Journal Track Paper Abstracts

Jörg Hoffmann and Alan Fern

Editors

Simple Regret Optimization in Online Planning for Markov Decision Processes

Zohar Feldman and Carmel Domshlak (Technion, Haifa, Israel)

We consider online planning in Markov decision processes (MDPs). In online planning, the agent focuses on its current state only, deliberates about the set of possible policies from that state onwards and, when interrupted, uses the outcome of that exploratory deliberation to choose what action to perform next. Formally, the performance of algorithms for online planning is assessed in terms of simple regret, the agent's expected performance loss when the chosen action, rather than an optimal one, is followed.

To date, state-of-the-art algorithms for online planning in general MDPs are either best effort, or guarantee only polynomial-rate reduction of simple regret over time. Here we introduce a new Monte-Carlo tree search algorithm, BRUE, that guarantees exponential-rate and smooth reduction of simple regret. At a high level, BRUE is based on a simple yet non-standard state-space sampling scheme, MCTS2e, in which different parts of each sample are dedicated to different exploratory objectives. We further extend BRUE with a variant of "learning by forgetting." The resulting parametrized algorithm, BRUE(alpha), exhibits even more attractive formal guarantees than BRUE. Our empirical evaluation shows that both BRUE and its generalization, BRUE(alpha), are also very effective in practice and compare favorably to the state-of-the-art.

This work has been partially supported by the Air Force Office of Scientific Research, USAF, under grant number FA8655-12-1-2096.

Energy Efficient Execution of POMDP Policies

Marek Grzes (University of Kent, UK), Pascal Poupart (University of Waterloo, Canada), Xiao Yang (Amazon, Canada) and Jesse Hoey (University of Waterloo, Canada)

Recent advances in planning techniques for partially observable Markov decision processes have focused on online search techniques and offline point-based value iteration. While these techniques allow practitioners to obtain policies for fairly large problems, they assume that a non-negligible amount of computation can be done between each decision point. In contrast, the recent proliferation of mobile and embedded devices has lead to a surge of applications that could benefit from state of the art planning techniques if they can operate under severe constraints on computational resources. To that effect, we describe two techniques to compile policies into controllers that can be executed by a mere table lookup at each decision point. The first approach compiles policies induced by a set of alpha vectors (such as those obtained by point-based techniques) into approximately equivalent controllers, while the second approach performs a simulation to compile arbitrary policies into approximately equivalent controllers. We also describe an approach to compress controllers by removing redundant and dominated nodes, often yielding smaller and yet better controllers. Further compression and higher value can sometimes be obtained by considering stochastic controllers. The compilation and compression techniques are demonstrated on benchmark problems as well as a mobile application to help persons with Alzheimer's to way-find. The battery consumption of several POMDP policies is compared against finite-state controllers learned using methods introduced in this paper. Experiments performed on the Nexus 4 phone show that finite-state controllers are the least battery consuming POMDP policies.

Envisioning the Qualitative Effects of Robot Manipulation Actions Using Simulation-Based Projections

Lars Kunze (University of Birmingham, UK) and Michael Beetz (University of Bremen, Germany)

Autonomous robots that are to perform complex everyday tasks such as making pancakes have to understand how the effects of an action depend on the way the action is executed. Within Artificial Intelligence, classical planning reasons about whether actions are executable, but makes the assumption that the actions will succeed (with some probability). In this work, we have designed, implemented, and analyzed a framework that allows us to envision the physical effects of robot manipulation actions. We consider envisioning to be a qualitative reasoning method that reasons about actions and their effects based on simulation-based projections. Thereby it allows a robot to infer what could happen when it performs a task in a certain way. This is achieved by translating a qualitative physics problem into a parameterized simulation problem; performing a detailed physics-based simulation of a robot plan; logging the state evolution into appropriate data structures; and then translating these sub-symbolic data structures into interval-based first-order symbolic, qualitative representations, called timelines. The result of the envisioning is a set of detailed narratives represented by timelines which are then used to infer answers to qualitative reasoning problems. By envisioning the outcome of actions before committing to them, a robot is able to reason about physical phenomena and can therefore prevent itself from ending up in unwanted situations. Using this approach, robots can perform manipulation tasks more efficiently, robustly, and flexibly, and they can even successfully accomplish previously unknown variations of tasks.

This work has been supported by the EU FP7 project RoboHow (grant number 288533) and by the EU FP7 project STRANDS (grant number 600623).

Distributed Heuristic Forward Search for Multi-agent Planning

Raz Nissim (Yahoo Labs, Israel) and Ronen Brafman (Ben-Gurion University, Israel)

This paper deals with the problem of classical planning for multiple cooperative agents who have private information about their local state and capabilities they do not want to reveal. Two main approaches have recently been proposed to solve this type of problem – one is based on reduction to distributed constraint satisfaction, and the other on partial-order planning techniques. In classical single-agent planning, constraint-based and partial-order planning techniques are currently dominated by heuristic forward search. The question arises whether it is possible to formulate a distributed heuristic forward search algorithm for privacy-preserving classical multi-agent planning. Our work provides a positive answer to this question in the form of a general approach to distributed state-space search in which each agent performs only the part of the state expansion relevant to it. The resulting algorithms are simple and efficient – outperforming previous algorithms by orders of magnitude – while offering similar flexibility to that of forward-search algorithms for single-agent planning. Furthermore, one particular variant of our general approach yields a distributed version of the A* algorithm that is the first cost-optimal distributed algorithm for privacy-preserving planning.

The authors were supported in part by ISF grant 933/13 and the Lynn and William Frankel Center for Computer Science.