AN OVERVIEW OF EXTENSIONS OF BAYESIAN NETWORKS TOWARDS FIRST-ORDER LOGIC

András MILLINGHOFFER
Advisors: Tadeusz DOBROWIECKI and Péter ANTAL

I. Introduction

Bayesian networks have become the most widely used models for representing abstract, uncertain probabilistic knowledge. A Bayesian network model consists of two representational levels: (1) the (directed acyclic) graph codes the dependency model of the underlying distribution, (2) a conditional probability distribution at each node, storing the distribution of the node given its parents (see [1]).

One main reason for the popularity of the Bayesian networks is their ability of efficiently decomposing and thus representing the whole distribution. On the other hand the graph provides easily understandable qualitative information about the underlying knowledge to human users as well.

Despite these advantageous features, the computational complexity and other costs which occur during model building (e.g. collecting measurement data, consulting with experts) make the efficient reuse of such models in complex problems difficult. This problem is analogue to the extension of the propositional logic to the first-order logic, where the quantification makes it possible to store reusable knowledge at a higher level, from which we can gain multiple propositional models (in the context of Bayesian networks, the monolithic models used in everyday systems).

This paper gives an enumeration and a short description of the most important representational languages which extend monolithic Bayesian networks.

II. The examined extensions

The following methods are similar in the respect that their basic aim is to construct Bayesian networks (or more generally any kind of decision support systems) automatically with the help of some higher-level information source. This could allow us to avoid expensive learning algorithms (or at least to speed them up significantly) or equivalently, the extraction of expert knowledge could be done only once (i.e. while creating this higher-level source).

A. Knowledge-based model construction (KBMC)

The basic concept of KBMC is relying on the use of knowledge bases (KBs) containing (probabilistic) logical rules. In logical databases information concerning types of objects can easily be stored while on the contrary decision support systems usually handle instances.

In KBMC systems abstract information about concepts and their relations is stored in the background KB. When considering a given problem, after enumerating the instances of objects involved, their relations (the model's structure) can be extracted from the KB. We can avoid thus the expensive structure learning methods. For further details see [2].

B. Object-oriented Bayesian networks (OOBNs)

In almost every application domain (i.e. where training data contain missing entries) learning cannot be applied to parts of the model, only to the whole of it. This means that after making even minor changes to a model (adding a new or deleting an already existing node) the whole learning process must be repeated.

OOBN models are built up from smaller pre-trained network fragments, the so called objects. Objects consist of attributes (BN nodes) which can be of three types: input, output and encapsulated. An

object can be connected to other objects which determine its properties trough its input attributes, while it provides information trough the output attributes. Encapsulated attributes store inner information.

In the OOBN framework we can build our models from larger parts assuming that the connections at the object level do not influence the relation in an object. Consequently we can teach each object separately, before building the whole model. Then we only have to adjust the parameters of inter-object connections. The OOBN concept is explained in detail in [3].

C. Probabilistic relational models (PRMs)

PRMs are constructed analogously to relational databases, which are the most widely used data sources. A PRM object has its own attributes (like in DBs) and has reference slots to other objects (with the analogy of outer keys).

PRMs make it possible to incorporate all the information accessible in the source database, without having to 'flatten' the data (i.e. not considering the information stored in the structure). See [4].

D. Relational Bayesian networks

The original BN models can be used to model first-order predicates as well. In this case the result of a query in the presence of some evidence is the probability of the given instantiation of the variables. The weakness of this method is that it can be applied to model relations between the features of the same random event, but not relations of different events (e.g. comparing different instantiations).

In the framework of relational Bayesian networks each node represents a predicate, indicating its true or false value. Here the probability distribution of a node can incorporate information of every instantiation of the parents by combining their probabilities using proper functions, such as *max*, *noisy-or*, etc. Relational BNs are described in detail in [5].

E. Graph-grammars

In some domains (e.g. in medicine) it is common that typical substructures occur multiple times in the domain-wide model. The method described in [6] tries to exploit this property by introducing graph-grammars. A graph-grammar consists of transformation rules and operates on graph structures. Using this methodology during the creation of considered structures we can assure that those structures will be examined which have the needed feature.

III.   Future work

After having examined the above methods theoretically I plan to test them in various application fields and compare them concerning their ability of representing background knowledge and the ease of the construction of application models. Then the best one or a combination of several ones could be chosen, possibly a new language could be constructed for further evaluation in a real-world application domain.

References

[1] J. Pearl, *Causality: Models, Reasoning, and Inference*, Cambridge University Press, 2000.

[2] M. P. Wellman, J. S. Breese, and R. P. Goldman, "From knowledge bases to decision models," *The knowledge engineering review*, 7:35–53.

[3] D. Koller and A. Pfeffer, "Object-oriented bayesian networks," in *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence (UAI-97)*, pp. 302–313, 1997.

[4] D. Koller and A. Pfeffer, "Probabilistic frame-based systems," in *AAAI/IAAI*, pp. 580–587, 1998.

[5] M. Jaeger, "Relational Bayesian networks," in *Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence*, M. Kaufmann, Ed., pp. 266–273, 1997.

[6] J. Egar and M. Musen, "Graph-grammar assistance for automated generation of influence diagrams," in *Proceedings of the 1st Annual Conference on Uncertainty in Artificial Intelligence (UAI-85)*. Elsevier Science Publishing Comapny, Inc., 1985.