Bounded Approximations for Linear Multi-Objective Planning under Uncertainty

(Extended Abstract)

Diederik M. Roijers [a] Joris Scharpff [b] Matthijs T.J. Spaan [a]

Frans A. Oliehoek [a] Mathijs M. De Weerdt [b] Shimon Whiteson [a]

{d.m.roijers, f.a.oliehoek, s.a.whiteson}@uva.nl
[a] *University of Amsterdam, The Netherlands*
{j.c.d.scharpff, m.t.j.spaan, m.m.deweerdt}@tudelft.nl
[b] *Delft University of Technology, The Netherlands*

Abstract

Planning under uncertainty poses a complex problem in which multiple objectives often need to be balanced. When dealing with multiple objectives, it is often assumed that the relative importance of the objectives is known a priori. However, in practice human decision makers often find it hard to specify such preferences exactly, and would prefer a decision support system that presents a range of possible alternatives. We propose two algorithms for computing these alternatives for the case of linearly weighted objectives. First, we propose an anytime method, approximate optimistic linear support (AOLS), that incrementally builds up a complete set of ϵ-optimal plans, exploiting the piecewise-linear and convex shape of the value function. Second, we propose an approximate anytime method, scalarised sample incremental improvement (SSII), that employs weight sampling to focus on the most interesting regions in weight space, as suggested by a prior over preferences. We show empirically that our methods are able to produce (near-)optimal alternative sets orders of magnitude faster than existing techniques, thereby demonstrating that our methods provide sensible approximations in stochastic multi-objective domains.

1 Introduction

Many real-world planning problems involve both uncertainty as well as multiple objectives. This type of problems is expressed naturally using the multi-objective Markov decision process (MOMDP) framework [4]. Following [4] we assume the existence of a *scalarisation function*, i.e. a function that translates multi-dimensional rewards into a scalar value. However, using such a function for planning requires complete knowledge of its parameters, or *weights*, beforehand. When such knowledge is not available, solving an MOMDP requires finding the set of optimal solutions for all possible weights.

In this paper[1], we consider only linear scalarisation functions. Therefore, it suffices to focus on the *Convex Coverage Set* of an MOMDP. Existing methods such as optimistic linear support (OLS) [5] exploit the value piecewise-linear convexity in the optimal value function over all weights, present when the scalarisation function is linear, to minimise the number of scalarised MDPs that need to be solved. However, OLS can only guarantee this when the scalarised MDPs are solved optimally and is therefore not directly applicable to large realistic planning problems.

2 Our Contributions

We propose new methods that rely on approximate MDP solving techniques to produce (near-)optimal CCSs. The first algorithm we propose is approximate optimistic linear support (AOLS) that, given an

[1] This is an extended abstract of our paper [3] at ICAPS 2014.

ϵ-bounded MDP approximation, is guaranteed to produce an ϵ-approximate CCS. The second algorithm, scalarised sample-based iterative improvement (SSII), exploits available prior knowledge on the distribution of weights and concentrates its effort within such a *prior*.[2] Although SSII can in practice produce a better approximate CCS over this prior, we cannot provide a bound on the CCS quality because SSII relies on sampling.

Both AOLS and SSII use an approximate single-objective solver as a subroutine. AOLS can use any solver. SSII requires an anytime method. In this paper we use UCT* [2] for both.

3 Evaluation

We performed experiments on instances of the maintenance planning problem [6], a 2-objective, probabilistic and numerical planning domain, and compared the optimal OLS method with our approximate AOLS and SSII. The optimal solutions have been computed using SPUDD [1] and approximations using the UCT* algorithm from PROST [2]. We compared the outcomes in terms of runtime, average CCS error ε_{exp} and maximal CCS ε_{max} error. The results are presented in Table 1 below.

			$[0, 1]$			$[0.5, 1]$		
Algorithm	Runtime	$\|CCS\|$	ε_{exp}	ε_{max}	%OPT	ε_{exp}	ε_{max}	%OPT
OLS + SPUDD	2390.819	9.250	-	-	-	-	-	-
AOLS + UCT* 0.01s	8.612	3.389	0.701	325.354	0.000	0.692	325.025	0.000
AOLS + UCT* 1s	19.940	4.111	0.119	65.668	0.167	0.117	65.426	0.167
AOLS + UCT* 10s	65.478	4.528	0.084	56.439	0.333	0.091	56.381	0.333
AOLS + UCT* 20s	165.873	5.694	0.044	38.667	0.417	0.048	38.627	0.417
SSII 1s, no prior	18.795	4.306	0.118	70.244	0.167	0.116	70.195	0.167
SSII 10s, no prior	59.336	3.889	0.061	51.800	0.333	0.068	51.747	0.333
SSII 1s, prior	17.892	3.944	0.221	95.189	0.000	0.125	61.667	0.167
SSII 10s, prior	59.154	4.083	0.141	71.290	0.083	0.057	43.006	0.333

Table 1: *Comparison of averaged performance of the algorithms presented in this paper for various parameters, shown for two regions of the scalarised reward space. Runtimes are in seconds, the expected error ε_{exp} and maximum error ε_{max} are relative to the optimum CCS and %OPT denotes the fraction of instances that were solved optimally.*

From the table we can conclude that both AOLS and SSII are able to produce reasonable, and sometimes even optimal, solutions much faster than OLS. Also, SSII is competitive with AOLS without exploiting additional knowledge but when SSII uses the prior it produces a slightly better CCS within the targeted weight region $w_1 \in [0.5, 1]$ and $w_2 = 1 - w_1$.

References

[1] Jesse Hoey, Robert St-Aubin, Alan Hu, and Craig Boutilier. SPUDD: Stochastic Planning Using Decision Diagrams. In *Proc. of the Fifteenth conference on Uncertainty in artificial intelligence*, pages 279–288, 1999.

[2] Thomas Keller and Malte Helmert. Trial-based Heuristic Tree Search for Finite Horizon MDPs. In *Proc. of the International Conference on Automated Planning and Scheduling*, 2013.

[3] Diederik M. Roijers, Joris Scharpff, Matthijs T.J. Spaan, Frans A. Oliehoek, Mathijs de Weerdt, and Shimon Whiteson. Bounded approximations for linear multi-objective planning under uncertainty. In *Proceedings of the 24th International Conference on Automated Planning and Scheduling (ICAPS)*, pages 262–270, June 2014.

[4] Diederik M. Roijers, Peter Vamplew, Shimon Whiteson, and Richard Dazeley. A Survey of Multi-Objective Sequential Decision-Making. *Journal of Artificial Intelligence Research*, 47:67–113, 2013.

[5] Diederik M. Roijers, Shimon Whiteson, and Frans Oliehoek. Linear Support for Multi-Objective Coordination Graphs. In *Proc. of the Autonomous Agents and Multi-Agent Systems conference*, 2014.

[6] Joris Scharpff, Matthijs T. J. Spaan, Leentje Volker, and Mathijs de Weerdt. Planning Under Uncertainty for Coordinating Infrastructural Maintenance. *Proc. of the International Conference on Automated Planning and Scheduling*, 2013.

[2] Note that such a prior expresses some — but not complete — knowledge about the weights.