Proceedings of the Nineteenth AAAI/SIGAI Doctoral Consortium

Probabilistic Planning with Reduced Models

Luis Pineda
School of Computer Science
University of Massachusetts
Amherst, MA 01003, USA
lpineda@cs.umass.edu

1 Introduction

Markov decision processes (MDP) (Puterman 1994) offer a rich model that has been extensively used by the AI community for planning and learning under uncertainty. Some applications include planning for mobile robots, network management, optimizing software on mobile phones, and managing water levels of river reservoirs. MDPs have polynomial complexity in the size of the state space, but the state space itself is exponential in the description size. Therefore, algorithms that try to find complete optimal plans are often impractical. Developing effective ways to tackle this complexity barrier is a challenging research problem.

Determinization-based algorithms for solving MDPs have gained popularity in recent years (Yoon *et al.* 2008; Teichteil-Königsbuch *et al.* 2010; Keyder and Geffner 2008), motivated by the surprising success of the FF-Replan solver (Yoon *et al.* 2007). The main idea is to generate a deterministic version of the underlying MDP and solve it using a classical deterministic planner, resulting in a *partial plan* for the original problem. When confronted by an unexpected state during plan execution, the planning process is repeated using the current state as the initial state. The advantage of this approach is its ability to quickly generate partial plans, particularly in intractable probabilistic domains.

Despite their success, determinization-based algorithms have drawbacks because they consider action outcomes in isolation. This leads to an overly optimistic view of the domain and can result in plans arbitrarily worse than optimal. Furthermore, even when optimal plans could be obtained using isolated outcomes, it is not always clear, nor intuitive, which outcomes should be included in the determinization.

In my work I introduce and study a more general paradigm in which the single-outcome variant of FF-Replan is just one extreme point on a *spectrum of MDP reductions* that differ from each other along two dimensions: (1) the number of outcomes per state-action pair that are fully accounted for in the reduced model, and (2) the number of occurrences of the remaining outcomes that are planned for in advance. Similar treatments of exceptional outcomes have been explored in fault-tolerant planning (Jensen *et al.* 2004; Domshlak 2013; Pineda *et al.* 2013).

The overall scope of my thesis is to develop robust and scalable approaches for concurrent planning and execution using reduced models, focusing on the following objectives:

1. Design new representations and a broad spectrum of reduced models that can be used in planning, and a disciplined way to perform model reduction and overcome the inherent drawbacks of determinization. These reduced models should allow for faster planning times when compared to solving full MDPs, but result in more robust plans than determinization-based methods.
2. Develop analytical approaches to evaluate the comprehensive value of a given reduced model and use it to guide the automated construction of good reduced models. The value of a reduced model should be measured by the performance of the resulting plans on the original problem.
3. Develop a continual planning paradigm that allows planning and plan execution to be conducted in parallel. Identify conditions under which this paradigm can produce near-optimal results and derive error bounds for the approach.
4. Perform a comprehensive evaluation and comparison of the approach with existing model reduction approaches, particularly determinization, and other approximate MDP solvers.
5. Develop an automated mechanism for solving reduced models via compilation to a standard MDP and using standard MDP solvers (using existing problem description languages such as PPDDL). A compilation scheme with complete support of a widely-used language such a PPDDL will make it easy to leverage existing probabilistic planners in the solution of reduced models.

2 Current progress

In a recent paper (Pineda and Zilberstein 2014) we formally defined a new family of MDP reductions and introduced the concepts of *primary outcomes*, namely, outcomes that are fully accounted for by the model, and *exceptional outcomes*, i.e., outcomes that the planner considers up to a maximum number of occurrences. An $\mathcal{M}_l^k$-reduction of an MDP is one in which up to k occurrences of exceptions are considered and the set of primary outcomes associated to any action in

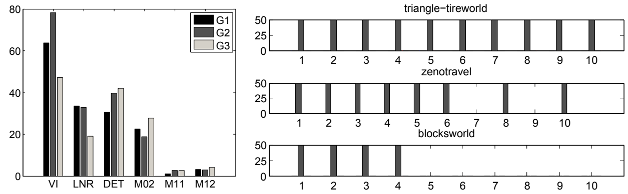

Figure 1: Left: Relative increase in total cost with respect to the optimal cost (racetrack domain). Right: Number of rounds ending up in success for 10 instances of various PPDDL domains.

the domain is of cardinality no greater than l. The benefit of using a well chosen $\mathcal{M}_l^k$-reduction is that the set of *reachable states* can become much smaller, which is desirable because the runtime of heuristic search algorithms for solving MDPs such as LAO* (Hansen and Zilberstein 1998; 2001) and LRTDP (Bonet and Geffner 2003) depends heavily on the size of the reachable state space.

We also proposed a continual planning paradigm to handle the case where more than k exceptions occur during plan execution, allowing execution to continue without delays. A benefit of this paradigm is that it is amenable to a precise evaluation of the benefits of planning with reduced models. We also investigated how to generate a good reduced model, be it a determinization or not, and showed that the choice of primary outcomes is non-trivial, even when reductions are limited to determinization ($\mathcal{M}_1^0$-reductions). As a baseline approach, we introduced a simple greedy algorithm that can produce good reduced models automatically.

We evaluated our continual planning paradigm on the well-known racetrack domain and showed that $\mathcal{M}_l^1$-reductions can be used to quickly compute near-optimal plans. The goal was to minimize the combined cost of planning and execution time, accounting for 1 second of execution time per action. Figure 1 (left) shows the relative increase in expected combined cost with respect to a theoretical lower bound—optimal cost ignoring planning time— for 6 planning methods. The best results were obtained using $\mathcal{M}_1^1$- and $\mathcal{M}_2^1$-reductions (M11 and M12, respectively).

Additionally, using an initial version of a PPDDL compiler, we applied our approach to several IPPC'08 domains using the IPC-style of evaluation: giving the planner a fixed amount of time to solve several rounds of the same problem. Figure 1 (right) shows that a planner using $\mathcal{M}_1^1$-reductions successfully solves many problem instances, results that are on par with those reported for state-of-the-art planners in these domains (Trevizan and Veloso 2012).

3 Research Plan

In the coming years I intend to tackle several open research questions related to planning with reduced models. One is the development of an efficient general-domain method that finds good reduced models for given problem instances; a potential approach is using sampling to estimate or bound the regret of removing outcomes from the primary set. Another important question is how to best redistribute the probabilities among the primary outcomes after the bound on exceptions is reached, possibly using various measures of structural similarity between primary and exceptional outcomes. I also plan to develop an anytime version of the continual planning algorithm. For the short term, my plan is to expand the support of the currently limited PPDDL compiler and perform a more thorough evaluation of the current continual planning method, gaining insight on how to best approach the more complex research problems ahead.

References

Blai Bonet and Hector Geffner. Labeled RTDP: Improving the convergence of real-time dynamic programming. In *Proceedings of the 13th International Conference on Automated Planning and Scheduling (ICAPS'03)*, pages 12–21, 2003.

Carmel Domshlak. Fault tolerant planning: Complexity and compilation. In *Proceedings of the 23rd International Conference on Automated Planning and Scheduling (ICAPS'13)*, pages 64–72, 2013.

Eric A. Hansen and Shlomo Zilberstein. Heuristic search in cyclic AND/OR graphs. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI'98)*, pages 412–418, Madison, Wisconsin, 1998.

Eric A. Hansen and Shlomo Zilberstein. LAO*: A heuristic search algorithm that finds solutions with loops. *Artificial Intelligence*, 129(1-2):35–62, 2001.

Rune M. Jensen, Manuela M. Veloso, and Randal E. Bryant. Fault tolerant planning: Toward probabilistic uncertainty models in symbolic non-deterministic planning. In *Proceedings of the 14th International Conference on Automated Planning and Scheduling (ICAPS'04)*, pages 335–344, 2004.

Emil Keyder and Hector Geffner. The HMDPP planner for planning with probabilities. *The ICAPS 3rd International Probabilistic Planning Competition (IPPC'08)*, 2008.

Luis Pineda and Shlomo Zilberstein. Planning under uncertainty using reduced models: Revisiting determinization. In *Proceedings of the 24th International Conference on Automated Planning and Scheduling (ICAPS'14)*, Portsmouth, USA, 2014.

Luis Pineda, Yi Lu, Shlomo Zilberstein, and Claudia V. Goldman. Fault-tolerant planning under uncertainty. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI'13)*, pages 2350–2356, Beijing, China, 2013.

Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., New York, NY, USA, 1994.

Florent Teichteil-Königsbuch, Ugur Kuter, and Guillaume Infantes. Incremental plan aggregation for generating policies in MDPs. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'10)*, pages 1231–1238, 2010.

Felipe W. Trevizan and Manuela M. Veloso. Short-sighted stochastic shortest path problems. *Proceedings of the 22nd International Conference on Automated Planning and Scheduling (ICAPS'12)*, pages 288–296, 2012.

Sung Wook Yoon, Alan Fern, and Robert Givan. FF-Replan: A baseline for probabilistic planning. In *Proceedings of the 17th International Conference on Automated Planning and Scheduling (ICAPS'07)*, pages 352–359, 2007.

Sungwook Yoon, Alan Fern, Robert Givan, and Subbarao Kambhampati. Probabilistic planning via determinization in hindsight. In *Proceedings of the 23rd National Conference on Artificial Intelligence (AAAI'08)*, pages 1010–1016, 2008.