International Journal of Approximate Reasoning 58 (2015) 39–56

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Chain graph interpretations and their relations revisited

Dag Sonntag *, Jose M. Peña

ADIT, IDA, Linköping University, Sweden

ARTICLE INFO

Article history:
Received 28 April 2014
Received in revised form 2 December 2014
Accepted 3 December 2014
Available online 9 December 2014

Keywords:
Chain graphs
Lauritzen–Wermuth–Frydenberg interpretation
Andersson–Madigan–Perlman interpretation
Multivariate regression interpretation

ABSTRACT

In this paper we study how different theoretical concepts of Bayesian networks have been extended to chain graphs. Today there exist mainly three different interpretations of chain graphs in the literature. These are the Lauritzen–Wermuth–Frydenberg, the Andersson–Madigan–Perlman and the multivariate regression interpretations. The different chain graph interpretations have been studied independently and over time different theoretical concepts have been extended from Bayesian networks to also work for the different chain graph interpretations. This has however led to confusion regarding what concepts exist for what interpretation.

In this article we do therefore study some of these concepts and how they have been extended to chain graphs as well as what results have been achieved so far. More importantly we do also identify when the concepts have not been extended and contribute within these areas. Specifically we study the following theoretical concepts: Unique representations of independence models, the split and merging operators, the conditions for when an independence model representable by one chain graph interpretation can be represented by another chain graph interpretation and finally the extension of Meek's conjecture to chain graphs. With our new results we give a coherent overview of how each of these concepts is extended for each of the different chain graph interpretations.

1. Introduction

Chain graphs (CGs) are hybrid graphs with two types of edges representing different types of relationships between the random variables of interest. These are the directed edges representing asymmetric relationships and a secondary type of edge representing symmetric relationships. Hence CGs extend Pearl's classical interpretation of directed and acyclic graphs (DAGs), i.e. Bayesian networks (BNs). However, there exist three different interpretations of CGs in research. These are the Lauritzen–Wermuth–Frydenberg (LWF) interpretation presented by Lauritzen, Wermuth and Frydenberg in the late nineteen eighties [9,11], the Andersson–Madigan–Perlman (AMP) interpretation presented by Andersson, Madigan and Perlman in 2001 [2] and the multivariate regression (MVR) interpretation presented by Cox and Wermuth in the nineteen nineties [6,7]. A fourth interpretation of CGs can also be found in a study by Drton [8] but this interpretation has not been further studied and will not be discussed in this paper.

Each interpretation has a different separation criterion and does therefore represent different independence models. Many papers have studied these independence models and extended many theoretical concepts regarding independence models from BNs to also work for CGs. Most of these papers have however only looked at one interpretation at a time,

* Corresponding author.
E-mail addresses: dag.sonntag@liu.se (D. Sonntag), jose.m.pena@liu.se (J.M. Peña).

http://dx.doi.org/10.1016/j.ijar.2014.12.001

which has led to an incoherent picture of what theoretical concepts exist for the different CG interpretations. Moreover, this has caused research on some concepts to be missing.

In this paper we do therefore look into some of these concepts and study how they are extended to the different CG interpretations to give a coherent overview of the research performed. More importantly, we do also identify where the concepts have not yet been extended to certain CG interpretations and contribute in different ways within these areas. Specifically we look into four areas that in different ways connect to the independence models of CGs. The first area is what unique representations exist for the different independence models representable by the different CG interpretations. Having such unique representations is important since there might exist multiple CGs representing the same independence model even for the same CG interpretation. The second area concerns the feasible split and feasible merging operators. These operators are used for altering the structure of a CG without altering which Markov equivalence class it belongs to. The third area we look into is what the conditions are for when an independence model represented by one CG interpretation also can be represented by another CG interpretation. This is important since it allows us to see when the different CG interpretations overlap in terms of representable independence models. The fourth and final area concerns Meek's conjecture and whether it can be extended to the different CG interpretations. Meek's conjecture states that given two DAGs G and H, s.t. the independence model represented by G includes the independence model represented by H, we can transform G into H through a sequence of operations s.t. the independence model represented by G includes the independence model of H for all intermediate DAGs G. The operations consist in adding a single directed edge to G, or replacing G with a Markov equivalent DAG. The validity of the conjecture was proven by Chickering in 2002 [4] and has allowed several learning algorithms for DAGs to be constructed.

Our contribution, in addition to a study of previous research in the area, is then the following definitions, examples and algorithms, together with their proofs of correctness, that previously have been missing:

- The definitions of the feasible split and feasible merging operators for AMP CGs and proof that for any two Markov equivalent AMP CGs G and H there exists a sequence of feasible splits and mergings that transforms G into H.
- An example showing there are no unique representatives of equivalence classes of MVR CGs that are MVR CGs.
- An algorithm that from any AMP CG G outputs the Markov equivalent AMP essential CG H.
- The necessary and sufficient conditions for when an independence model represented by a MVR CG can be perfectly represented by a CG in another interpretation and vice versa.
- An example that proves that Meek's conjecture does not hold for MVR CGs.

The remainder of the article is organized as follows. In the next section we present the notation we use throughout the article. In Section 3 we discuss the unique representations and in Section 4 we define the feasible split and merging operators. Section 5 contains the necessary and sufficient conditions for when an independence model represented by a CG in one interpretation can be perfectly represented by a CG in another interpretation. In Section 6 we then discuss Meek's conjecture and prove that this does not hold for MVR CGs. Finally we do a short summary and conclusion in Section 7.

To improve readability of the article we have chosen to move most of the theorems, lemmas and proofs to appendices. The article does therefore include three appendices, Appendices A, B and C, that contain the theorems, lemmas and proofs of Sections 3, 4 and 5 respectively.

2. Notation

All graphs are defined over a finite set of discrete or continuous random variables V. If a graph G contains an edge between two nodes V_1 and V_2, we denote with $V_1 \rightarrow V_2$ a *directed edge*, with $V_1 \leftrightarrow V_2$ a *bidirected edge* and with $V_1 - V_2$ an *undirected edge*. By $V_1 \circ\!\!\rightarrow V_2$ we mean that either $V_1 \rightarrow V_2$ or $V_1 \leftrightarrow V_2$ is in G. By $V_1 -\!\circ V_2$ we mean that either $V_1 \rightarrow V_2$ or $V_1 - V_2$ is in G. By $V_1 \circ\!-\!\circ V_2$ we mean that there exists an edge between V_1 and V_2 in G while we with $V_1 \cdots V_2$ mean that there might or might not exist an edge between V_1 and V_2. By a *non-directed* edge we mean either a bidirected edge or an undirected edge. A set of nodes is said to be *complete* if there exist edges between all pairs of nodes in the set.

The *parents* of a set of nodes X of G is the set $pa_G(X) = \{V_1 | V_1 \rightarrow V_2$ is in G, $V_1 \notin X$ and $V_2 \in X\}$. The *children* of X is the set $ch_G(X) = \{V_1 | V_2 \rightarrow V_1$ is in G, $V_1 \notin X$ and $V_2 \in X\}$. The *spouses* of X is the set $sp_G(X) = \{V_1 | V_1 \leftrightarrow V_2$ is in G, $V_1 \notin X$ and $V_2 \in X\}$. The *neighbors* of X is the set $nb_G(X) = \{V_1 | V_1 - V_2$ is in G, $V_1 \notin X$ and $V_2 \in X\}$. The *boundary* of X is the set $bd_G(X) = pa_G(X) \cup nb_G(X) \cup sp_G(X)$. The *adjacents* of X is the set $ad_G(X) = bd_G(X) \cup ch_G(X)$.

To exemplify these concepts we can study the graph G with five nodes shown in Fig. 1a. In the graph we can see two bidirected edges, one between B and D and one between D and E. Hence we know the spouses of D are B and E. G also contains two directed edges between A and B and B and E and we can see that E is the only child of B and B is the only child of A. Finally G also contains one undirected edge between C and D and hence C is a neighbor of D. All and all this means that the boundary of B is A and D while the adjacents of B also contains E in addition to A and D.

A *route* from a node V_1 to a node V_n in G is a sequence of nodes $V_1, \ldots, V_n$ s.t. $V_i \in ad_G(V_{i+1})$ for all $1 \leq i < n$. A *section* of a route is a maximal (w.r.t. set inclusion) non-empty set of nodes $B_1, \ldots, B_n$ s.t. the route contains the subpath $B_1 - B_2 - \ldots - B_n$. It is called a *collider section* if $B_1, \ldots, B_n$ together with the two neighboring nodes in the route, A and C, form the subpath $A \rightarrow B_1 - B_2 - \ldots - B_n \leftarrow C$. For any other configuration the section is a non-collider section. A *path* is a route containing only distinct nodes. The length of a path is the number of edges in the path. A path is *descending* if