International Journal of Approximate Reasoning 49 (2008) 503–521

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

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

Belief update in CLG Bayesian networks with lazy propagation

A.L. Madsen

HUGIN EXPERT A/S, Gasværksvej 5, DK-9000 Aalborg, Denmark

ARTICLE INFO

Article history:
Received 2 January 2008
Received in revised form 8 May 2008
Accepted 9 May 2008
Available online 23 May 2008

Keywords:
CLG Bayesian networks
Belief update
Propagation

ABSTRACT

In recent years, Bayesian networks with a mixture of continuous and discrete variables have received an increasing level of attention. In this paper, we focus on the restricted class of mixture Bayesian networks known as conditional linear Gaussian Bayesian networks (CLG Bayesian networks) and present an architecture for exact belief update for this class of mixture networks.

The proposed architecture is an extension of lazy propagation using operations of Lauritzen and Jensen [S.L. Lauritzen, F. Jensen, Stable local computation with mixed Gaussian distributions, Statistics and Computing 11(2) (2001) 191–203] and Cowell [R.G. Cowell, Local propagation in conditional Gaussian Bayesian networks, Journal of Machine Learning Research 6 (2005) 1517–1550]. By decomposing clique and separator potentials into sets of factors, the proposed architecture takes advantage of independence and irrelevance properties induced by the structure of the graph and the evidence. The resulting benefits are illustrated by examples and assessed by experiments.

The performance of the proposed architecture has been evaluated using a set of randomly generated networks. The results indicate a significant potential of the proposed architecture.

1. Introduction

The framework of Bayesian networks is an efficient knowledge representation for reasoning under uncertainty [19,5,6]. Traditionally, the variables of a Bayesian network are assumed to be either discrete or continuous. In recent years, Bayesian networks with a mixture of continuous and discrete variables have, however, received an increasing level of attention. In this paper, we focus on the restricted class of mixture Bayesian networks known as CLG Bayesian networks where continuous variables are assumed to have a conditional linear Gaussian distribution and where discrete variable may have discrete parents only. In particular, we present an algorithm for exact belief update in CLG Bayesian networks where belief update is defined as the task of computing all single posterior marginals given a set of evidence.

Extending the class of Bayesian networks containing discrete (or continuous) variables only to the class of Bayesian networks containing both discrete and continuous variables is not simple. The work by Pearl [19] on Bayesian networks containing continuous variables imposed three constraints on the variables in the network. The interaction between variables is linear, the sources of uncertainty are Gaussian distributed and uncorrelated, and the causal network is singly connected. Later, Shachter and Kenley [24] described how to solve Gaussian influence diagrams under similar constraints, but allowing multiply connected causal networks.

Lauritzen [8] presents a scheme for modeling and exact belief update in CLG Bayesian networks. This scheme is more general than the scheme proposed by Pearl. The conditional distribution of the continuous variables given the discrete variables is assumed to be multivariate Gaussian. We consider only models where the continuous variables have a linear additively Gaussian distribution. The asymmetry between continuous and discrete variables induces a number of constraints on the

E-mail address: Anders.L.Madsen@hugin.com

0888-613X/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ijar.2008.05.001

model specification and the inference structure. Using a similar approach, Chang and Fung [1] extend the SPI algorithm [22] to solve arbitrary queries against CLG Bayesian networks.

The basic operations of the scheme proposed by Lauritzen are combination and addition of conditional Gaussian distributions. The combination operation is based on a canonical characteristics representation while the marginalization operation is based on a moment characteristics representation. This implies that it is necessary to switch between the two representations during belief update. The numerical instability resulting from the switch of representation motivated the introduction of a computational architecture based solely on the moment characteristics representation [9].

Recently, Cowell [4] introduced an alternative architecture for belief update based on message passing in an elimination tree using the arc-reversal operation of Shachter and Kenley [24] (referred to as the EXCHANGE operation) as the basic operation for variable elimination during belief update. By performing message passing in an elimination tree the need for complex matrix operations is eliminated. The elimination of complex matrix operation greatly simplifies the implementation of the algorithm.

We introduce a new architecture for belief update in CLG Bayesian networks. The architecture is an extension of the lazy AR propagation architecture [16] based on the PUSH and EXCHANGE operations introduced by Lauritzen and Jensen [9] and Cowell [4], respectively. In lazy AR propagation messages are computed using arc-reversal operations and barren variable removals for variable elimination.

In the proposed architecture, belief update proceeds by message passing in a strong junction tree structure where messages are computed using arc-reversal operations for discrete variable elimination and EXCHANGE operations for continuous variable elimination. To simplify the presentation, the EXCHANGE operation is extended to include arc-reversal between discrete variables. Variables are eliminated using a sequence of EXCHANGE operations and barren variable removals whereas posterior marginal distributions are computed using EXCHANGE and PUSH operations. By decomposing clique and separator potentials and eliminating variables using EXCHANGE and barren variable removal operations, it is possible to take advantage of independence and irrelevance properties induced by the structure of the graph and the evidence to reduce the cost of belief update.

In an experimental evaluation, we investigate the computational efficiency of the proposed architecture using a set of randomly generated CLG Bayesian networks. In addition, we analyze the performance of various steps of belief update such as computing posterior distributions and the importance of the order in which evidence on continuous variables is inserted into the strong junction tree structure.

The paper is organized as follows. In Section 2, preliminaries and the notation used in the paper are presented. Section 3 describes the lazy propagation architecture for belief update in CLG Bayesian networks as proposed in this paper. A comparison between the proposed architecture and existing architectures for belief update in CLG Bayesian networks is presented in Section 4. The results of an empirical performance evaluation are presented in Section 5. Finally, Section 6 concludes the paper.

2. Preliminaries and notation

2.1. CLG Bayesian network

A CLG Bayesian network $\mathcal{N} = (\mathcal{X}, \mathcal{G}, \mathcal{P}, \mathcal{F})$ over variables $\mathcal{X}$ consists of an acyclic, directed graph (DAG) $\mathcal{G} = (V, E)$, a set of conditional probability distributions $\mathcal{P} = \{P(X|\pi(X)) : X \in \Delta\}$ where $\pi(X)$ is the set of variables corresponding to the parents of the vertex representing X in $\mathcal{G}$, and a set of CLG density functions $\mathcal{F} = \{f(Y|\pi(Y)) : Y \in \Gamma\}$ where Δ is the set of discrete variables and Γ is the set of continuous variables such that $\mathcal{X} = \Delta \cup \Gamma$. The vertices V of $\mathcal{G}$ correspond one to one with the variables of $\mathcal{X}$.

A CLG Bayesian network $\mathcal{N}$ induces a multivariate normal mixture density over $\mathcal{X}$ on the form:

$$P(\Delta) \cdot f(\Gamma|\Delta) = \prod_{X \in \Delta} P(X|\pi(X)) \cdot \prod_{Y \in \Gamma} f(Y|\pi(Y)).$$

Let $Y \in \Gamma$ with $I = \pi(Y) \cap \Delta$ and $Z = \pi(Y) \cap \Gamma$, then Y has a CLG distribution if

$$\mathcal{L}(Y|I = i, Z = z) = N(\alpha(i) + \beta(i)z, \sigma^2(i)), \tag{1}$$

where the mean value of Y depends linearly on the values of the continuous parent variables Z, while the variance is independent of Z. In (1), $\alpha(i)$ is a table of real numbers, $\beta(i)$ is a table of $|Z|$-dimensional vectors, and $\sigma^2(i)$ is a table of non-negative values.

Belief update is the task of computing posterior marginals given evidence. Evidence on a variable $X \in \Delta$ is assumed to be an instantiation, i.e., $X = x$. Evidence on a variable $Y \in \Gamma$ is an assignment of a value y to Y, i.e., $Y = y$. We let ϵ_Δ and ϵ_Γ denote the set of evidence on variables of Δ and Γ, respectively, such that $\epsilon = \epsilon_\Delta \cup \epsilon_\Gamma$. Furthermore, we let $\mathcal{X}_\epsilon$ denote the set of variables instantiated by evidence ϵ.

Definition 2.1 (*Barren variable*). A variable X is a *barren variable* w.r.t. a set of variables $T \subseteq \mathcal{X}$, evidence ϵ, and DAG $\mathcal{G}$, if $X \notin T$, $X \notin \epsilon$ and X only has barren descendants in $\mathcal{G}$ (if any).