



## A simple graphical approach for understanding probabilistic inference in Bayesian networks

C.J. Butz<sup>\*</sup>, S. Hua, J. Chen, H. Yao

Department of Computer Science, University of Regina, 3737 Wascana Parkway, Regina, Canada S4S 0A2

### ARTICLE INFO

#### Article history:

Received 7 September 2007

Received in revised form 17 September 2008

Accepted 26 October 2008

#### Keywords:

Bayesian networks

Join tree propagation

Probabilistic inference

Conditional independence

### ABSTRACT

We present a simple graphical method for understanding exact probabilistic inference in discrete *Bayesian networks* (BNs). A conditional probability table (conditional) is depicted as a directed acyclic graph involving one or more black vertices and zero or more white vertices. The probability information propagated in a network can then be graphically illustrated by introducing the *black variable elimination* (BVE) algorithm. We prove the correctness of BVE and establish its polynomial time complexity. Our method possesses two salient characteristics. This purely graphical approach can be used as a pedagogical tool to introduce BN inference to beginners. This is important as it is commonly stated that newcomers have difficulty learning BN inference due to intricate mathematical equations and notation. Secondly, BVE provides a more precise description of BN inference than the state-of-the-art discrete BN inference technique, called LAZY-AR. LAZY-AR propagates potentials, which are not well-defined probability distributions. Our approach only involves conditionals, a special case of potential.

© 2008 Elsevier Inc. All rights reserved.

### 1. Introduction

*Bayesian networks* (BNs) [6,8,12,19,21,26,34,37,40,44,45] are an established framework for uncertainty management in artificial intelligence. For instance, they have been applied in building intelligent agents, such as Office Assistant, and adaptive user interfaces by Microsoft [14,16,30], process control by NASA [15,17] and Lockheed [28], software diagnosis by Hewlett-Packard [3,42] and Nokia [2], and medical diagnosis such as the heart disease program [29] at the Massachusetts Institute of Technology and the Pathfinder Project for lymph-node diseases [13] at the Stanford University. A BN consists of a *directed acyclic graph* (DAG) and a corresponding set of *conditionals* (conditional probability tables) [40]. The *probabilistic conditional independencies* (CIs) [1,43] encoded in the DAG indicate the product of the conditionals is a *joint probability distribution* [4]. Thus, BNs provide a clear semantic modelling tool facilitating the acquisition of probabilistic knowledge.

On the other hand, it is widely acknowledged [7,12,18,19,35,38] that understanding inference algorithms in BNs is an arduous task. Jensen [19] explicitly states that probabilistic reasoning literature is not meant for readers looking for a way into the field. Understanding a probabilistic inference algorithm is not trivial [18]. Although practical applications require *multiply connected* BNs [9], newcomers are usually introduced to inference algorithms conducted on *singly connected* BNs (SCBNs), a special case of BNs. Kim and Pearl [21] proposed one such algorithm, called *propagation in singly connected BNs* (PIS) [6]. Even for inference in SCBNs, Russell and Norvig [38] state that some of the mathematics and notation are unavoidably intricate. Furthermore, Neapolitan [35] writes that inference in a tree BN, a special case of SCBN, is not very transparent.

<sup>\*</sup> Corresponding author. Tel.: +1 306 585 4856; fax: +1 306 585 4745.

E-mail addresses: [butz@cs.uregina.ca](mailto:butz@cs.uregina.ca) (C.J. Butz), [huash111@cs.uregina.ca](mailto:huash111@cs.uregina.ca) (S. Hua), [chen209j@cs.uregina.ca](mailto:chen209j@cs.uregina.ca) (J. Chen), [yao2zhong@cs.uregina.ca](mailto:yao2zhong@cs.uregina.ca) (H. Yao).

We advocate that the use of *potentials* [6,12,40] is the biggest hindrance to the comprehension of probabilistic inference in BNs. Potentials do not have clear physical interpretation [6] as they are not well-defined probability distributions [44]. Yet, potentials play an integral part in the numerous inference algorithms developed for multiply connected BNs [11,20,24,27,31,33,35,37,38,41,46]. Even worse, inference in SCBNs [6,21,35,38] is more complicated in the sense that two types of potentials are needed, namely,  $\lambda$ -potentials are propagated upwards in the SCBN, while  $\rho$ -potentials are propagated downwards. Thus, it is not surprising that beginners experience difficulty grasping probabilistic inference.

In this paper, we propose a simple graphical approach for understanding exact probabilistic inference in discrete BNs. A conditional is represented as a *graphional*, which is a DAG involving one or more black vertices and zero or more white vertices. More specifically, given a conditional  $Pr(X|Y)$ , the corresponding graphional  $Gr(X|Y)$  has a directed edge from each variable in  $Y$  to each variable in  $X$ . Those variables in  $X$  are represented by *black* vertices, while variables in  $Y$  are depicted using *white* vertices. BN inference is described using the *join tree propagation* (JTP) framework [20,24,31,41]. In JTP, messages are systematically passed in a *join tree* (JT) [37,40] constructed from the DAG of the BN. We present an algorithm, called *black variable elimination* (BVE), to graphically illustrate the conditionals being passed from a JT node to a neighbour. We establish the polynomial time complexity of BVE. In addition, we prove the correctness of BVE by showing that the conditionals depicted by our approach are exactly the same as the messages passed by *LAZY-AR* [23,32,33], the state-of-the-art algorithm for exact inference in discrete BNs.

There are two favourable features of our approach. Our purely graphical method can be used as a pedagogical tool to introduce SCBN inference. We propose the *SCBN2JT* algorithm to build a special JT from a SCBN. We show that the messages passed in our constructed JT using the BVE algorithm are exactly the same as those passed in the given SCBN using the PIS algorithm. The difference being BVE is purely graphical, whereas PIS requires intricate mathematical equations and two types of potentials. Our method also provides a more detailed perspective on inference in multiply connected BNs for experienced practitioners. *LAZY-AR* propagates potentials in a JT. On the contrary, our approach graphically describes this same probability information in terms of conditionals. There should be an improvement in clarity, since conditionals are a special case of potentials.

This paper is organized as follows. Section 2 contains definitions. A graphical approach to probabilistic inference is given in Section 3. Section 4 provides the theoretical foundation of our approach. We show how our approach can be used to introduce SCBN inference in Section 5. Section 6 shows how our method can clarify inference in multiply connected BNs. Conclusions are provided in Section 7.

## 2. Definitions

Here we review discrete Bayesian networks and two approaches for exact inference therein.

### 2.1. Bayesian networks

Let  $U = \{v_1, v_2, \dots, v_n\}$  denote a finite set of discrete random variables. Each variable  $v_i$  is associated with a finite domain, denoted  $dom(v_i)$ , representing the values  $v_i$  can take on. For a subset  $X \subseteq U$ , we write  $dom(X)$  for the Cartesian product of the domains of the individual variables in  $X$ . Each element  $x \in dom(X)$  is called a *configuration* of  $X$ .

**Definition 1** [12]. A *potential* on  $dom(X)$  is a function  $\phi$  on  $dom(X)$  such that  $\phi(x) \geq 0$ , for each configuration  $x \in dom(X)$ , and at least one  $\phi(x)$  is positive.

**Example 1.** Three potentials  $\phi(a)$ ,  $\phi(b, c)$  and  $\phi(d, e, f)$  are illustrated in Fig. 1, where all variables are binary.

For brevity, we refer to a potential as a probability distribution on  $X$  rather than  $dom(X)$ , and we call  $X$ , not  $dom(X)$ , its domain [40]. Also, for simplified notation, we may write  $XY$  to denote  $X \cup Y$ . A *joint probability distribution* (JPD) [4] on  $U$ , denoted  $Pr(U)$ , is a potential on  $U$  that sums to one.

**Definition 2** [40]. Let  $X$  and  $Y$  be two disjoint subsets of  $U$ . A *conditional* for  $Y$  given  $X$ , denoted  $Pr(Y|X)$ , is a nonnegative function on  $XY$ , satisfying the following condition: for each configuration  $x \in dom(X)$ ,  $\sum_{y \in dom(Y)} Pr(Y = y|X = x) = 1.0$ .

a	$\phi(a)$	b	c	$\phi(b, c)$	d	e	f	$\phi(d, e, f)$
0	0.423	0	0	0.496	0	0	0	0.000
1	0.000	0	0	0.061	0	0	1	0.059
		0	1	0.000	0	1	0	0.441
		0	1	0.504	1	1	0	0.000
		1	0	0.435	1	0	0	0.227
		1	0	0.029	1	0	1	0.038
		1	1	0.000	1	1	0	0.000
		1	1	0.475	1	1	1	0.023

Fig. 1. Potentials  $\phi(a)$ ,  $\phi(b, c)$  and  $\phi(d, e, f)$ .

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات