On Intercausal Interactions in Probabilistic Relational Models

Silja Renooij
Linda C. van der Gaag

Department of Information and Computing Sciences, Utrecht University, The Netherlands

Philippe Leray
LS2N UMR CNRS 6004, DUKe research group, University of Nantes, France

Abstract

Probabilistic relational models (PRMs) extend Bayesian networks beyond propositional expressiveness by allowing the representation of multiple interacting classes. For a specific instance of sets of concrete objects per class, a ground Bayesian network is composed by replicating parts of the PRM. The interactions between the objects that are thereby induced, are not always obvious from the PRM. We demonstrate in this paper that the replicative structure of the ground network in fact constrains the space of possible probability distributions and thereby the possibly patterns of intercausal interaction.

Keywords: PRM instances, qualitative constraints on probability distributions, intercausal interaction.

1. Introduction

Real-world problem domains often have a relational structure involving multiple interacting object classes, which cannot be appropriately modelled by a Bayesian network. By allowing the representation of relational information, probabilistic relational models, or PRMs, extend on the propositional expressiveness of Bayesian networks [2, 3, 4]. More specifically, a PRM describes the object classes from a relational schema by a graphical dependency structure and supplements this structure with probabilistic information. Inference with an instance composed of sets of concrete objects per class is performed in a ground Bayesian network which is derived from the PRM by replicating parts for the concrete objects involved. The PRM thus is essentially a meta-model covering all possible instances.

Upon constructing a ground Bayesian network for a given instance, interactions may be induced between the random variables modelling the PRM’s feature variables, that are not always obvious from the meta-level model. In this paper, we study the possible interactions resulting for an instance from relational dependencies between the features of different classes. We focus more specifically on many-to-one dependencies: for a given instance, the value of a feature variable $E$ of a concrete object for some class $X$ may then depend on the values of the variable $C$ in multiple objects for a class $X'$. In the PRM, the precise dependency involved is described by an aggregation function which serves to summarise a multiset of values of $C$ into a single such value [3]. Example aggregation functions are the logical OR, the MODE and the STOCHASTIC-MODE [5]. In this paper, we will demonstrate that, in addition to the aggregation function used, the replicative structure of the ground network constrains the intercausal interactions induced among multiple objects. Our analysis informs the choice of aggregation function to attain the desired type of interaction in a PRM instance.

2. Running example

As running example, we consider the small PRM shown in Figure 1(a) and assume all variables to be binary-valued. The feature variables $C$ and $E$ belong to different classes (left implicit). The relational dependency between $C$ and $E$ thus extends over the class boundaries, and we assume it to be of type many-to-one, indicated by an arc associated with an aggregation function. We now consider an instance with two objects with the variable $C$, whose replicates are denoted $C_1$ and $C_2$ as shown in Figure 1(b). For the node $C$, encoding the aggregation of the two $C_i$ values, in essence four conditional probabilities are specified:

\[
\begin{align*}
    p_1 &= \Pr(c | c_1, c_2) \\
    p_2 &= \Pr(c | c_1, \bar{c}_2) \\
    p_3 &= \Pr(c | \bar{c}_1, c_2) \\
    p_4 &= \Pr(c | \bar{c}_1, \bar{c}_2)
\end{align*}
\]
We address the patterns of interaction between the two and combinations below it in explaining-away. To aid in the combinations above this surface result in a value is known for the aggregation node $C$. Properties of the PRM in fact dictate to a large extent the set of possible probability values $p_i, i = 1, 2, 3, 4$, for the aggregation node $C$, and thereby the patterns of interaction that can be induced. Given the constraint $p_2 = p_3$, Figure 2(a) shows the surface of all $p_1$ combinations that result in a zero product synergy, and hence in a lack of interaction between $C_1$ and $C_2$ given $C = c$. Probability combinations above this surface result in $X^+(\{C_1, C_2\}, c)$ and combinations below it in explaining-away. To aid in the interpretation of Figure 2(a), Figure 3 shows two slices of the surface: the left one is the horizontal slice for $p_1 = 0.8$, and the right one the vertical slice at $p_2 = p_3 = 0.5$.

Another qualitative property of probability distributions is described by the concept of qualitative influence [7], denoted by $S^\delta(C, C)$, which captures the sign of the direct influence between two variables $C_i$ and $C$. More specifically, the influence is positive if $Pr(c|C_i, C_j) \geq Pr(c|\bar{C}_i, C_j)$ for all values of $C_j, i, j = 1, 2, i \neq j$; assuming $c > \bar{c}$ and $C_j \succ C_i$, such an influence captures the idea that higher values of $C_i$ make higher values of $C$ more likely, regardless of other influences on $C$. In the ground Bayesian network, the direct influences are constrained by $S^\delta(C_i, C), i = 1, 2$ [6], that is, the set of possible probability values $p_i, i = 1, 2, 3, 4$, is restricted by $p_1 \geq \{p_2, p_3\} \geq p_4$. The plane representing the constraint $p_1 = p_2 = p_3$ is shown by dotted lines in Figure 2(b); dashed lines show the plane representing the constraint $p_2 = p_3 = p_4$. The figure shows that these con-
Restraints significantly reduce the possible $p_i$ combinations that induce a zero product synergy, and leave very little room for capturing positive intercausal interactions in a PRM at hand. We further note that any deterministic aggregation function, encoded by a degenerate distribution for $C$ with $p_1 = 1$ and $p_4 = 0$, will necessarily result in negative intercausal interactions given $C = c$. If positive intercausal interactions or a lack of interaction are desired therefore, the use of a stochastic aggregation function is called for.

The product synergies discussed capture the intercausal interaction upon observation of $C = c$. The intercausal interactions induced by observing $C = \bar{c}$ are found basically by mirroring the surface across all three axes in the figures.

4. Conclusion and Future Research

Our analyses show that to induce the type of intercausal interaction desired for an application domain requires careful tuning of the probabilities for the observed value of aggregation node $C$. As the probabilities for this node in essence encode an aggregation function, studying the possible probability values for $C$ thus informs the choice of an appropriate aggregation function for the PRM.

In view of possible instantiations of the PRM we are typically interested in intercausal effects due to the observation of the effect node $E$ rather than aggregation node $C$. An observation for node $E$ is essentially indirect evidence for $C$ in the ground Bayesian network, and capturing the induced intercausal effects then involves the product synergies for both values of $C$ [1]. For instantiations with more than two objects, we can use a generalised definition of product synergy to similarly study all pair-wise interactions [1]. The concepts of product synergy and qualitative influence are not restricted to binary variables, but are in fact defined for discrete variables that allow a total order on their values. By considering all possible pairs of values $c_i > c_j$ of the variables involved, our results generalise straightforwardly beyond the binary case. A more formal underpinning of further generalisations is topic of future research.

References


