Redesigning Stochastic Environments for Maximized Utility

Sarah Keren, Avigdor Gal, Erez Karpas
Technion — Israel Institute of Technology

Luis Pineda, Shlomo Zilberstein
University of Massachusetts Amherst

Abstract

We present the Utility Maximizing Design (UMD) model for optimally redesigning stochastic environments to achieve maximized performance. This model suits well contemporary applications that involve the design of environments where robots and humans co-exist an co-operate, *e.g.*, vacuum cleaning robot. We discuss two special cases of the UMD model. The first is the equi-reward UMD (ER-UMD) in which the agents and the system share a utility function, such as for the vacuum cleaning robot. The second is the goal recognition design (*GRD*) setting, discussed in the literature, in which system and agent utilities are independent. To find the set of optimal modifications to apply to a UMD model, we present a generic method, based on heuristic search. After specifying the conditions for optimality in the general case, we present an admissible heuristic for the ER-UMD case. We also present a novel compilation that embeds the redesign process into a planning problem, allowing use of any off-the-shelf solver to find the best way to modify an environment when a design budget is specified. Our evaluation shows the feasibility of the approach using standard benchmarks from the probabilistic planning competition.

Introduction

We are surrounded by environments that are designed and manipulated with the intention of maximizing some benefit. Hospitals may be designed to minimize the daily distance covered by staff, supermarkets are constantly rearranged to make sure users buy as much as possible, airports may be designed to increase passenger spending, computer networks are structured to maximize message throughput, *etc.*

Common to all these environments is that their design is controllable. Such environments can be designed and often later redesigned to accommodate a specific objective. In addition, such environments need to account for different forms of uncertainty.

We aim at providing a generic model to support the offline design of such environments. Therefore, we present a model of *Utility Maximizing Design* in which a problem of redesigning non-deterministic environments in order to maximize system utility is specified. Non-determinism is expressed by stochastic outcomes of actions performed by agents. The setting we propose takes as input a stochastic environment, a set of allowed modifications, a set of constraints and a system utility criteria. It then finds an optimal set of modifications to apply to the environment for maximizing expected utility under the constraints.

Example 1 *Consider Figure 1(left), where a vacuum cleaning robot is placed in a living room. The utility of the robot may be expressed in various ways; it may try to clean an entire room as quickly as possible or cover as much space as possible before its battery runs out. In any case, (re)moving a piece of furniture from or within the room (Figure 1(center)) may increase the robot's utility. Accounting for uncertainty, there may be specific locations in which the robot tends to slip, ending up in a different location than intended. Increasing friction,* e.g.*, by introducing a high friction tile (Figure 1(right)), may reduce the probability of undesired outcomes in particular locations. Both types of modifications are applied offline (since such robots typically perform their task unsupervised) and should be applied economically in order to maintain usability of the environment.*

The proposed *Utility Maximizing Design* (UMD) model is a general model whose instantiations provide common grounds for comparative analysis and identification of efficient methods for special cases. A key observation with the UMD model is that utility may differ between the system and the agents acting in it. While Example 1 illustrates an *Equi-Reward* UMD(ER-UMD) case where agent and system share a utility function, earlier works on goal recognition design (Keren, Gal, and Karpas 2014; Wayllace et al. 2016)(*GRD*) assumed optimal agents while the system aims at minimizing expected goal recognition time. We show that different assumptions on the relation between agent and system utility induce different solution techniques.

To support various assumptions on the model, the *Utility Maximizing Design* (UMD) model we propose consists of four elements. The *environment component* describe the possible settings in which agents act by applying stochastic actions. The *agents component* describes the different types of agents that may act in the environment. The *system component* specifies both the way the system accumulates rewards and the ways by which it can redesign the environment. Finally, the initial environment describes the environment that is modified for maximizing system utility.

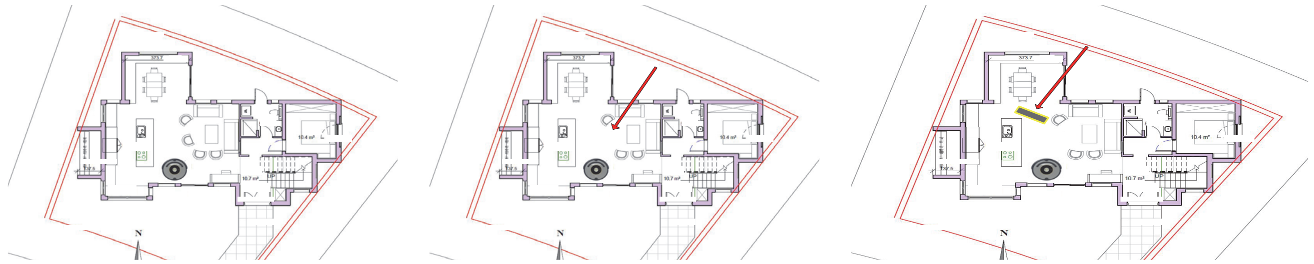

Figure 1: An example of a *Utility Maximizing Design* problem

In this work we assume a fully observable stochastic setting and use Markov decision processes (Bellman 1957; Mausam 2012) to model the agent environment. We offer a general solution to the redesign problem by using heuristic search that yields optimal design strategies when using admissible heuristics. We formulate the conditions for admissibility for UMD settings and propose a heuristic based on simplifications of the environment, which we show to be admissible for the ER-UMD case but not for *GRD*. In addition, inspired by the compilation techinque of Göbelbecker et al. (2010), we exploit the alignment of system and agent utility in ER-UMD settings, to show a way to piggyback on the search for optimal policy to find an optimal set of modifications. Finally, for settings where practicality is prioritized over optimality, we discuss ways to acquire sub-optimal solutions.

The contributions of this work are threefold. First, we describe a new general model, namely *Utility Maximizing Design*, which involves the offline redesign stochastic environments for improving utility and show how goal recognition design is a special case of this setting. In particular, changing probability distributions offers a wide range of subtle (more realistic) modifications to be applied to a model, *e.g.*, reducing the probability of a slipping rather than eliminating it altogether. Second, we present a general method for solving UMD problems using informed heuristic search and specify the conditions under which an optimal solution can be found. Finally, for the special case where agent and system utility function is the same, which we refer to as *equi-reward UMD* (ER-UMD) we formulate and compare three approaches for finding an optimal set of modifications to apply given a budget, namely an informed search approach with an admissible heuristic, a compilation-based method that embeds design into the definition of a planning problem, and a sub-optimal solver.

We evaluate our work using probabilistic benchmarks from the International Planning Competitions, where a variety of stochastic shortest path MDPs are introduced (Bertsekas 1995). Our evaluation aims at measuring the effect of a budget on the utility of a ER-UMD problem, as well as the performance of the different techniques for solving a ER-UMD problem. We used six PPDDL domains from the probabilistic tracks of the sixth and eighth International Planning Competition[1] (IPPC06 and IPPC08) representing stochastic shortest path MDPs with uniform action cost: Box World (IPPC08), Blocks World (IPPC08), Exploding Blocks World (IPPC08), Triangle Tire (IPPC08) and Elevators (IPPC06).

With the exception of Exploding Blocks World, results show the reduction in expected cost increases with the budget increase, demonstrating the applicability of the UMD problem. When comparing solution performance, in all case, the use of informed search outperformed the exhaustive approach on all domains. However, the dominating heuristic approach varied between domains, encouraging a further exploration of various heuristic techniques for solving UMD problems.

[1] http://icaps-conference.org/index.php/main/competitions

References

Bellman, R. 1957. A markovian decision process. *Indiana Univ. Math. J.* 6:679–684.

Bertsekas, D. P. 1995. *Dynamic programming and optimal control*, volume 1. Athena Scientific Belmont, MA.

Göbelbecker, M.; Keller, T.; Eyerich, P.; Brenner, M.; and Nebel, B. 2010. Coming up with good excuses: What to do when no plan can be found. *Cognitive Robotics* (10081).

Keren, S.; Gal, A.; and Karpas, E. 2014. Goal recognition design. In *ICAPS Conference Proceedings*.

Mausam, A. K. 2012. Planning with markov decision processes: an ai perspective. *Morgan & Claypool Publishers*.

Wayllace, C.; Hou, P.; Yeoh, W.; and Son, T. C. 2016. Goal recognition design with stochastic agent action outcomes. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI 2016)*.