POMDP compression and decomposition via belief state analysis

Li, Xin

Date of Award:
2009

POMDP Compression and Decomposition via Belief State Analysis

LI Xin

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Principal Supervisor: Dr. William K. CHEUNG

Hong Kong Baptist University

April 2009

Abstract

Partially observable Markov decision process (POMDP) is a commonly adopted mathematical framework for solving planning problems in stochastic environments. However, computing the optimal policy of POMDP for large-scale problems is known to be intractable, where the high dimensionality of the underlying belief state space is one of the major causes. This thesis focuses on studying two different paradigms, namely *POMDP compression* and *POMDP decomposition*, for addressing the POMDP's tractability issue.

To reduce POMDP's complexity via compression (c.f. dimension reduction), belief compression and value-directed compression are the two representative approaches recently proposed in the literature. However, both bear their own limitations in terms of policy quality and computational efficiency. In this thesis, the use of non-negative matrix factorization (NMF) for belief compression is proposed, which is then further integrated with a value-directed compression framework for ensuring the quality of the computed policies as far as possible. The proposed hybrid approach has been tested empirically based on some commonly used benchmark problems. It is found that the proposed approach is effective in speeding up the convergence of the underlying value iteration process. Also, the policies obtained are of superior quality when compared with those obtained using belief compression and value-directed compression alone. However, they are still not as good as those obtained without the proposed compression applied.

Given that the main cause of the limited policy accuracy is due to the proposed value-directed framework being ill-posed, an orthogonality constraint is proposed to be incorporated into the framework which then leads to the need of a novel orthogonal NMF. Updating rules corresponding to the proposed orthogonal NMF are derived with their convergence

theoretically proved. Also, it has been empirically demonstrated that this orthogonal NMF is effective in making a trade-off among (1) reducing the POMDP's dimension, (2) maintaining the orthogonality of the NMF projection matrix, and (3) ensuring the optimally of the computed policies.

In addition to POMDP compression, decomposing a POMDP problem is another direction where a conquer-and-divide approach is taken. In this thesis, application of data clustering techniques to POMDP's belief state space is proposed for "decomposing" the POMDP. The clustering criterion function is designed so that the transition probabilities between clusters are minimized as far as possible, and thus reduce the loss in formulation accuracy incurred due to the decomposition. Via experiments, it has been shown that such a belief clustering technique can readily be combined with non-linear and linear belief compression methods to tackle the POMDP's tractability.

To further study the scalability of our proposed compression framework and the compressibility of different POMDP problems, the application of interior-point gradient acceleration to the proposed orthogonal NMF and the use of an eigenvalue analysis are proposed. Again via experiments, the former is shown to be effective in further reducing the NMF overhead needed for the compression and the latter is validated to be more or less consistent to the best ratio of POMDP compression which can be empirically obtained for different benchmark problems.

A number of future research directions are also proposed in the thesis, including (1) optimizing the belief clustering quality and extending it to support hierarchical decomposition, (2) alternative POMDP decomposition techniques derived based on the eigenvalue-based analysis over the generalized transition functions, (3) extending the decomposition setting to a multi-agent with the hope to obtain a better and more dynamic belief clustering algorithm, and (4) extending our current approaches to support online learning of POMDPs.

Keywords: POMDP, belief compression, non-negative matrix factorization, value-directed compression, belief clustering

Table of Contents