International Journal of Approximate Reasoning 70 (2016) 99–122

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Uncertain and negative evidence in continuous time Bayesian networks

Liessman Sturlaugson *, John W. Sheppard

Department of Computer Science, Montana State University, Bozeman, MT 59717, USA

ARTICLE INFO

Article history:
Received 16 January 2015
Received in revised form 16 December 2015
Accepted 17 December 2015
Available online 23 December 2015

Keywords:
Continuous time Bayesian network
Uncertain evidence
Negative evidence
Exact inference
Importance sampling

ABSTRACT

The continuous time Bayesian network (CTBN) enables reasoning about complex systems by representing the system as a factored, finite-state, continuous-time Markov process. Inference over the model incorporates evidence, given as state observations through time. The time dimension introduces several new *types* of evidence that are not found with static models. In this work, we present a comprehensive look at the types of evidence in CTBNs. Moreover, we define and extend inference to reason under uncertainty in the presence of uncertain evidence, as well as negative evidence, concepts extended to static models but not yet introduced into the CTBN model.

1. Introduction

Probabilistic models, such as Bayesian networks, provide a mathematically rigorous framework for reasoning under uncertainty. Given observations (i.e., evidence) about the system that the model represents, the model can update the posterior probabilities of other states of the system in light of that evidence. However, the representation of evidence in Bayesian networks has been further extended to allow for uncertainty in the evidence, in which there is uncertainty in the observations themselves. For example, a medical test might have false positive and false negative rates. Thus, the test result can be trusted only to a certain degree, which uncertain evidence is able to capture. Uncertain evidence provides one generalization of evidence, in which the evidence also has an associated likelihood.

Furthermore, one can think of negative evidence. Instead of an observation of the system being in a certain state, we can observe the system to *not* be in certain states, but it could be any of several other states. For example, we may be modeling a system in which sensors may only be able to detect a subset of the possible states. A negative sensor reading would imply that the system is in one of those other states. Negative evidence provides another generalization of evidence, in which multiple states can be ruled out instead of a single state being given.

By lifting restrictions on the types of evidence that the models can support, generalizations of evidence make the models more powerful and versatile. Rather than having to transform the observations into a form of supported evidence, such as treating all observations as certain or ignoring incomplete observations, generalizing evidence allows the observations to be used as evidence directly.

As the CTBN is a relatively new model, current CTBN inference algorithms only support certain and positive evidence, in which all of the temporal state evidence is trusted with complete confidence and in which the temporal state evidence

* Corresponding author.
E-mail addresses: listurlaugson@gmail.com (L. Sturlaugson), john.sheppard@cs.montana.edu (J.W. Sheppard).

http://dx.doi.org/10.1016/j.ijar.2015.12.013

dictates what the state *is*, never what it is not. However, in many applications, the evidence may contain errors and can be trusted only to a certain degree. In some cases, only subsets of the states may be observable. Here, the model can benefit from ruling out certain states by reasoning with negative evidence. Uncertain and negative evidence for CTBNs has not yet been defined nor have CTBN inference algorithms been extended to allow for incorporating negativity and uncertainty into the temporal evidence. Therefore, this work proposes a definition for uncertain and negative evidence and shows how to support the definitions in the context of CTBNs. The introduction of timing information adds another dimension into the types of evidence we can apply. We show how combinations of certain/uncertain evidence and positive/negative evidence interact. By so doing, we show how the definitions of uncertain and negative evidence provide a generalization of certain positive evidence.

2. Background work

In this section, we give the formal definition of the Bayesian network (BN), as well as its temporal version, the dynamic Bayesian network (DBN). Next we give the formal definition of the CTBN model, compare and contrast the CTBN and the DBN, and then survey existing CTBN applications and extensions.

2.1. Bayesian networks

Bayesian networks are probabilistic graphical models that use nodes and arcs in a directed acyclic graph to represent a joint probability distribution over a set of variables [1]. Let $P(\mathbf{X})$ be a joint probability distribution over n variables $\mathbf{X} = \{X_1, \ldots, X_n\}$. A Bayesian network $\mathcal{B}$ is a directed, acyclic graph in which each variable X_i is represented by a node in the graph. Let $\mathbf{Pa}(X_i)$ denote the parents of node X_i in the graph. The graph representation of $\mathcal{B}$ factors the joint probability distribution as:

$$P(\mathbf{X}) = \prod_{i=1}^{n} P(X_i|\mathbf{Pa}(X_i)). \tag{1}$$

This factorization is induced by the conditional independences in the underlying distribution. Without any factorization, the number of parameters required to define the full joint probability distribution is exponential in the number of variables. By factoring the joint probability distribution to consider only the relevant variable interactions, represented by the parent–child relationships in the network, often the complexity of the distribution can be managed.

2.2. Dynamic Bayesian networks

The Bayesian network defined above is a static model. However, we can introduce the concept of time into the network by assigning discrete timesteps to the nodes to create a dynamic Bayesian network, a temporal version of a BN.

A dynamic Bayesian network (DBN) is a special type of Bayesian network that uses a series of connected timesteps, each of which contains a copy of a regular Bayesian network defined over $\mathbf{X}$ indexed by time t (denoted $\mathbf{X}_t$). The probability distribution of a variable at a given timestep can be conditioned on states of that variable (or even other variables) throughout any number of previous timesteps as well as on other variables within the same timestep. In first-order DBNs, the nodes in each timestep are conditionally independent of any nodes further back than the immediately previous timestep, given that previous timestep. Therefore, the joint probability distribution for a first-order DBN factors as:

$$P(\mathbf{X}_0, \ldots, \mathbf{X}_k) = P(\mathbf{X}_0) \prod_{t=0}^{k} P(\mathbf{X}_{t+1}|\mathbf{X}_t). \tag{2}$$

Spanning multiple timesteps, the DBN can include any evidence gathered throughout that time and use it to help reason about state probability distributions across different timesteps. Often, the conditional probability tables of the DBN can be defined compactly by defining a prior network over $\mathbf{X}_0$ and a single temporal network over $\mathbf{X}_t$ that represents $P(\mathbf{X}_{t+1}|\mathbf{X}_t)$ for every t. The temporal network $\mathbf{X}_t$ is then “unrolled” to consider $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_k$ for k timesteps.

2.3. Continuous time Bayesian networks

The continuous time Bayesian network was first introduced in [2]. Although its name attempts to draw parallels between the conditional independence encoded by Bayesian networks, the CTBN represents a factored Markov process.

Let $\mathbf{X}$ be a set of conditional Markov processes $\{X_1, X_2, \ldots, X_n\}$, where each conditional process X_i has a finite number of discrete states. Formally, a continuous time Bayesian network $\mathcal{N} = \langle \mathcal{B}, \mathcal{G} \rangle$ over $\mathbf{X}$ consists of two components. The first is a Bayesian network $\mathcal{B}$ with nodes corresponding to $\mathbf{X}$. This Bayesian network is only used for determining $P(\mathbf{X}_0)$, the initial distribution of the process. The second is a continuous-time transition model $\mathcal{G}$, which describes the evolution of the process from its initial distribution, specified as: