Dynamic Execution of Temporal Plans with Sensing Actions and Bounded Risk

Pedro H. Santana, Brian C. Williams
Massachusetts Institute of Technology, Computer Science and Artificial Intelligence Laboratory, MERS
32 Vassar St. Room 32-224, Cambridge, MA 02139, {psantana,williams}@mit.edu

Abstract

This thesis focuses on the problem of temporal planning under uncertainty with explicit safety guarantees, which are enforced by means of chance constraints. We aim at elevating the level in which operators interact with autonomous agents and specify their desired behavior, while retaining a keen sensitivity to risk. Instead of relying on unconditional sequences, our goal is to allow contingent plans to be dynamically scheduled and conditioned on observations of the world while remaining safe. Contingencies add flexibility by allowing goals to be achieved through different methods, while observations allow the agent to adapt to the environment. We demonstrate the usefulness of our chance-constrained temporal planning approaches in real-world applications, such as partially observable power supply restoration and collaborative human-robot manufacturing.

1 Introduction

Trusting autonomous systems with critical missions in real-world situations requires that they develop a keen sensitivity to risk and incorporate uncertainty into their decision-making. By risk, we mean the common notion of the probability of some failure event (crashing against obstacles, missing deadlines, running out of battery, crossing no-fly zones, etc.) happening. In the context of this thesis, failure denotes the event of an agent operating under temporal and resource constraints failing to meet its deadlines; violating one or more plan constraints; or not achieving all of its original goals. It draws great inspiration from interactions with groups in the Woods Hole Oceanographic Institution (WHOI), NASA's Jet Propulsion Laboratory (JPL), and National ICT Australia (NICTA), all with the responsibility of operating multimillion dollar systems in extreme environments.

Due to the lack of guarantees that plans will be carried out within their very stringent safety requirements, the current practice for ensuring mission safety generally requires groups of engineers to reason over a very large number of potential decisions and scenarios that might unfold during execution, which is a challenging, time-consuming, and error-prone process. Given the overwhelming number of possible scenarios, one opts, in many cases, to follow the "safest", most predictable strategy, in which the impact of uncertainty in the plan is limited. Such "safe" precomputed sequences of actions, however, tend to be far from ideal in terms of utility or brittle to disturbances due to their inability to adapt to the environment. Extensions of this planning paradigm have been proposed in order to improve robustness, such as conditioning execution on the state of the world and dynamic task scheduling. However, mission operators resist to incorporate those improvements, due to a lack of explicit guarantees of correctness in terms of risk of mission failure.

This thesis aims at developing tools that will enable autonomous systems to operate well (as defined by some measure of utility) while ensuring that hard safety guarantees are met. More specifically, it addresses the problem of extracting chance-constrained temporal execution policies in planning domains with uncertainty. We do so by guiding agents to make active use of sensing information in order to improve their belief states, while providing correctness guarantees expressed in the form of chance constraints. An important advancement towards value-optimizing temporal planning was the introduction of Temporal Plan Networks (TPN's) [Kim *et al.*, 2001], which allow preferences to be placed over the choices on a Disjunctive Temporal Problem (DTP)[Stergiou and Koubarakis, 2000]. Later, Temporal Plan Networks with Uncertainty (TPNU's) [Effinger, 2012] extended TPN's by allowing uncontrollable choices to be incorporated into the plan description. The Probabilistic Temporal Plan Networks (pTPN) we introduced in [Santana and Williams, 2014] extend TPNU's by considering probabilistic models for uncontrollable choices and allowing chance constraints to be imposed on the violation of temporal constraints.

1.1 Contributions

Departing from previous conservative risk-minimization strategies, this thesis extends the chance-constrained approach of [Ono *et al.*, 2012] to high-level temporal activity planning in partially observable domains, where the system must actively incorporate sensing information into its decision-making. We improve upon prior work on Constrained Partially Observable Markov Decision Processes (POMDP's) [Undurti and How, 2010; Poupart *et al.*, 2015] handling particular forms of chance constraints by deriving new ways of dynamically computing mission risk without the assumption that constraint violations are observable or cause execution to be terminated. We increase the temporal robustness of our chance-constrained policies by scheduling activities in real-time, as opposed to choosing a feasible schedule beforehand. Our algorithms propagate families of feasible, least commitment schedules in the form of dispatchable temporal networks, where *dispatchable* depends on whether all temporal durations are controllable or if some of them can only be observed. Finally, since plan execution in partially observable domains requires accurate real-time inference based on sensing data, we propose novel algorithms capable of learning expressive hybrid dynamical models automatically from data, therefore improving the quality of real-time state estimation in support of plan execution.

2 Progress towards goals and next steps

In [Santana and Williams, 2014], we extend the current notions of weak and strong temporal plan consistency to a chance-constrained setting, while providing efficient algorithms for determining (or refuting) them. Weak and strong consistency are useful concepts when planning missions for agents whose embedded hardware has very limited computation and telecommunication power. Chance-constrained weak consistency (CCWC) is a useful concept for missions where agents operate in static or slow changing environments after an initial scouting mission. Chance-constrained strong consistency (CCSC), on the other hand, tries to determine the existence of a solution that, with probability greater than some threshold, will succeed irrespective of the outcomes of uncertainty. Despite its conservativeness, strong consistency is appealing to mission operators, for strongly consistent policies require little to no embedded sensing and decision-making; and reduce or completely eliminate the need to coordinate between multiple agents. We efficiently explore the space of feasible temporal plans while keeping risk bounded by leveraging a "diagnostic" approach, which continuously learns subsets of conflicting constraints and generalizes them to a potentially much larger set of pathological scenarios. For the problem of extracting a strongly consistent policy from a contingent plan description, our numerical results showed significant gains in scalability for our approach when compared to current methods based on chronological search.

Current work focuses on dynamic execution of chance-constrained temporal plans. Towards this goal, we have recently derived new methods for computing risk dynamically and allowing it to depend on real-time sensor observations, which were used to propose RAO*, and extension of AO* [Nilsson, 1982] that computes optimal, deterministic, chance-constrained policies for finite horizon POMDP's. We demonstrate the usefulness of our results by using RAO* to solve chance-constrained instances of the partially observable power supply restoration domain [Thiébaux and Cordier, 2001], a challenging problem of practical interest that was beyond previous approaches. We also validate our results on the context of collaborative human-robot manufacturing, in which humans and robots should collaborate in order to achieve common goals. By means of computing optimal chance-constrained policies, we ensure that robots have a fluid interaction with their human coworkers while limiting the frequency of unwanted events, such as missing manufacturing deadlines, or having robots needlessly disturb human colleagues. The next step is to extend RAO* with the ability to propagate dispatchable temporal networks, therefore achieving the goal of computing chance-constrained, dynamically consistent (optimal) temporal policies.

Dynamic execution of temporal plans with sensing actions requires accurate real-time inference about the system currently under control to be performed in real-time. In support of that, we propose in [Santana *et al.*, 2015] the first algorithm capable of learning Probabilistic Hybrid Automata (PHA), an expressive class of hybrid models with guarded transitions, purely from data, therefore precluding the need of human effort in describing these models. In our experiments, we demonstrate that PHA models have negligible computational overhead and yield significant performance improvement in hybrid filtering, when compared with widely-used Jump Markov Linear System (JMLS) models.

References

[Effinger, 2012] Robert T Effinger. *Risk-minimizing program execution in robotic domains*. PhD thesis, MIT, 2012.

[Kim *et al.*, 2001] P. Kim, B. Williams, and M. Abramson. Executing reactive, model-based programs through graph-based temporal planning. In *IJCAI*, 2001.

[Nilsson, 1982] Nils J Nilsson. *Principles of artificial intelligence*. Springer, 1982.

[Ono *et al.*, 2012] Masahiro Ono, Yoshiaki Kuwata, and J Balaram. Joint chance-constrained dynamic programming. In *CDC*, pages 1915–1922, 2012.

[Poupart *et al.*, 2015] Pascal Poupart, Aarti Malhotra, Pei Pei, Kee-Eung Kim, Bongseok Goh, and Michael Bowling. Approximate linear programming for constrained partially observable markov decision processes. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence*, 2015.

[Santana and Williams, 2014] P. Santana and B. Williams. Chance-constrained consistency for probabilistic temporal plan networks. In *24th ICAPS*, 2014.

[Santana *et al.*, 2015] Pedro Santana, Spencer Lane, Brian Williams, and Carlos Forster. Learning hybrid models with guarded transitions. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence*, 2015.

[Stergiou and Koubarakis, 2000] K. Stergiou and M. Koubarakis. Backtracking algorithms for disjunctions of temporal constraints. *Artificial Intelligence*, 120(1):81–117, 2000.

[Thiébaux and Cordier, 2001] S. Thiébaux and M.-O. Cordier. Supply restoration in power distribution systems - a benchmark for planning under uncertainty. In *Proc. 6th ECP*, 2001.

[Undurti and How, 2010] Aditya Undurti and Jonathan P. How. An online algorithm for constrained pomdps. In *IEEE International Conference on Robotics and Automation*, pages 3966–3973, 2010.