Approximating Credal Network Inferences by Linear Programming

Alessandro Antonucci, Cassio P. de Campos, David Huber, and Marco Zaffalon

Istituto Dalle Molle di Studi sull’Intelligenza Artificiale (IDSIA)
Galleria 2, Manno-Lugano, Switzerland
{alessandro,cassio,david,zaffalon}@idsia.ch

Abstract. An algorithm for approximate credal network updating is presented. The problem in its general formulation is a multilinear optimization task, which can be linearized by an appropriate rule for fixing all the local models apart from those of a single variable. This simple idea can be iterated and quickly leads to very accurate inferences. The approach can also be specialized to classification with credal networks based on the maximality criterion. A complexity analysis for both the problem and the algorithm is reported together with numerical experiments, which confirm the good performance of the method. While the inner approximation produced by the algorithm gives rise to a classifier which might return a subset of the optimal class set, preliminary empirical results suggest that the accuracy of the optimal class set is seldom affected by the approximate probabilities.

1 Introduction

Credal networks [5] are a generalization of Bayesian networks (e.g., [11]) based on the notion of credal sets. A credal set is a set of probability mass functions, thus representing a quite general and expressive model of uncertainty. Other uncertainty models like belief functions [14] or possibility measures can be regarded as (special classes of) credal sets. A Bayesian network can be turned into a credal network by simply replacing the local models, which are conditional probability mass functions, with conditional credal sets over the same variables. Exactly as a Bayesian network defines a joint probability mass function over its whole set of variables, a credal network defines a joint credal set, which is (the convex closure of) the set of all joint mass functions obtained from the Bayesian networks consistent with the local credal sets.

Compared to the case of Bayesian networks, inference in credal networks is considerably harder. For instance, a marginalization task corresponds to a multilinear optimization problem (updating is a fractional multilinear task) [7]. This is known to be NP-hard even for singly connected networks [8], while the analogous inference in Bayesian networks can be performed in polynomial time [11].

* This work was supported by the Swiss NSF grants nos. 200020_134759 / 1, 200020_137680 / 1, and by the Hasler foundation grant n. 10030.

© Springer-Verlag Berlin Heidelberg 2013
Despite the hardness of the problem, some algorithms are known to perform reasonably well under certain conditions. Exact approaches have been proposed that implement some branch-and-bound method with local searches \[4,6,8,9\]. Unfortunately they all suffer from serious efficiency issues unless the credal network is very simple. For instance, none of these methods can deal well with a binary node having four ternary parents, because this setting is already equivalent to \(3^4 = 81\) free optimization variables to be chosen, meaning a space of \(2^{81}\) possible solutions just locally to this node! On the other hand, approximate methods either are fast and provide no accuracy guarantee \[3,4,6\] or provide theoretical guarantees but are as slow as exact methods \[13\]. Moreover, all these approximate methods are only capable of treating credal networks under a vertex-based representation, while a constraint-based specification of credal networks still lacks any practical algorithm.

In this paper we present a fast approximate algorithm for inferences in credal networks based on solving a sequence of linear programming problems. It uses a constraint-based specification, which allows us to deal with domains where the local credal sets are given by their linear constraints. It does not suffer from many parents and large credal sets because the optimization is done by compact linear problems. To the best of our knowledge, this is the first method for general credal networks to truly run the inference with a constraint-based specification. We describe the method and some heuristic ideas to improve its accuracy. Unlike similar ideas already proposed in the literature \[6\], our approach does not require an explicit enumeration of the extreme points of the credal sets and should be therefore used when the number of extreme points in the local credal sets is exponentially large (e.g., variables with many states and/or parents, credal sets defined by probability intervals, etc). We also discuss how the method can be used for decision making under the maximality criterion \[15\].

Sections 2 and 3 review the basic notation and definitions of Bayesian and credal networks. The proposed procedure is presented in Sections 4 and 5. Numerical experiments show that the proposed method compares favorably against other available methods in the literature (Section 7). Results are particularly positive when the algorithm is specialized to the case of classification in credal networks based on the maximality criterion. Although this problem is shown to be even harder than the marginalization inferences (discussed in Section 6), classifications based on our approximate algorithm are empirically shown to coincide with those based on exact methods.

2 Bayesian Networks

Consider a set of variables \(X := (X_0, X_1, \ldots, X_n)\) in one-to-one correspondence with the nodes of an acyclic directed graph \(G\). For each \(i = 0, \ldots, n\), the joint variable \(\Pi_i \subseteq X\) denotes the parents of \(X_i\) according to \(G\). All these variables are categorical: \(X_i\) takes its values on the finite set \(\Omega_{X_i}\) and so does \(\Pi_i\) in \(\Omega_{\Pi_i} := \times_{j \in \Pi_i} \Omega_{X_j}\), for each \(i = 0, \ldots, n\). The graph \(G\) represents stochastic

\footnote{Symbol \(\times\) denotes Cartesian set product.}