Probabilistic Planning with Influence Diagrams

Junkyu Lee
Donald Bren School of Information and Computer Sciences
University of California, Irvine, Irvine, CA 92697, USA

Introduction

Motivations and Goal Graphical models provide a powerful framework for reasoning under uncertainty, and an influence diagram (*ID*) is a graphical model of a sequential decision problem that maximizes the total expected utility of a non-forgetting agent. Relaxing the regular modeling assumptions, an ID can be flexibly extended to general decision scenarios involving a limited memory agent or multi-agents. The approach of probabilistic planning with IDs is expected to gain computational leverage by exploiting the local structure as well as representational flexibility of influence diagram frameworks. My research focuses on graphical model inference for IDs and its application to probabilistic planning, targeting online *MDP*/*POMDP* planning as testbeds in the evaluation.

Related Works The computation of the maximum expected utility (*MEU*) of IDs involves the sum of utility functions as well as the product of conditional probability functions. Thus, most algorithms for solving IDs use a pair of probability and utility functions, called a potential, to trace the partial expected utility while marginalizing or combining functions (Jensen, Jensen, and Dittmer 1994). Previous works on search algorithms for solving IDs are depth-first AND/OR search exploiting the decomposition of graphical model (Marinescu 2010), and depth-first branch and bound with a heuristic generated by relaxing the constrained elimination ordering (Yuan, Wu, and Hansen 2010). Although depth-first search only requires linear memory, the space complexity for finding the optimal strategy is exponential in the length of history due to the non-forgetting assumption. Local search algorithms improving a subset of policies are proposed for *ID*s with limited memory in (Lauritzen and Nilsson 2001) and (Mauá and Cozman 2016).

On the other hand, translation based approaches allow applying existing inference algorithms to planning. MPD/POMDPs are translated as a mixture of finite horizon dynamic Bayesian networks and planning is formulated as likelihood maximization in (Toussaint, Harmeling, and Storkey 2006). Belief propagation algorithms are presented in (Liu and Ihler 2012) when utilities are multiplicative. A generic translation scheme between MEU and marginal *MAP* was shown in (Mauá 2016) , hence any marginal MAP inference could solve IDs. However, a naive application of translation often produces larger scope size functions and larger number of variables.

Online planning considers generating a single action at a time while interacting with a simulator in the loop of alternating planning and execution phases, hence it avoids exponential complexity for finding optimal decisions for all possible trajectories. The baseline algorithms and benchmark problems are available at the previous *ICAPS* planning competitions.

Current Progress

Generalized Dual Decomposition for *ID*s Generalized Dual Decomposition method (*GDD*) is a generalization of Lagrangian dual decomposition to L^p space with Hölder's inequality (Ping, Liu, and Ihler 2015). I presented a decomposition based approximate inference algorithm in (Lee, Ihler, and Dechter 2018). In this work, an ID is decomposed as a join-graph, a graph of clusters of subproblems with bounded number of variables (Dechter, Kask, and Mateescu 2002), Hence, the space and time complexity is bounded by the maximum cluster size. Then, the GDD is applied to the join-graph; equality constraints for the variables shared between clusters are augmented to the relaxed problem to enforce local consistency, and non-negative weights for each variable are distributed to bound MEU by Hölder's inequality. To optimize the bound, current implementation employs a gradient based local search, called *GDD-ID*. Since the bound from GDD-ID is formulated under the potential representation of IDs, it is free from translation and the bound can be computed by combining local expected utilities. The experimental evaluation on random factored MDP/POMDP instances demonstrated fast convergence to local optima, and generation of bounds for problem instances that an offline POMDP planner failed because of the large number of states.

Anytime AND/OR Search AND/OR search space for graphical model captures problem decomposition by conditioning, and it exhibits exponential improvement from OR search space (Dechter and Mateescu 2007). In collaboration with Radu Marinescu, I investigated several anytime AND/OR search algorithms. In (Lee et al. 2016), weighted best-first search algorithms were introduced to produce anytime marginal MAP solutions by inflating heuristic. The best-first scheme is in favor of expanding less number of nodes since evaluation of summation node involves high computational cost. On the other hand, the depth-first scheme can generate suboptimal solutions faster. Hybrids of both schemes can produce upper and lower bounds from each component and empirical evaluation demonstrated the effectiveness of best+depth-first search algorithms (Marinescu et al. 2017). The AND/OR search space for IDs can be defined by incorporating expected utilities as the value of each node. There are four types of nodes: the decision variable OR node for selecting the best action, the decision value AND node for combining expected utilities of subproblems, the chance variable AND node for computing the average expected utility of its children, and the chance value AND node for also combining expected utilities of subproblems.

Translation of Planning Benchmarks In earlier works, I presented translation schemes from a subset of *PPDDL* to dynamic Bayesian network for probabilistic conformant planning. Anytime AND/OR search algorithms for marginal MAP were applied to solve translated instances (Lee, Marinescu, and Dechter 2016). The marginal MAP guarantees the optimality of the plan in terms of both length and probability of reaching the goal. However, the translation didn't scale up because of the large number of constraints required to encode frame axioms, mutual exclusivity of actions, etc. Recently, I also implemented a translation from a subset of *RDDL* to grounded IDs to avoid the overhead of aforementioned constraints.

Research Plan

Short Term Plan My short term plan by Feb. 2018 is to integrate currently developed components into a single online planner. The basic architecture of the online planner can be described as follows. At the preparation step, the planner reads an RDDL planning domain and converts it to grounded IDs unrolled up to desired time steps. Then, GDD-ID generates heuristics and suboptimal policies. In the online planning loop, anytime best+depth-first AND/OR search traverses the search space until the end of planning phase and submits the best action; the search restarts after the execution phase.

There are several issues need to be addressed. The current AND/OR search algorithms are guided by static heuristic. Since planning problem requires look-ahead of sufficiently long time horizon, dynamic heuristic is worth being considered especially for POMDP problems of which the tree width grows linearly with the length of history. Most planning problems contain deterministic relations that often generate huge scope size factors which can't be even stored as a discrete table. Therefore, it is desirable to separate deterministic constraints and encode them over the join-graph independently of GDD-ID. Such deterministic information should be propagated by a stand-alone CP/SAT solver before addressing the GDD-ID heuristic to prune search space.

Long Term Plan The GDD bound for MEU could be improved; the local search could discover the global minimum or escape local optima more efficiently; and more importantly, the convex dual of the GDD bound could be formulated to provide more efficient algorithms with better theoretical guarantees. The bottleneck of AND/OR search is evaluation of summation subproblems rooted at chance nodes, so sampling methods could be integrated to search and reduce the computational burden as Monte Carlo tree search. The large scope size factors are unavoidable when IDs are translated from relational and symbolic languages like RDDL. Thus, the alternative approach of learning IDs from data would be also promising, which will generate compact and scalable models.

References

Dechter, R., and Mateescu, R. 2007. And/or search spaces for graphical models. *Artificial intelligence* 171(2-3):73–106.

Dechter, R.; Kask, K.; and Mateescu, R. 2002. Iterative join-graph propagation. In *UAI*, 128–136.

Jensen, F.; Jensen, F. V.; and Dittmer, S. L. 1994. From influence diagrams to junction trees. In *UAI*, 367–373.

Lauritzen, S. L., and Nilsson, D. 2001. Representing and solving decision problems with limited information. *Management Science* 47(9):1235–1251.

Lee, J.; Marinescu, R.; Dechter, R.; and Ihler, A. T. 2016. From exact to anytime solutions for marginal map. In *AAAI*, 3255–3262.

Lee, J.; Ihler, A.; and Dechter, R. 2018. Generalized dual decomposition for bounding maximum expected utility of influence diagrams with perfect recall. In *The AAAI-18 Workshop on Planning and Inference WS-12*.

Lee, J.; Marinescu, R.; and Dechter, R. 2016. Applying search based probabilistic inference algorithms to probabilistic conformant planning: Preliminary results. In *ISAIM*.

Liu, Q., and Ihler, A. 2012. Belief propagation for structured decision making. In *UAI*, 523–532.

Marinescu, R.; Lee, J.; Ihler, A. T.; and Dechter, R. 2017. Anytime best+ depth-first search for bounding marginal map. In *AAAI*, 3775–3782.

Marinescu, R. 2010. *A New Approach to Influence Diagrams Evaluation*. Springer London.

Mauá, D. D., and Cozman, F. G. 2016. Fast local search methods for solving limited memory influence diagrams. *Int. J. Approx. Reasoning* 68(C):230–245.

Mauá, D. D. 2016. Equivalences between maximum a posteriori inference in bayesian networks and maximum expected utility computation in influence diagrams. *Int. J. Approx. Reasoning* 68(C):211–229.

Ping, W.; Liu, Q.; and Ihler, A. T. 2015. Decomposition bounds for marginal map. In *NIPS*, 3267–3275.

Toussaint, M.; Harmeling, S.; and Storkey, A. 2006. *Probabilistic inference for solving (PO)MDPs*. Technical Report EDI-INF-RR-0934, University of Edinburgh.

Yuan, C.; Wu, X.; and Hansen, E. A. 2010. Solving multistage influence diagrams using branch-and-bound search. In *UAI*, 691–700.