height}(B,H), \mathsf{incr}(H,H'), \mathsf{clear}(A,\mathsf{t}), \mathsf{on}(A,C), \mathsf{broke}(\mathsf{f})) \\
\mathit{eff}: & \begin{array}{ll} [\,0.70\,] & (\mathsf{on}(A,B), \mathsf{height}(A,H), \mathsf{clear}(A,\mathsf{t}), \mathsf{clear}(B,\mathsf{f}), \mathsf{hold}(\mathsf{nil}), \mathsf{clear}(C,\mathsf{t})) \\ [\,0.30\,] & (\mathsf{on}(A,\mathsf{table}), \mathsf{clear}(A,\mathsf{t}), \mathsf{height}(A,H), \mathsf{hold}(\mathsf{nil}), \mathsf{clear}(\mathsf{C},\mathsf{t}), \mathsf{broke}(\mathsf{t})) \end{array}
\end{array}$$

Essentially, we want to come up with an initial, high-probability trajectory (an *envelope* of states) to the goal; then, we refine the policy by sampling from our model of the dynamics and incorporating nearby *fringe* states into the envelope. These fringe states correspond to possible action outcomes that are not already accounted for in the envelope. This incremental approach avoids the wild growth of purely propositional techniques by restricting attention to a useful subset of states.

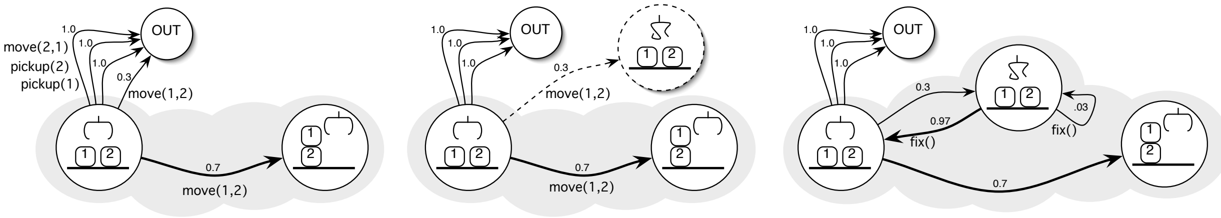


Figure 1: On the far left, an initial envelope corresponding to a simple one-step plan for stacking two blocks. The next two diagrams show how the envelope looks after some rounds of fringe sampling and envelope expansion.

**Conclusions:** Using the relational envelope method, we can take real advantage of relational generalization to produce good initial plans efficiently, and use envelope-growing techniques to improve the robustness of our plans incrementally as time permits. REBP is a planning system that tries to dynamically reformulate an apparently intractable problem into a small, easily handled problem at run time by restricting attention to a small, useful subset of a large MDP space.

However, there is plenty remaining to be done. For instance, current envelope-extension method is relatively undirected; it might be possible to diagnose more effectively which fringe states would be most profitable to add. Future directions include methods for deciding when to stop envelope growth and managing the eventual interleaving of envelope-growth and execution. Additionally the states in the envelope are essentially atomic; a key avenue to pursue is in exploiting the factored nature of relational representations to allow abstraction in the MDP model, with aggregate "states" in the MDP actually representing sets of states in the underlying world.

**Research Support:** This work was supported by an NSF Graduate Research Fellowship, by the Office of Naval Research contract #N00014-00-1-0298, and by NASA award #NCC2-1237.

**References:**

[1] Avrim L. Blum and John C. Langford. Probabilistic planning in the graphplan framework. In *5th European Conference on Planning*, 1999.

[2] Craig Boutilier, Raymond Reiter, and Bob Price. Symbolic dynamic programming for first-order MDPs. In *IJCAI*, 2001.

[3] Thomas Dean, Leslie Pack Kaelbling, Jak Kirman, and Ann Nicholson. Planning under time constraints in stochastic domains. *Artificial Intelligence*, 76, 1995.

[4] Natalia H. Gardiol and Leslie Pack Kaelbling. Envelope-based planning in relational MDPs. In *Advances in Neural Information Processing Systems 16 (NIPS-03)*, Vancouver, CA, 2004.

[5] B. Cenk Gazen and Craig A. Knoblock. Combining the expressivity of UCPOP with the efficiency of graphplan. In *Proc. European Conference on Planning (ECP-97)*, 1997.

[6] H. Geffner and B. Bonet. High-level planning and control with incomplete information using POMDPs. In *Fall AAAI Symposium on Cognitive Robotics*, 1998.

[7] Jesse Hoey, Robert St-Aubin, Alan Hu, and Craig Boutilier. Spudd: Stochastic planning using decision diagrams. In *Fifteenth Conference on Uncertainty in Artificial Intelligence*, 1999.

[8] J. Koehler, B. Nebel, J. Hoffmann, and Y. Dimopoulos. Extending planning graphs to an ADL subset. In *Proc. European Conference on Planning (ECP-97)*, 1997.

[9] Daniel S. Weld. Recent advances in AI planning. *AI Magazine*, 20(2):93–123, 1999.

[10] Daniel S. Weld, Corin R. Anderson, and David E. Smith. Extending graphplan to handle uncertainty and sensing actions. In *Proceedings of AAAI '98*, 1998.