

Planning under Uncertainty with Weighted State Scenarios

Walraven, Erwin; Spaan, Matthijs T. J.

Publication date

2015

Document Version

Final published version

Published in

AAAI 2015 Fall Symposium on Sequential Decision Making for Intelligent Agents

Citation (APA)

Walraven, E., & Spaan, M. T. J. (2015). Planning under Uncertainty with Weighted State Scenarios. In *AAAI 2015 Fall Symposium on Sequential Decision Making for Intelligent Agents* (pp. 93-94). American Association for Artificial Intelligence (AAAI).
<http://www.aaai.org/ocs/index.php/FSS/FSS15/paper/view/11689/11514>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Planning Under Uncertainty with Weighted State Scenarios

Erwin Walraven

Delft University of Technology
Mekelweg 4, 2628 CD
Delft, The Netherlands

Matthijs T. J. Spaan

Delft University of Technology
Mekelweg 4, 2628 CD
Delft, The Netherlands

Abstract

External factors are hard to model using a Markovian state in several real-world planning domains. Although planning can be difficult in such domains, it may be possible to exploit long-term dependencies between states of the environment during planning. We introduce weighted state scenarios to model long-term sequences of states, and we use a model based on a Partially Observable Markov Decision Process to reason about scenarios during planning. Experiments show that our model outperforms other methods for decision making in two real-world domains.

The full version of this extended abstract has been published in the Proceedings of the Conference on Uncertainty in Artificial Intelligence (Walraven and Spaan 2015).

Introduction

Markov Decision Processes can be used to model planning problems in various domains (Puterman 1994). However, external factors can be difficult to predict, which makes it hard to define a Markovian model with the right state features and an appropriate level of detail (Witwicki et al. 2013). Decision making in such domains can be challenging since it is often difficult to estimate transition probabilities and it can be hard to predict uncertain events.

An example of a domain with hard-to-model external factors is a smart grid, in which wind influences the availability of renewable electricity supply. The most severe problems in electricity grids occur during peak hours, when the demand of consumers is high and renewable generation is interrupted (Moura and De Almeida 2010). In such cases, the supply of electricity may not be sufficient to satisfy the demand. This problem can be addressed by scheduling flexible loads of consumers in such a way that they can be supplied when renewable supply is available, but it can be hard to model uncertain renewable supply using a Markovian state.

In this work we propose a framework which enables agents to reason about future states in domains with hard-to-model external factors. Rather than defining a state-based transition model, we use scenarios to reason about states that may be encountered in the future.

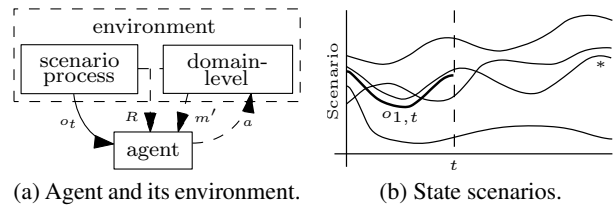


Figure 1: Agent-environment interaction and state scenarios.

Weighted State Scenarios

We assume that an agent interacts with an environment as shown in Figure 1a. The environment consists of a process for which the domain-level state changes to m' after executing an action a . There is another process for which a compact Markovian model does not exist, called the scenario process. We assume that an agent observes a numerical-valued state o_t of this process at time t , which we call a state observation. Actions do not influence the state transitions of the scenario process.

The rewards received by the agent are dependent on both the domain-level state, as well as the state of the scenario process. Therefore, it is required to reason about future states of the scenario process during planning. We use scenarios which encode long-term sequences of numerical states of this process, and we assume that a large set containing possible scenarios can be obtained from historical data, as visualized in Figure 1b. When interacting with the environment, the first t state observations are encoded in a vector $o_{1,t}$, which is represented by the bold line. We use the vector to assign weights to scenarios, corresponding to the likelihood that a scenario predicts the future. In order to do this, we compare the vector $o_{1,t}$ with each scenario and we compute the sum of squared errors. The weights assigned to the scenarios are inversely proportional to the computed errors.

The illustrative example in Figure 1b shows four scenarios and the state observations until time t . If the scenarios are accurate, then it is most likely that the scenario marked with the star provides an accurate prediction of the future. For this scenario the computed error is relatively small, and therefore the weight assigned to this scenario is high compared to other scenarios.

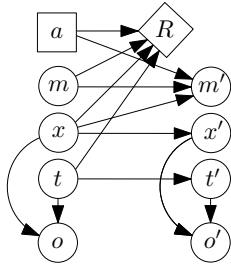


Figure 2: Dynamic Bayesian network of the POMDP model.

Planning with Scenarios

We present a model that is based on a Partially Observable Markov Decision Process (POMDP) to reason about scenarios during planning, for which the corresponding dynamic Bayesian network is shown in Figure 2. Each state s of the POMDP can be factored into a tuple $s = (m, x, t)$, where m is the observable domain-level state of the environment, x is the scenario of the scenario process, and t is a variable representing the current time. State variable m represents the state of multiple flexible loads in a smart grid, but our model can also be applied to other domains. State variable x is not fully observable since this would imply certainty regarding the future states of the scenario process. In state $s = (m, x, t)$ the observation o corresponds to the state in scenario x at time t . Before selecting the next action, we derive a probability distribution over state variable x from the weights associated with the scenarios. Then we invoke the POMCP algorithm for planning (Silver and Veness 2010), which is a Monte-Carlo planning algorithm that is capable of dealing with a large number of states. Our POMCP implementation samples scenarios based on weights that have been assigned.

Experiments: Scheduling in Smart Grids

Our experiments focus on a smart grid domain, where loads of consumers have to be scheduled in such a way that wind power is used as much as possible and power consumption from the grid is minimized. In each experiment we consider a collection of flexible loads (e.g., household appliances, electric vehicles), and we formulate the scheduling problem using our POMDP model. Scenarios define wind speed during the day, and have been derived from historical data that we obtained from a wind farm in Spain.

In the experiments we compare our method with a consensus algorithm (Ströhle et al. 2014) and optimal omniscient schedules that have been computed using mixed-integer programming. We also make a comparison with an MDP planner which models wind speed as a Markov chain. The results are shown in Figure 3, where the cost of optimal omniscient schedules is represented by 1, and the distributions visualize the cost increase compared to the cost of these omniscient schedules. Outliers have been omitted for readability reasons. We use two variants of our planner. We consider the situations in which the algorithm encounters a scenario that is already present in the set (1), and we consider the situation in which the algorithm only encounters new scenarios (2).

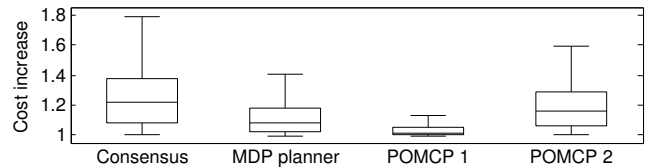


Figure 3: Comparison of scheduling methods in smart grids.

Our method outperforms other methods in case it encounters known scenarios, and performs competitively in case new scenarios are encountered. In the paper (Walraven and Spaan 2015) we present similar results for a planning problem in financial stock markets.

Conclusions

We proposed a method to reason about future states in domains with hard-to-model external factors. We defined weighted scenarios that represent long-term sequences of states, and we presented a POMDP model to reason about scenarios during planning. Our model has been shown to outperform other methods in smart grids and stock markets, which can also be considered as a motivating example of promising applications for sequential decision making. In future work we aim to investigate how wind prediction methods based on analog ensembles (Van den Dool 1989) can be combined with our work.

Acknowledgements

This work is funded by the Netherlands Organisation for Scientific Research (NWO), as part of the URSES program.

References

- Moura, P. S., and De Almeida, A. T. 2010. The role of demand-side management in the grid integration of wind power. *Applied Energy* 87(8):2581–2588.
- Puterman, M. L. 1994. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons.
- Silver, D., and Veness, J. 2010. Monte-Carlo Planning in Large POMDPs. In *Advances in Neural Information Processing Systems*, 2164–2172.
- Ströhle, P.; Gerding, E. H.; De Weerd, M. M.; Stein, S.; and Robu, V. 2014. Online Mechanism Design for Scheduling Non-Preemptive Jobs under Uncertain Supply and Demand. In *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems*, 437–444.
- Van den Dool, H. M. 1989. A New Look at Weather Forecasting through Analogues. *Monthly Weather Review* 117(10):2230–2247.
- Walraven, E., and Spaan, M. T. J. 2015. Planning under Uncertainty with Weighted State Scenarios. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence*.
- Witwicki, S.; Melo, F. S.; Capitán, J.; and Spaan, M. T. J. 2013. A Flexible Approach to Modeling Unpredictable Events in MDPs. In *Proceedings of the International Conference on Automated Planning and Scheduling*, 260–268.