

Explanation in Bayesian and Credal Networks

Alberto Piatti
IDSIA, Switzerland
alberto.piatti@idsia.ch

Alessandro Antonucci
IDSIA, Switzerland
alessandro@idsia.ch

Marco Zaffalon
IDSIA, Switzerland
zaffalon@idsia.ch

Abstract

A *Knowledge-based system* (KBS) is a computer program designed to support human experts in complex diagnostic, identification, or decision problems through a mathematical model of the available knowledge. To improve the acceptance of the system by human users and to help debugging the system, it is important to provide them with explanation facilities. Addressing this need is particularly important for KBSs developed on the basis of normative formalisms, whose structure and automatic reasoning might be far from human attitudes. In this paper, we focus on the explanation of normative KBSs based on Bayesian Networks (BNs) and Credal Networks (CNs). The need for explanation facilities in KBSs based on BNs has been highlighted by Suermondt [3] and Druzdzel [1], who have proposed several approaches to the explanation of BNs. Recently, Piatti et al. [2] have proposed a procedure for building KBSs based on CNs that allows to overcome two well-known problems of KBS based on BNs: the problem of obtaining from the experts the precise probabilities required for the quantification of the BN and the problem of modeling adequately the *observational process*. The need for explanation facilities in KBSs based on BNs is clearly valid also for KBSs based on CNs, but explanation methods for CNs have never been investigated so far. In this paper, we address for the first time the problem of the explanation of CNs. In the first part of the paper, the existing explanation methods for BNs are reviewed in the light of their applicability to CNs. In the second part of the paper, we generalize and adapt two existing explanation methods for BNs to CNs. The first method has been proposed by Suermondt [3] to explain the inferences produced by a BN. It consists in identifying the most informative findings, respectively the irrelevant ones, by comparing the inference produced according to the available evidence with the inferences produced by assuming that some of the previously observed variables are missing. The method assumes (implicitly) that the missingness process is *missing at random*. The application of this method to CNs allows the missingness process to be modeled in a more flexible and reliable way. We propose a generalization of the method to CNs based on a conservative approach, called *conservative inference rule*, producing robust explanations of the inferences returned by a CN. The second method that we discuss in detail has been proposed by Druzdzel [1]. It consists in transforming a BN with binary or ordinal variables into a *qualitative probabilistic network* (QPN) to produce qualitative verbal explanations for BN inferences in the light of the collected evidence. While the transformation of a BN into a QPN is not problematic, the transformation of a CN into a QPN is not straightforward and has never been addressed in general. We propose here a general procedure for transforming a CN with binary variables into a QPN and we show that the adapted method is able to produce meaningful explanations.

Keywords. Credal networks, explanation, observational process, qualitative probabilistic networks.

References

- [1] M. Druzdzel. *Probabilistic reasoning in decision support systems: from computation to common sense*. PhD thesis, Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA, 1993.
- [2] A. Piatti, A. Antonucci, and M. Zaffalon. Building knowledge-based systems using credal networks: A tutorial. Nova Science Publishers, New York, 2009. Accepted for publication.
- [3] H.J. Suermondt. *Explanation in Bayesian Belief Networks*. PhD thesis, Department of Computer Science, Stanford University, Stanford, CA, 1992.