ARTICLE IN PRESS

Neurocomputing ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Efficient solutions of interactive dynamic influence diagrams using model identification

He Wu, Jian Luo *

School of Information Science and Technology, Xiamen University, Xiamen 361005, China

ARTICLE INFO

Article history:
Received 14 September 2015
Received in revised form
16 June 2016
Accepted 27 July 2016
Communicated by Dr. Grana Manuel

Keywords:
Interactive dynamic influence diagram
Multi-agent dynamic decision making
Mutual information
Model identification

MSC:
91A35
90C40

ABSTRACT

Interactive dynamic influence diagram (I-DID) is one of the graphical frameworks for sequential decision making in partially observable environment. Subject agent in I-DID maintains beliefs over not only physical states of the environment, but also over models of the other agents. Consequently, solving I-DIDs suffers from the exponential growth of models ascribed to the other agents over time. Previous methods to solve I-DIDs aim at clustering equivalent models by comparing the entire or partial policy trees of the candidate models, which is time-consuming. In this paper, we present a new method for further reducing the model space by identifying the true model of the other agent and pruning the other irrelevant models. Toward this, we use an information-theoretic method—mutual information to measure the relevance between the candidate models and the true model in terms of predicted and observed actions of the other agent. We construct a dynamic Bayesian network to learn the value of parameters needed in the computation of mutual information. This approach bounds the model space by containing only the true model of the other agent. We evaluate our approach on multiple problem domains and empirically demonstrate the efficiency in solving I-DIDs.

1. Introduction

Sequential decision making is a research hotspot in multi-agent system [1], especially when the environment is partially observable to the agents. Some frameworks formalize the problem of agent decision making in uncertain settings that is shared with other agents. One of these frameworks is partially observable Markov decision process (POMDP) [2–4]. POMDP is used to guide an agent's actions regardless of the existence of other agents. Then, interactive partially observable Markov decision process (I-POMDP) generalizes POMDP to multi-agent settings providing a framework for decision making in multi-agent environments [5]. I-POMDP takes a decision-maker's perspective to make decision by including other agents' models in the state space together with the physical environment. The other agents' models encompass all the information including their beliefs, capabilities and preferences.

More recently, interactive dynamic influence diagram (I-DID) may be viewed as the graphical model of I-POMDP [6–8]. I-DID represents the world by mapping various variables into chance, decision, utility nodes and the dependency links between them. I-DID may be used to compute the policy of an agent given its initial belief as the agent acts and observes in an environment that is populated by other agents. This graphical model provides an intuitive language for modeling the problem structure thereby serving as an very important tool to enable multi-agent planning. Due to the compact representation, I-DID is proved to have computational advantage over I-POMDP [6]. First, I-DID is built by factored representation of state space and solution of I-DID applies the lookahead and backup methods exploiting the conditional independence. Furthermore, allowing the graphical representation of lower level model also makes the improvement of computation. However, it still suffers from the curse of both dimensionality and history. This is because I-DID's state space includes not only the traditional physical states, but also the models of other agents. As these agents act and observe, I-DIDs must track the evolution of the models. Consequently, the number of candidate models grows exponentially over time. That makes I-DIDs suffer from the curse of history that afflicts the modeling (subject) agent, also from the update of the modeled agents. This is further complicated when it comes to more agents and more nested levels of the state space.

The computational complexity of solving I-DIDs may limit their application to the real world. To settle this problem, some methods are presented to reduce the dimensionality of the interactive state space [9,10]. The main idea is to cluster candidate models which share the same or similar policy tree. These models are called behaviorally equivalent models [11–13].

In this article, we present a new approach for solving I-DIDs online and further speed up the solution. We compress the model

* Corresponding author.
E-mail addresses: wuhe@stu.xmu.edu.cn (H. Wu), jianluo@xmu.edu.cn (J. Luo).

http://dx.doi.org/10.1016/j.neucom.2016.07.052
0925-2312/

Please cite this article as: H. Wu, J. Luo, Efficient solutions of interactive dynamic influence diagrams using model identification, Neurocomputing (2016), http://dx.doi.org/10.1016/j.neucom.2016.07.052

space maximumly by including only the true model and excluding all the other models. In order to do this, we propose a method to identify the true model in the model space by using an information-theoretical tool—mutual information(MI).[1] We seek to dig some relevance between the candidate models and the true model. Given predictions of the candidate models and previous observations of other agents' actions, we use Bayesian learning to update the likelihood of each model. Through the learned classification function, we obtain MI which provides a natural way of measuring of the dependence between the candidate models and the true model. Then gradually prune the candidate models with the lowest MI, we finally identify the true model and reduce the model space. Our approach speeds up solution of I-DID and makes it scalable to solve the complex domain.

The remainder of this paper is structured as follows. In Section 2, we discuss some related works. In Section 3, we briefly review the graphical model of I-ID and its dynamic form—I-DID as well as its solutions. In Section 4, we present the definition of relevant model and mutual information. We build a dynamic Bayesian network to learn the parameters which are used in the computation of mutual information. In Section 5, we propose the formal algorithm for the solution and discuss the computational savings of our method. We then provide, in Section 6, the experimental results that show the performance of our algorithm and compare it with previous solutions. At last, in Section 7, we discuss and conclude this paper and remark some future works.

2. Related work

Interactive influence diagram (I-ID), the static part of I-DID, is a type of networks of influence diagram (NID) [14]. The graphical formalism contributes to a growing line of work on multi-agent decision making including multi-agent influence diagram (MAID) [15]. These formalisms explicitly model the structure presented in real world by decomposing the problem into chance and decision nodes, and the dependencies between them. They open up a promising area of research that represents multi-agent interaction problems more transparently. However, the applicability of both NIDs and MAIDs is limited to static single play games. MAIDs represent multi-agent games objectively and analyze the game from an external viewpoint adopting Nash equilibrium as a solution concept. I-DIDs differ from these representations by modeling decision making from an individual agent's perspective (subject agent) and explicitly representing other agents' candidate models in the unique “model node”. I-DIDs update the subject agent's beliefs over these models by “model update link” over multiple time steps. In recent years, some works are presented based on I-DID. Communicative I-DID is proposed to enable agents to have the ability of communication [16–18]. TC-IDID studies time-critical dynamic decision making by providing time index to each node in I-DID [19,20].

The complexity of solving I-DID is always a problem needed to be settled [21–23]. A method is presented to solve I-DIDs approximately by clustering the models using K-means and selecting a representative set made of K models from the clusters so as to limit the number of candidate models to a constant at each time step [24]. Behaviorally equivalent (BE) is first applied to solve I-POMDPs exactly, then employed to I-DIDs, becoming state-of-the-art technique to solve I-DIDs [9]. This efficient method groups together those models whose policy tree is completely the same and selects only one model in each BE sets thereby reducing the cardinality of the model space. Furthermore, a new method called discriminative model updates (DMU) is proposed based on BE to update only those models which result in predictive behaviors that are distinct from others in the updated model space [25]. This approach improves on BE as it generates less possible models prior to selection at each time step. As BE becomes a mature and effective method to solve I-DIDs, Zeng et al present a novel approach for identifying behaviorally equivalent models of agents. This approach focuses on partial policy trees for comparison and determines how much of the policy trees to consider instead of the traditional method which needs to compare the whole policy trees. This approach is an approximating solution of I-DIDs which trades off solution quality for efficiency [26]. Later, Zeng et al further improve this approach and develop an incremental comparison method to identify BE models [27]. This approach is the most efficient solution to I-DIDs so far, to the best of our knowledge.

3. Background

We first introduce interactive influence diagrams (I-IDs), then their extensions to dynamic settings, I-DIDs for two agents interactions.

3.1. Interactive influence diagrams

An influence diagram (ID) has three types of nodes—the chance, decision and utility nodes. The chance nodes S, O_i, A_j model the physical state, the observations of agent i, and the actions of agent j, respectively. The decision node A_j models the actions of agent i and the utility node R_i models i's reward function. Besides these usual nodes, I-IDs include a new type of node named model node ($M_{j,l-1}$, hexagonal node in Fig. 1(a)). $M_{j,l-1}$ denotes the other agent j that is modeled in the strategy level $l-1$.

In I-IDs, agent i is modeled in a higher level l. The level represents the reasoning between the two agents—what does agent i think that agent j thinks that i thinks. Similar recursive modeling can be found in decision theory [28,29]. The recursion terminates when the level 0 agent is an ID or just a probability distribution that doesn't have the ability to model others. We note that the probability distribution over the chance node, S, and the model node together represent agent i's belief over its interactive state

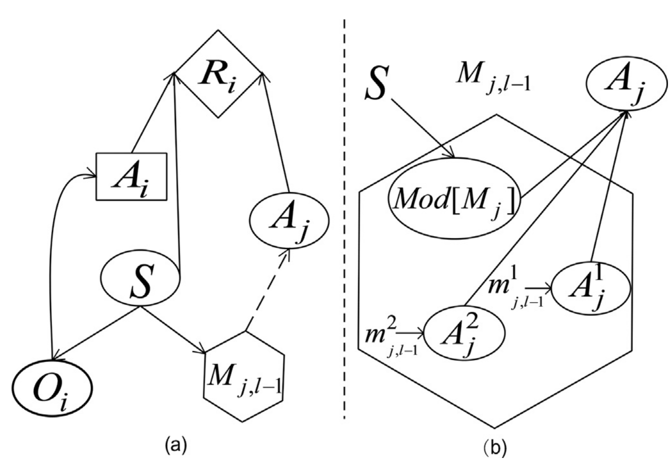

Fig. 1. A level l I-ID and model node. (a) A generic level l I-ID for agent i situated with one other agent j. The hexagon is the model node ($M_{j,l-1}$) and the dashed arrow is the policy link; (b) Members of model node could be I-ID themselves or IDs $\left(m^1_{j,l-1}, m^2_{j,l-1}\right)$ whose decision nodes are mapped to the corresponding chance nodes (A_j^1, A_j^2).

[1] We use MI as a terse expression for both, ‘mutual information’ and ‘value of mutual information’. Appropriate usage will be self-evident.

Please cite this article as: H. Wu, J. Luo, Efficient solutions of interactive dynamic influence diagrams using model identification, Neurocomputing (2016), http://dx.doi.org/10.1016/j.neucom.2016.07.052