International Journal of Approximate Reasoning 54 (2013) 602–614

Contents lists available at SciVerse ScienceDirect

International Journal of Approximate Reasoning

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

DemocraticOP: A democratic way of aggregating Bayesian network parameters

Kobra Etminani [a,*], Mahmoud Naghibzadeh [a], Jose M. Peña [b]

[a] Department of Computer Engineering, Ferdowsi University of Mashhad, Iran
[b] ADIT, Department of Computer and Information Science, Linköping University, Sweden

ARTICLE INFO

Article history:
Received 1 March 2012
Received in revised form 10 December 2012
Accepted 13 December 2012
Available online 21 January 2013

Keywords:
Aggregation function
Bayesian networks
Parameter aggregation

ABSTRACT

When there are several experts in a specific domain, each may believe in a different Bayesian network (BN) representation of the domain. In order to avoid having to work with several BNs, it is desirable to aggregate them into a single BN. One way of finding the aggregated BN is to start by finding the structure, and then find the parameters. In this paper, we focus on the second step, assuming that the structure has been found by some previous method.

DemocraticOP is a new way of combining experts' parameters in a model. The logic behind this approach is borrowed from the concept of democracy in the real world. We assume that there is a ground truth and that each expert represents a deviation from it – the goal is to try to find the ground truth based on the experts' opinions. If the experts do not agree, then taking a simple average of their opinions (as occurs in classical aggregation functions such as LinOP and LogOP) is flawed. Instead, we believe it is better to identify similar opinions through clustering, and then apply averaging, or any other aggregation function, over the cluster with the highest number of members to obtain the aggregated parameters that are closest to the ground truth. In other words, respect the majority as is done in democratic societies instead of averaging over all experts' parameters. The new approach is implemented and tested over several BNs with different numbers of variables and parameters, and with different numbers of experts. The results show that DemocraticOP outperforms two commonly used methods, LinOP and LogOP, in three key metrics: the average of absolute value of the difference between the true probability distribution and the one corresponding to the aggregated parameters, Kullback–Leibler divergence, and running time.

1. Introduction

Bayesian networks (BNs) are a popular graphical formalism for representing probability distributions. A BN consists of structure and parameters. The structure, a directed and acyclic graph (DAG), induces a set of independencies that the represented probability distribution satisfies. The parameters specify the conditional probability distribution of each node given its parents in the structure. The BN represents the probability distribution that results from the product of these conditional probability distributions. Typically, a single expert (or learning algorithm such as [1–3]) is consulted to construct a BN of the domain at hand. Therefore, there is a risk that a BN constructed in this way is not as accurate as it could be, e.g., if the expert has a bias or overlooks certain details. One way to minimize this risk is to obtain multiple BNs of the domain from multiple experts and combine them into a single BN. This approach has received significant attention in the literature [4–14].

* Corresponding author: Tel./fax: +98 511 8763306.
E-mail address: etminani@wali.um.ac.ir (K. Etminani).

0888-613X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ijar.2012.12.002

The most relevant of these references is probably [9], because it shows that even if the experts agree on the BN structure, no method for combining the experts' BNs produces a consensus BN that respects some reasonable assumptions and whose structure is the agreed BN structure. Unfortunately, this problem is often overlooked. To avoid it, we proposed combining the experts' BNs in two steps [12,4,11], by first, finding the consensus BN structure and then finding the consensus parameters for the consensus BN structure.

The consensus Bayesian network structure can be obtained by any existing method. In particular, we recommend [12,8,11], because these methods discard the BN parameters provided by the experts and only combine the BN structures provided by the experts, which makes it possible to avoid the problem pointed out by Pennock and Wellman [9]. Even if the experts agree on the BN structure, no method for combining the experts' BNs (structures + parameters) produces a consensus BN that respects some reasonable assumptions and whose structure is the agreed BN structure. Hereafter, we assume that the experts have adopted the consensus structure and thus they differ only in the parameters. In this paper, the second step of the BN combination strategy is studied. Specifically, we introduce a new way of pooling the experts' parameters into one aggregated set of parameters.

In this study, we assume the existence of a ground truth and that experts represent deviations from the ground truth. It seems that this may be a reasonable assumption in many domains such as medicine, where a physical mechanism or system is being modeled. The experts may still disagree due to, for instance, personal beliefs or experiences. However, we do not claim that the assumption of a ground truth is valid in every domain. For instance, if the experts are sport experts that are consulted about the results of the next season, then the idea of the existence of a ground truth is open to discussion, to say the least.

Two commonly used methods for aggregating the BN parameters are linear opinion pool (LinOP) [15] and logarithmic opinion pool (LogOP) [16], which obtain the weighted arithmetic and geometric means respectively. Both methods suffer from two main problems. First, they are slow, since they have to compute the probability of each state of the world and there are exponentially many states. Second, they can be misled by outliers, i.e., non-experts, especially when there is a ground truth and experts believe in a deviation from this truth. One possible solution to the first problem is to do family aggregation, i.e., locally combine the opinions of the experts for each conditional probability distribution [4]. However, this solution still suffers from the second problem. In this paper, we try to address these problems by introducing a novel and smart family aggregation called DemocraticOP.

The main idea of DemocraticOP is to start by clustering the experts' parameters, and then obtain the aggregated parameter by combining the parameters of those experts who are located in the largest cluster. The underlying logic is to respect the majority, i.e., experts in the largest cluster, and disregard others as they are outliers and would distort the result. Moreover, DemocraticOP is fast because it does not need to compute the probability of each state of the world.

We have implemented DemocraticOP and compared it to two previously well studied methods, LinOP and LogOP, in several BNs with different numbers of variables and parameters, and various numbers of experts as well. It shows superior performance in KL divergence, average of the absolute value of difference between true probability distribution and the one corresponding to the aggregated parameters, and particularly in running time.

The rest of this paper is organized as follows. First the preliminaries are discussed in Section 2. Related works in this area are described in Section 3. We introduce DemocraticOP, its time complexity and the properties it satisfies in Section 4. Section 5 includes the experimental results of implementing three aggregation functions DemocraticOP, LinOP and LogOP, and a comparison of the results with respect to different criteria. Finally, we conclude in Section 6.

2. Preliminaries

In this section, the notation used in this paper is introduced. Let m be the number of random variables $V = \{V_1, \ldots, V_m\}$ in a BN and π_i be the parent set of each variable. Each variable has r_i possible states and $q_i = \prod_{V_j \in \pi_i} r_j$ possible parent configurations. Therefore the number of free parameters, N, can be calculated as $N = \sum_{i=1}^{m} q_i(r_i - 1)$. The set of states of the world is denoted by $\Omega = \{\omega_1, \ldots, \omega_A\}$, where $A = \prod_{i=1}^{m} r_i$ is the total number of states of the world.

The number of experts is denoted by n. The experts' opinions can be aggregated by combining all of their opinions about each state of the world, $p_l(\omega_j)$, or each free parameter, $p_l(V_i = v_k|\pi_i = \mu_j)$, where $p_l(x)$ is the probability chosen by expert e_l for event x, v_k is one of the $(r_i - 1)$ possible values of variable V_i, and μ_j is one of the q_i possible parent configurations of V_i. The conditional probability $p_l\,(V_i = v_k|\pi_i = \mu_j)$ is denoted by p_l^z for short, where $z = 1, \ldots, N$.

Note that p_0 is the aggregated parameter and the function f which aggregates the experts' parameters is called the opinion pool:

$$p_0 = f(p_1, \ldots, p_n).$$

3. Previous work

In this section, we explain the background of the various opinion pools and their problems. When combining the parameters of different experts, experts may agree or disagree on the parameters. Experts may all be knowledgeable in the