

Efficient exploration and learning of structure in factored-state MDPs

Carlos Diuk, Michael L. Littman, Alexander L. Strehl

Abstract: Algorithms for provably experience-efficient exploration of MDPs have been generalized to factored-state MDPs specified by dynamic Bayes nets (DBNs). The Factored E^3 (Kearns and Koller 1999), Factored Rmax (Guestrin et al. 2002), and Factored MBIE (Strehl and Littman 2006) algorithms are all known to behave near optimally, with high probability, in all but a polynomial number of timesteps. Here, the polynomial depends on the size of the conditional probability tables of the DBNs that specify the transition probabilities, which can be exponentially smaller than the total number of states. A shortcoming of all these algorithms is that, while they learn the necessary conditional probabilities, they must be provided with the DBN structures in advance. We provide a method for learning the DBN structures and prove it provides an efficient learning algorithm when combined with factored RMax.

References

Kearns, M. J., Koller D. 1999. Efficient reinforcement learning in factored MDPs. In Proceedings of the 16th International Joint Conference on Artificial Intelligence (IJCAI), pages 740--747.

Guestrin, C. Patrascu R., Schuurmans, D. 2002. Algorithm-directed exploration for model-based reinforcement learning in factored MDPs. In Proceedings of the International Conference on Machine Learning. Pages 235--242.

Strehl, A. L., Littman, M. L., 2006. Model-based reinforcement learning in factored MDPs. Submitted.