Compact Approximations of Mixture Distributions for State Estimation in Multiagent Settings

(Extended Abstract)

Prashant Doshi
Dept. of Computer Science
University of Georgia
Athens, GA 30602
pdoshi@cs.uga.edu

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Multiagent Systems

General Terms

Theory, Performance

Keywords

state estimation, particle filters, representation

1. INTRODUCTION

In order to act rationally, an agent must track the state of the environment over time. In the presence of other agents who themselves act, observe, and update their beliefs the agent must track not only the physical state but also the possible states of others. This is because others' actions may affect the evolution of the physical state and the agent's payoffs. One approach is to generalize the Bayes filter to multiagent settings, in which an agent tracks the evolution of the *interactive state* [2]. In practice, the estimation may be carried out using the *interactive PF* (I-PF) [2] that generalizes the PF to the multiagent setting.

Applications of I-PF are confined to simple problems with a small number of discrete physical states. This is because a large number of particles must be used, at the expense of computational efficiency, to achieve good approximation quality. This limitation is acute for the I-PF because the interactive state space from which the particles are sampled tends to get large as it includes the nested beliefs of other agents. Often, the structure of the problem allows further statistical efficiency. Doshi [1] marginalizes out some dimensions of the interactive space and updates those dimensions analytically while sampling the particles from the remaining space. This procedure is equivalent to *Rao-Blackwellising* the I-PF.

Although efficient, marginalization generates beliefs represented as mixtures whose components grow exponentially over time. Thus, the belief representation does not exhibit a closed form and gradually becomes intractably large. To address this, we derive *compact* approximate representations for updated beliefs. Specifically, we desire an accurate representation that consumes as little memory as possible and is computationally efficient to evaluate in comparison to the original. Snelson and Ghahramani [3] investigate a similar problem in the context of supervised learning. We apply their general framework for approximation to multiagent state estimation.

Our approach is to fit the marginal posterior with a mixture of a constant number, K, densities of exactly the same form as the component densities in the original belief. Fitting involves minimizing the distance between the approximate and exact posteriors, for which we utilize the KL-Divergence. We show that this representation is both compact and flexibly accurate. It enables the application of the Rao-Blackwellised I-PF for efficiently approximating the estimation process.

Cite as: Compact Approximations of Mixture Distributions for State Estimation in Multiagent Settings, (Extended Abstract), Prashant Doshi, *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009),* Decker, Sichman, Sierra and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary, pp. 1207–1208
Copyright © 2009, International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org), All rights reserved.

2. BACKGROUND

We consider an agent i, that is interacting with one other agent j. Consider a space of physical states S. We will call agent j's belief over S a 0^{th} level belief, $b_{j,0}$. Additionally, j can be modeled by specifying its set of actions A_j, set of observations Ω_j, transition and observation functions, T_j and O_j, a reward function R_j, and an optimality criterion OC_j. j's 0^{th} level model is then $\theta_{j,0} = \langle b_{j,0}, A_j, \Omega_j, T_j, O_j, R_j, OC_j \rangle$, which is a POMDP. Agent i's 1^{st} level beliefs are defined over the physical states and the 0^{th} level models of j. This enriched state space has been called a set of **interactive** states. Thus, let $IS_{i,1}$ denote a set of interactive states, $IS_{i,1} = S \times \Theta_{j,0}$, and $IS_{i,0} = S$, where $\Theta_{j,0}$, is the set of 0^{th} level models of agent j. Let us rewrite $\theta_{j,0}$ as, $\theta_{j,0} = \langle b_{j,0}, \hat{\theta}_j \rangle$, where $\hat{\theta}_j \in \hat{\Theta}_j$, is the agent j's *frame*. To keep matters simple, we limit our focus to i's singly nested beliefs.

State estimation in multiagent settings is complex because of two reasons: First, the prediction of how the physical state changes must be made based on the predicted actions of the other agent. The probabilities of other's actions are based on its models. Second, changes in other's models have to be included in the update. Specifically, update of the other agent's beliefs due to its new observation must be included. $b^t_{i,1}(is^t) \stackrel{def}{=} Pr(is^t | a_i^{t-1}, o_i^t, b_{i,1}^{t-1})$:

$$b^t_{i,1}(is^t) = \alpha \int_{is^{t-1}:\hat{\theta}_j^{t-1}=\hat{\theta}_j^t} b^{t-1}_{i,1}(is^{t-1}) \sum_{a_j^{t-1}} Pr(a_j^{t-1}|\theta_{j,0}^{t-1}) \times O_i(s^t, a_i^{t-1}, a_j^{t-1}, o_i^t)\, T_i(s^{t-1}, a_i^{t-1}, a_j^{t-1}, s^t) \sum_{o_j^t} \delta_D(SE_{\hat{\theta}_j^t}(b_{j,0}^{t-1}, a_j^{t-1}, o_j^t) - b_{j,0}^t)\, O_j(s^t, a_i^{t-1}, a_j^{t-1}, o_j^t)\, d\, is^{t-1} \quad (1)$$

where α is the normalization constant, δ_D is the Dirac-delta function, $Pr(a_j^{t-1}|\theta_{j,0}^{t-1})$ is the probability that a_j^{t-1} is Bayes rational for agent described by $\theta_{j,0}^{t-1}$, and $SE_{\hat{\theta}_j^t}$ denotes update of the complete belief using the transition and observation functions in $\hat{\theta}_j^t$.

We may decompose the state estimation (Eq. 1) into two factors, one of which represents the update of the belief over the physical

states, and the other is the update of the belief over the other agent's models conditioned on a physical state:

$$Pr(is^t|a_i^{t-1}, o_i^t, b_{i,1}^{t-1}) = \int_{is^{t-1}} \sum_{a_j^{t-1}} Pr(s^t|a_i^{t-1}, a_j^{t-1}, o_i^t, s^{t-1}) \\ Pr(\theta_{j,0}^t|s^t, a_i^{t-1}, a_j^{t-1}, o_i^t, \theta_{j,0}^{t-1})\ Pr(a_j^{t-1}|\theta_{j,0}^{t-1})\ b_{i,1}^{t-1}(is^{t-1}) \quad (2)$$

The I-PF [2] propagates a sampled representation of agent i's nested beliefs, over time. The I-PF may become computationally intensive and the problem is exacerbated when we consider large physical state spaces. One way to alleviate this difficulty is to simulate the belief update over the physical state space using the traditional PF while updating agent i's belief over j's beliefs conditioned on the physical state, as exactly as possible [1]. N particles are sampled from agent i's belief over the physical state space, $b_{i,1}^{t-1}(s^{t-1})$, resulting in a set of particles $\{s^{(n)}\}_{n=1}^N$ that together approximate the belief over the large state space. The belief over the complete interactive state space is then given by the following set of particles, $\{(s^{(n)}, b_{i,1}^{t-1}(\theta_{j,0}^{t-1}|s^{(n)})\}_{n=1}^N$, which represent $b_{i,1}^{t-1}(is^{t-1})$ in Eq. 2. On substitution and simplification, we get:

$$b_{i,1}^t(is^t) \approx \frac{\alpha}{N} \sum_{a_j^{t-1}} \sum_{n=1}^{N} \rho_{a_j^{t-1}}^{(n)}(s^t)\ \kappa_{a_j^{t-1}}^{(n)}(\theta_{j,0}^t|s^t) \quad \text{where:}$$

$$\rho_{a_j^{t-1}}^{(n)}(s^t) \stackrel{def}{=} O_i(s^t, a_i^{t-1}, a_j^{t-1}, o_i^t)\ T_i(s^{(n)}, a_i^{t-1}, a_j^{t-1}, s^t)$$

$$\kappa_{a_j^{t-1}}^{(n)}(\theta_{j,0}^t|s^t) \stackrel{def}{=} \sum_{o_j^t} O_j(s^t, a_j^{t-1}, a_i^{t-1}, o_j^t) \int_{\theta_{j,0}^{t-1}} \delta_D(SE_{\hat{\theta}_j^t} \\ (b_{j,0}^{t-1}, a_j^{t-1}, o_j^t) - b_{j,0}^t)\ Pr(a_j^{t-1}|\theta_{j,0}^{t-1})\ b_{i,1}^{t-1}(\theta_{j,0}^{t-1}|s^{(n)})\ d\theta_{j,0}^{t-1}$$

Starting with a single prior density, the conditional posterior $\kappa_{a_j^{t-1}}^{(n)}$ is a mixture of multiple conditional densities – at most $|\Omega_j|$. In general, there will be at most $|A_j||\Omega_j|$ densities that make up i's level 1 belief, after one step of the belief update. After t steps, there will be a maximum of $(|A_j||\Omega_j|)^t$ distinct components in the mixture [1]. As the number of mixture components grows exponentially with the updates, compact representations of the mixture are needed.

3. COMPACT APPROXIMATION

Our aim is to find a compact approximation, $Pr(\theta_{j,0}^t|s^t, \mathbf{w})$ parameterized by $\mathbf{w}$. Thus, we seek the best $\mathbf{w}$ that approximates $\kappa_{a_j^{t-1}}^{(n)}$ parameterized by $\mathbf{v}$, as closely as possible and which saves on memory. Note that $\mathbf{v}$ is a vector of parameters, $\mathbf{v}$=$\{v_1, v_2, \ldots, v_{|\Omega_j|}\}$, where each v_p parameterizes a component density in $\kappa_{a_j^{t-1}}^{(n)}$.

We begin by obtaining X samples, $\mathcal{D} = \{\theta_{j,0}^{t,(x)}\}_{x=1}^X$, from the mixture. At this stage, we could obtain $\mathbf{w}$ as the point estimate of $\mathbf{v}$ from the X samples using the standard maximum likelihood or maximum-a-posteriori approach. However, the resulting density would be an inaccurate approximation of the original mixture density. Instead, we compute $Pr(v|\mathcal{D})$, using one of many accurate parameter estimation approaches such as Laplace's.

We will seek the approximation that minimizes a distance measure such as the KL-Divergence (KLD) from $\kappa_{a_j^{t-1}}^{(n)}(\theta_{j,0}^t|s^t, \mathbf{v})$:

$$\hat{\mathbf{w}} = \arg\min_{\mathbf{w}} \text{KLD}(\kappa_{a_j^{t-1}}^{(n)}(\theta_{j,0}^t|s^t, \mathbf{v})||Pr(\theta_{j,0}^t|s^t, \mathbf{w})) \\ = \arg\max_{\mathbf{w}} \int_{\theta_{j,0}^t} \kappa_{a_j^{t-1}}^{(n)}(\theta_{j,0}^t|s^t, \mathbf{v}) \log Pr(\theta_{j,0}^t|s^t, \mathbf{w}) \quad (3)$$

A straightforward step is to select $\mathbf{w}$ so that $Pr(\theta_{j,0}^t|s^t, \mathbf{w})$ is a mixture of densities of the same form as the original densities, but with fewer components. Thus, we may sample K parameter sets, $\{v_k\}_{k=1}^K$, from $Pr(v|\mathcal{D})$ to obtain an equally weighted mixture:

$$Pr(\theta_{j,0}^t|s^t, \mathbf{w}) = \frac{1}{K} \sum_{k=1}^{K} \tilde{\kappa}_{a_j^{t-1}}^{(n)}(\theta_{j,0}^t|s^t, v_k) \quad (4)$$

where $\tilde{\kappa}_{a_j^{t-1}}^{(n)}(\theta_{j,0}^t|s^t, v_k)$ denotes an individual component density of the same form as in $\kappa_{a_j^{t-1}}^{(n)}$. This implies that $\mathbf{w} = \{v_k\}_{k=1}^K$, and Eq. 3 reduces to finding the 'best' $\{v_k\}_{k=1}^K$. We select $K < |\Omega_j|$ thereby retaining fewer components in memory, and K may be flexibly varied to get as accurate an approximation as needed.

Next, we need an efficient way to compute the integral in Eq. 3. We could obtain samples from $\kappa_{a_j^{t-1}}^{(n)}(\theta_{j,0}^t|s^t, \mathbf{v})$ and utilize them to approximate the integral. Previously generated samples $\{\theta_{j,0}^{t,(x)}\}_{x=1}^X$ may not be used since they helped form the basis for sampling v and we wish to avoid incurring the cost of sampling again from the original complex density. Snelson and Ghahramani [3] recommend generating 'fake data' by first obtaining a new set of $\{v_m\}_{m=1}^M$ from $Pr(v|\mathcal{D})$ where $M < |\Omega_j|$ and using it to approximate $\kappa_{a_j^{t-1}}^{(n)}$:

$$\kappa_{a_j^{t-1}}^{(n)}(\theta_{j,0}^t|s^t, \mathbf{v}) \approx \frac{1}{M} \sum_{m=1}^{M} \tilde{\kappa}_{a_j^{t-1}}^{(n)}(\theta_{j,0}^t|s^t, v_m) \quad (5)$$

The above approximation is easier to sample from than the original mixture density because the above has fewer components. We generate a one-time sample, $\{\theta_{j,0}^{t,(q)}\}_{q=1}^Q$, from the mixture above and utilize the 'fake data' for approximating the integral in Eq. 3:

$$\hat{\mathbf{w}} = \arg\max_{\mathbf{w}} \sum_{q=1}^{Q} \log Pr(\theta_{j,0}^{t,(q)}|s^t, \mathbf{w}) \quad (6)$$

Here, $Pr(\theta_{j,0}^{t,(q)}|s^t, \mathbf{w})$ is evaluated as in Eq. 4.

Notice that we have succeeded in approximating the original mixture density, $\kappa_{a_j^{t-1}}^{(n)}$, with a more compact approximation.

PROPOSITION 1. *The approximation, $Pr(\theta_{j,0}^{t,(q)}|s^t, \hat{\mathbf{w}})$, computed according to Eq. 4 keeps K sets of parameters $\{v_k\}_{k=1}^K$ in memory. In comparison, the original density, $\kappa_{a_j^{t-1}}^{(n)}$, stores $|\Omega_j|$ sets of parameters and the mixture weights in memory.*

Because we select $K < |\Omega_j|$, Proposition 1 implies that the approximation saves on memory. Additionally, it is more efficient to evaluate as well. Note that we could have used the equally weighted mixture in Eq. 5 to approximate $\kappa_{a_j^{t-1}}^{(n)}$. Typically, M in the mixture is greater than K in order to obtain a good set of samples, which makes the mixture in Eq. 5 less efficient. By minimizing KLD, we utilize a principled information-theoretic approach to produce an approximation that is not only compact but also flexibly accurate. On the other hand, we may lose the theoretical guarantees of Rao-Blackwellisation due to the approximation.

Acknowledgment: This research is supported in part by grant FA 9550-08-1-0429 from AFOSR to Prashant Doshi.

4. REFERENCES

[1] P. Doshi. Improved state estimation in multiagent settings with continuous or large discrete state spaces. In *AAAI*, pages 712–717, 2007.

[2] P. Doshi and P. J. Gmytrasiewicz. Approximating state estimation in multiagent settings using particle filters. In *AAMAS*, pages 320–327, 2005.

[3] E. Snelson and Z. Ghahramani. Compact approximations to bayesian predictive distributions. In *ICML*, pages 840–847, 2005.