A Simple Method for Identifying Compelled Edges in DAGs

S.K.M. Wong and **D. Wu**
Department of Computer Science
University of Regina
Regina Saskatchewan
Canada S4S 0A2
Email: {wong, danwu}@cs.uregina.ca

Abstract

Identifying compelled edges is important in learning the structure (i.e., the DAG) of a Bayesian network. A graphical method (Chickering 1995) was proposed to solve this problem. In this paper, we show that a joint probability distribution defined by a Bayesian network can be uniquely characterized by its *intrinsic factorization*. Based on such an algebraic characterization, we suggest a simple algorithm to identify the compelled edges of a Bayesian network structure.

1. Introduction

Bayesian networks have been widely used for managing uncertainty using probability (Pearl 1988). A Bayesian network consists of a graphical structure, a *directed acyclic graph* (DAG) and a set of *conditional probability distributions* (CPDs). These two components define a *joint probability distribution*.

The conditional independencies satisfied by a joint probability distribution can be represented by different DAGs. The notion of an equivalence class of DAGs was studied by (Verma & Pearl 1990; Chickering 1995). An equivalence class of DAGs consists of all the Bayesian networks which define the same joint probability distribution but differ in their respective DAGs.

Although different DAGs in the same equivalence class have different graphical structures, certain directed edges retain their directionality in all those equivalent DAGs. These edges are referred to as *compelled edges*. It is useful to identify these compelled edges in learning a Bayesian network from data. Chickering (Chickering 1995) proposed a graphical algorithm for identifying such edges.

In this paper, we show that a joint probability distribution defined by a Bayesian network can be uniquely expressed as a product of marginals divided by another product of marginals. We refer to such a characterization of a joint probability distribution as an *intrinsic factorization*. All Bayesian network structures having the same intrinsic factorization belong to the same equivalence class. Based on this algebraic classification of Bayesian networks, we suggest a simple method to identify all the compelled edges in a given DAG.

The paper is organized as follows. We first review the notion of equivalent DAGs in Section 2. Our algebraic characterization of a joint probability distribution defined by a Bayesian network is discussed in Section 3. In Section 4, we describe our algorithm for identifying compelled edges of a Bayesian DAG. The conclusion is given in Section 5.

2. Equivalent DAGs

In the paper, we use upper case letters to represent set of variables and lower case letters to represent a single variable unless otherwise mentioned. Let U denote a finite set of discrete variables. We use $p(U)$ to denote a *joint probability distribution* (jpd) over U. We call $p(V)$, $V \subset U$, the *marginal* (distribution) of $p(U)$ and $p(X|Y)$ the *conditional probability distribution* (CPD), or simply *conditional*. By XY, where $X \subseteq U, Y \subseteq U$, we mean $X \cup Y$. Similarly, by ab, where $a \in U, b \in U$, we mean $\{a\} \cup \{b\}$. According to the definition of conditional probability,

$$p(X|Y) = \frac{p(XY)}{p(Y)}, \text{whenever } p(Y) > 0,$$

we thus say that in the above expression, the denominator $p(Y)$ is *absorbed* by the numerator $p(XY)$ to yield conditional $p(X|Y)$. It is noted that $p(W)$ can be absorbed by $p(V)$ to yield $p(V - W|W)$ if and only if $W \subset V$.

A Bayesian network (BN) defined over a set $U = \{a_1, a_2, \ldots, a_n\}$ of variables consists of two components: (i) a *directed acyclic graph* (DAG) $\mathcal{D}$. Each vertex in $\mathcal{D}$ represents a variable $a_i \in U$. The parents of the vertex a_i are denoted $pa(a_i)$. The graphical structure of the DAG encodes a set of conditional independency information; (ii) a *quantification* of $\mathcal{D}$. Each variable a_i in $\mathcal{D}$ is quantified with a conditional probability $p(a_i|pa(a_i))$. These two components define (induce) a *joint probability distribution* (jpd) over U as follows:

$$p(U) \quad = \quad \prod_{i=1}^{n} p(a_i|pa(a_i)). \quad (1)$$

We call the equation in (1) the *Bayesian factorization*. It is worth mentioning that by looking at the factors in the above Bayesian factorization, the DAG can be drawn by directing an arrow from each node in $pa(a_i)$ to the node a_i. A $v - structure$ in a DAG $\mathcal{D}$ is an ordered triple of nodes

$(a,\ b,\ c)$ such that (1) $\mathcal{D}$ contains edges $a \rightarrow b$ and $c \rightarrow b$, and (2) a and c are not directly connected in $\mathcal{D}$. The *skeleton* of a DAG is the undirected graph obtained by dropping the directionality of every edge in the DAG. Since a BN is highlighted by its graphical structure, namely, the DAG, We thus will use the term BN or DAG interchangeably if no confusion arises.

It has been noted (Verma & Pearl 1990) that DAGs of different BNs may define the same jpd.

Definition 1 (Chickering 1995) Two DAGs $\mathcal{D}$ and $\mathcal{D}'$ are *equivalent* if for every BN induced by $\mathcal{D}$, there exists another BN induced by $\mathcal{D}'$ such that both induced BNs defines the same jpd, and vice versa.

We use $\mathcal{D}_1 \approx \mathcal{D}_2$ to denote that DAG $\mathcal{D}_1$ is equivalent to DAG $\mathcal{D}_2$, or equivalently, $BN1 \approx BN2$.

Graphical criteria (Verma & Pearl 1990; Chickering 1995) have been proposed to determine whether two given DAGs are equivalent or not.

Theorem 1 (Verma & Pearl 1990) Two DAGs are equivalent if and only if they have the same skeleton and the same v-structures.

It is not hard to see that the relation "$\approx$" induces an equivalence relation. Let $\{\mathcal{D}\}$ denote all the DAGs defined over a fixed set U of variables. Naturally, the characteristics of the equivalence relation $\approx$ can be used to group the DAGs in $\{\mathcal{D}\}$ into different equivalence classes.

It is noted that since equivalent DAGs define the same jpd, they certainly encode the same set of conditional independency information. Therefore, all the DAGs in the same equivalence class encode the same conditional independency information.
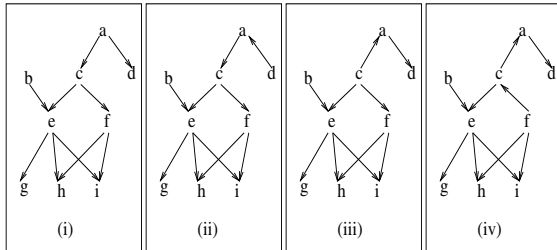

Figure 1: Four equivalent DAGs comprise an equivalence class.

Example 1 Consider the four DAGs shown in Figure 1. By Theorem 1, they are equivalent to each other. Moreover, no other DAGs defined over the same set of variables are equivalent to any of them, therefore, they comprise an equivalent class.

Given an equivalence class of DAGs, it is noticed that some edges retain their directionality in all the equivalent DAGs in the equivalence class.

Definition 2 An edge $x \rightarrow y$ in a DAG $\mathcal{D}$ is a *compelled* edge, if for any $\mathcal{D}' \approx \mathcal{D}$, $x \rightarrow y$ is also in $\mathcal{D}'$. Otherwise, it is *reversible*.

Example 2 It can be verified that in the DAGs shown in Fig 1, the edges $b \rightarrow e$, $c \rightarrow e$, $e \rightarrow g$, $e \rightarrow h$, $f \rightarrow h$, $e \rightarrow i$, and $f \rightarrow i$ appear in all of the equivalent DAGs, therefore, they are compelled edges. On the other hand, the edge $a \rightarrow c$, for example, is reversible .

It has been discussed in (Chickering 1995) that the compelled edges in a DAG have particular importance in learning the structure of Bayesian networks from observed data. Moreover, an algorithm was developed to identify all compelled edges in a given DAG. The algorithm first defines a total ordering over all the edges in the give DAG and then finds out all the compelled edges based on the total ordering. However, the presentation of the algorithm and the proof of its correctness are rather complicated and hard to be comprehended. For details of the discussion of the importance on identifying compelled edges and the algorithm, readers are referred to (Chickering 1995).

3. Algebraic Representation of Equivalent DAGs

In this section, we first briefly review different *graphical* representations of equivalence class of DAGs, we then give a novel *algebraic* representation of equivalent DAGs by studying the form of Bayesian factorization. This algebraic representation of equivalent DAGs will serve as the basis for a simple and easy to understand method for identifying compelled edges.

Since all the DAGs in the same equivalence class define the same jpd and represent the same conditional independency information, it is thus highly desirable to have a unique representation to represent the whole class of equivalent DAGs. Several graphical representations have been proposed. Verma (Verma & Pearl 1990) proposed the notion of *rudimentary pattern* and *complete pattern* to characterize the equivalence class using a partially directed graph. Anderson (Andersson, Madigan, & Perlman 1997) suggested the notion of *essential graph*, which is a special chain graph (Frydenberg 1990), to represent Markov equivalent DAGs. More recently, Studeny (Studeny 1998) used the notion of *largest chain graph* (different than the essential graph) to characterize the equivalence class of Bayesian networks.

Since all the equivalent DAGs also defines the *same* jpd, we believe that this "same" jpd should be the representation of the whole equivalence class and it should also possess the ability to discern different DAGs in the same equivalence class. In the following, we present a novel *algebraic* representation to represent the whole equivalence class based on the notion of Bayesian factorization.

Definition 3 Consider a DAG defined over a set $U = \{a_1,\ \ldots,\ a_n\}$ of variables with its Bayesian factorization as follows:

$$
\begin{aligned}
p(U) &= p(a_1) \cdot p(a_2|pa(a_2)) \cdot \ldots \cdot p(a_n|pa(a_n)), \quad (2)\\
&= \frac{p(a_1)}{1} \cdot \frac{p(a_2,\ pa(a_2))}{p(pa(a_2))} \cdot \ldots \cdot \frac{p(a_n,\ pa(a_n))}{p(pa(a_n))}, (3)\\
&= \prod_{i,\ j} \frac{p(a_i,\ pa(a_i))}{p(pa(a_j))}, \quad (4)
\end{aligned}
$$

where $\{a_i,\ pa(a_i)\} \neq \{pa(a_j)\}$ for any $1 \leq i,\ j \leq n$. Each $p(a_i|pa(a_i))$ in equation (2) is called a *factor*. We call the equation in (3) the *fraction factorization* of the DAG. The expression in equation (4) is obtained by canceling any applicable numerator and denominator in equation (3) and is called the *intrinsic factorization* of the DAG.

The intrinsic factorization has been proved to be the invariant property of an equivalence class of DAGs.

Theorem 2 (Wong & Wu 2002) Two DAGs are *equivalent* if and only if they have the same intrinsic factorization.

Theorem 2 implies that all the DAGs in the same equivalence class have the same intrinsic factorization. It immediately follows:

Corollary 1 The intrinsic factorization of a given DAG $\mathcal{D}$ is *unique* and it characterizes and describes algebraically the whole equivalence class that the given DAG $\mathcal{D}$ belongs to.

Example 3 Consider the four DAGs defined over $U = \{a,\ b,\ c,\ d,\ e,\ f,\ g,\ h,\ i\}$ as shown in Fig 1. The intrinsic factorization for the DAG in (i) is as follows in equation (6):

$$
\begin{aligned}
p(U) &= p(a) \cdot p(b) \cdot p(c|a) \cdot p(d|a) \cdot p(e|bc) \\
&\quad p(f|c) \cdot p(g|e) \cdot p(h|ef) \cdot p(i|ef) \quad (5) \\
&= \frac{p(b) \cdot p(ca) \cdot p(da) \cdot p(ebc) \cdot p(fc) \cdot p(ge)}{p(a) \cdot p(bc) \cdot p(c) \cdot p(e)} \cdot \\
&\quad \frac{p(hef) \cdot p(ief)}{p(ef) \cdot p(ef)}. \quad (6)
\end{aligned}
$$

Similarly, the intrinsic factorization for the DAGs in Fig 1 (ii), (iii) and (iv) can be obtained in the same fashion. It can be easily verified that the four different DAGs do have the same intrinsic factorization shown in equation (6).

4. Compelled Edges Identification

In this section, we will present a very simple method to identify compelled edges in a given DAG based on its intrinsic factorization. The idea behind our proposed method is very intuitive. As will be demonstrated shortly, the intrinsic factorization of a given DAG represents the whole equivalent class and it contains all the information needed to restore each equivalent DAG (more precisely, to restore its Bayesian factorization.) This observation gives rise to the proposed method for identifying compelled edges.

Example 4 Continue on Example 3 and consider the intrinsic factorization of the Bayesian network shown in Figure 1 (i) repeated as follows:

$$
\begin{aligned}
p(U) &= \frac{p(b) \cdot p(ca) \cdot p(da) \cdot p(ebc) \cdot p(fc) \cdot p(ge)}{p(a) \cdot p(bc) \cdot p(c) \cdot p(e)} \cdot \\
&\quad \frac{p(hef) \cdot p(ief)}{p(ef) \cdot p(ef)}. \quad (7)
\end{aligned}
$$

We now demonstrate how we can restore each equivalent DAG (its Bayesian factorization) in the equivalence class characterized by the above intrinsic factorization.

In order to transform the intrinsic factorization in equation (7) into a Bayesian factorization, we need to absorb all the denominators in equation (7) as follows.

$$
\begin{aligned}
p(U) &= \frac{p(b) \cdot p(ca) \cdot p(da) \cdot p(ebc) \cdot p(fc) \cdot p(ge)}{p(a) \cdot p(bc) \cdot p(c) \cdot p(e)} \cdot \\
&\quad \frac{p(hef) \cdot p(ief)}{p(ef) \cdot p(ef)}, \quad (8) \\
&= \frac{p(b)}{1} \cdot \frac{p(ca)}{p(a)} \cdot \frac{p(da)}{1} \cdot \frac{p(ebc)}{p(bc)} \cdot \frac{p(fc)}{p(c)} \cdot \\
&\quad \frac{p(ge)}{p(e)} \frac{p(hef)}{p(ef)} \cdot \frac{p(ief)}{p(ef)}, \quad (9) \\
&= p(b) \cdot p(c|a) \cdot p(da) \cdot p(e|bc) \cdot p(f|c) \cdot p(g|e) \\
&\quad \cdot p(h|ef) \cdot p(i|ef), \quad (10) \\
&= p(b) \cdot p(c|a) \cdot p(d) \cdot p(a|d) \cdot p(e|bc) \cdot p(f|c) \cdot \\
&\quad p(g|e) \cdot p(h|ef) \cdot p(i|ef), \quad (11)
\end{aligned}
$$

or

$$
\begin{aligned}
p(U) &= \frac{p(b)}{1} \cdot \frac{p(ca)}{1} \cdot \frac{p(da)}{p(a)} \cdot \frac{p(ebc)}{p(bc)} \cdot \frac{p(fc)}{p(c)} \cdot \\
&\quad \frac{p(ge)}{p(e)} \frac{p(hef)}{p(ef)} \cdot \frac{p(ief)}{p(ef)}, \quad (12) \\
&= p(b) \cdot p(c) \cdot p(a|c) \cdot p(d|a) \cdot p(e|bc) \cdot p(f|c) \cdot \\
&\quad p(g|e) \cdot p(h|ef) \cdot p(i|ef), \quad (13)
\end{aligned}
$$

or

$$
\begin{aligned}
p(U) &= \frac{p(b)}{1} \cdot \frac{p(ca)}{p(c)} \cdot \frac{p(da)}{p(a)} \cdot \frac{p(ebc)}{p(bc)} \cdot \frac{p(fc)}{1} \cdot \\
&\quad \frac{p(ge)}{p(e)} \frac{p(hef)}{p(ef)} \cdot \frac{p(ief)}{p(ef)}, \quad (14) \\
&= p(b) \cdot p(a|c) \cdot p(d|a) \cdot p(e|bc) \cdot p(f) \cdot p(c|f) \cdot \\
&\quad p(g|e) \cdot p(h|ef) \cdot p(i|ef), \quad (15)
\end{aligned}
$$

In equations (9), (12), and (14), we have absorbed each denominator in the intrinsic factorization by an appropriate numerator, these absorptions resulted in equations (11), (13), and (15), respectively. We thus have finally obtained three different Bayesian factorizations, which exactly correspond to the three equivalent DAGs shown in Fig 1 (ii), (iii) and (iv), respectively. Therefore, we have successfully restored all the DAGs that are equivalent to the one shown in Fig 1 (i).

There are a few observations that can be made with respect to the above demonstration in Example 4.

1. Different DAGs in the same equivalence class are obtained, depending on how each denominator in the intrinsic factorization is absorbed. It is obvious that different absorption will result in different Bayeisn factorizations, hence, produce different but equivalent DAGs.

2. It is noted that during the course of absorbing denominators, some denominator is "forced" to be absorbed by a particular fixed numerator, no other choices. In other words, some denominator has no flexibility so that it has to be absorbed by a fixed numerator.

Before we move on, let's scrutinize the denominator absorption we have made in Example 4. Note that for a denominator $p(X)$ to be absorbed by a numerator $p(Y)$, it must be the case that $X \subset Y$. Under this restriction, the possible absorption of the denominators a, bc, c, e, ef, and ef can be summarized by the following expressions,

$$a \implies \{ca, da\}, \quad (16)$$
$$bc \implies \{ebc\}, \quad (17)$$
$$c \implies \{ca, ebc, fc\}, \quad (18)$$
$$e \implies \{ebc, ge, hef, ief\}, \quad (19)$$
$$ef \implies \{hef, ief\}, \quad (20)$$
$$ef \implies \{hef, ief\}, \quad (21)$$

where $X \implies \{Y_1, \ldots, Y_n\}$ means that the denominator $p(X)$ can possibly be absorbed by $p(Y_1)$, $\ldots$, $p(Y_n)$. We further call the set $\{Y_1, \ldots, Y_n\}$, denoted $AS(X)$, the *absorption set* for X. It is noted that $Y_i \in AS(X)$ if $X \subset Y_i$. We will use $|AS(X)|$ to denote its cardinality. It is also noted that the set $D = \{X_i \mid p(X_i)$ is a denominator in the intrinsic factorization$\}$ is a *multiset*, and we will use $|X_i|$ to denote the number of occurrence of X_i in the multiset D.

Recall that the DAG of a Bayesian network can be drawn based on its Bayesian factorization, more precisely, based on its factor $p(a_i|pa(a_i))$, by directing an edge from each node in $pa(a_i)$ to the node a_i. Following this line of reasoning, it follows that if a denominator, say $p(X)$, has no choice but is "forced" to be absorbed by a fixed numerators, say $p(Y)$, then it is expected that the edges $y \to x$, where $y \in Y - X$, $x \in X$, will appear in every resulting Bayesian factorizations. Therefore, they will be compelled edges by definition. For instance, in expression (17), the only numerator that is applicable to absorb $p(bc)$ is $p(ebc)$, hence, the factor $p(e|bc)$ will be obtained by this absorption and the edges $b \to e$ and $c \to e$ will appear in all possible resulting Bayesian factorizations, which implies that they are compelled edges as can be verified by Fig 1. Special attention should be paid to the two identical denominators $p(ef)$ and $p(ef)$ in expressions (20) and (21), respectively. Although these two denominators are syntactically identical, both of them have to be absorbed in order to obtain the Bayesian factorization. The applicable numerators for both of them is the absorption set $\{hef, ief\}$, which contains exactly two elements, the same number of the occurrence of $p(ef)$ as denominators. This indicates that one of the $p(ef)$ must be absorbed by $p(hef)$, and the other $p(ef)$ must be absorb by $p(ief)$, no other choices. These absorptions imply that the factors $p(h|ef)$ and $p(i|ef)$ will be obtained and the edges $e \to h$, $f \to h$, $e \to i$, and $f \to i$ will appear in all possible resulting Bayesian factorizations, which implies that they are compelled edges as can be verified by Fig 1 as well. Since the numerators $p(hef)$, $p(ief)$, and $p(ebc)$ have been designated to absorb the denominators $p(ef)$, $p(ef)$, and $p(bc)$, respectively, this changes the absorption set for e from $\{ebc, ge, hef, ief\}$ shown in expression (19) to the new refined singleton set $\{ge\}$, which implies that the denominator $p(e)$ will have to be absorbed by the numerator $p(ge)$ to obtain $p(g|e)$. Therefore, the edge $e \to g$ is also a compelled edges as can be verified by Fig 1.

The above analysis resulted in the following refined expressions for the absorptions of denominators, contrasting with those in the expressions (16-21),

$$a \implies \{ca, da\}, \quad (22)$$
$$bc \implies \{ebc\}, \quad (23)$$
$$c \implies \{ca, fc\}, \quad (24)$$
$$e \implies \{ge\}, \quad (25)$$
$$ef \implies \{hef, ief\}, \quad (26)$$
$$ef \implies \{hef, ief\}, \quad (27)$$

from which the compelled edges can be found out right away if a denominator can only be absorbed by a fixed denominator.

Based on the above discussions, we thus propose the following procedure to identify compelled edges in a given DAG. The correctness of the procedure will be given shortly.

PROCEDURE Find-Compelled-Edges
Input: a given DAG $\mathcal{D}$.
Output: Compelled edges in $\mathcal{D}$ collected in the set E.
{
  1: Obtain the intrinsic factorization of the given $\mathcal{D}$, and let $E = \emptyset$.
  2: Let $D = \{d_1, \ldots, d_m\}$ be a *multiset*, where $p(d_i)$ is a denominator in the intrinsic factorization obtained in step 1.
  3: For each $d_i \in D$, $i = 1, \ldots, m$, compute d_i's absorption set $AS(d_i)$;
  4: For each $d_i \in D$ s.t. $|d_i| = 1$, $i = 1, \ldots, m$, If $AS(d_i)$ is a singleton set, $AS(d_j) = AS(d_j) - AS(d_i)$, for all $j \neq i$,
  5: For each $d_i \in D$ s.t. $|d_i| > 1$, $i = 1, \ldots, m$, If $|AS(d_i)| = |d_i|$, $AS(d_j) = AS(d_j) - AS(d_i)$, for all $j \neq i$,
  6: For each $d_i \in D$, $i = 1, \ldots, m$, If $k = |AS(d_i)| = |d_i|$, $E = E \cup \{p(Y_j - d_i|d_i)|Y_j \in AS(d_i), j = 1, ..., k\}$,
  7: Return E.
}

Compelled edges, from the perspective of the intrinsic factorization of a given DAG, is obtained by identifying whether a denominator $p(X)$ can be absorbed by only one numerator or can be possibly absorbed by multiple numerators. If a denominator $p(X)$ can only be absorbed by only one numerator $p(Y)$, then the factor $p(Y - X|X)$ will appear in all possible resulting Bayesian factorizations, therefore, the edges $y \to x$, where $y \in Y - X$, $x \in X$, will be compelled edges.

Step 1-5 in the procedure computes the absorption set for each denominator. In step 6, for any denominator $p(X)$ and its absorption set $AS(X)$, we try to see whether it can be absorbed by a unique numerator or by multiple choices of different numerators. There are two cases:

1. if $p(X)$ is a denominator such that $|X| = 1$, consider $AS(X)$.

- If $|AS(X)| > 1$, in other words, $p(X)$ can be absorbed by at least two different numerators, then there is no chance to have any compelled edges produced by absorbing $p(X)$.
- If $|AS(X)| = 1$, in other words, $p(X)$ will be absorbed by only one denominator $p(Y)$, then the factor $p(Y - X|X)$ obtained will give rise to compelled edges $y \rightarrow x$, where $y \in Y - X$, $x \in X$.

2. if $p(X)$ is a denominator such that $|X| > 1$, consider $AS(X)$.

- If $|AS(X)| = |X|$, then there exists a one-to-one correspondence between each member of $Y \in AS(X)$ and each occurrence of $p(X)$ as denominator, in which case every $p(X)$ will be absorbed by a unique $p(Y)$. Therefore, compelled edges will be obtained by such absorptions.
- If $|AS(X)| > |X|$, then it has more numerator $p(Y)$, where $Y \in AS(X)$ than the denominator $p(X)$, in which case each denominator is *not* uniquely absorbed by a numerator. Therefore, there is no chance to obtain the compelled edges.

The analysis of the two cases has been exactly captured by step 6 in the procedure.

The above discussion actually proves the following theorem.

Theorem 3 Given a DAG $\mathcal{D}$, the output of the procedure "Find-Compelled-Edges" above is exactly the compelled edges in $\mathcal{D}$.

The procedure "Find-Compelled-Edges" only involves some set operations and can be implemented easily.

5. Conclusion

In this paper, we have demonstrated that the notion of intrinsic factorization can serve as the algebraic characterization of the equivalence class of DAGs, based on which we have presented a simple method for identifying compelled edges in a Bayesian network. The new method is intuitively simple and can be easily implemented. It reveals the importance of the intrinsic factorization of an equivalence class of Bayesian networks. The application of intrinsic factorization in this paper suggests more research on the factorization of the jpd defined by a Bayesian network is worth pursuing.

References

Andersson, S.; Madigan, D.; and Perlman, M. 1997. A characterization of markov equivalence classes for acyclic digraphs. *The Annals of Statistics* 25(2):505–541.

Chickering, D. 1995. A transformational characterization of equivalent bayesian network structures. In *Eleventh Conference on Uncertainty in Artificial Intelligence*, 87–98. Morgan Kaufmann Publishers.

Frydenberg, M. 1990. The chain graph markov property. *Scandnavian Journal of Statistics* 11:333–353.

Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, California: Morgan Kaufmann Publishers.

Studeny, M. 1998. Bayesian networks from the point of view of chain graphs. In *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligernce (UAI-98)*, 496–503. Morgan Kaufmann.

Verma, T., and Pearl, J. 1990. Equivalence and synthesis of causal models. In *Sixth Conference on Uncertainty in Artificial Intelligence*, 220–227. GE Corporate Research and Development.

Wong, S., and Wu, D. 2002. An algebraic characterization of equivalent bayesian networks. In Neumann;, M. M. B., and Studer, R., eds., *Intelligent Information Processing, The 17th IFIP World Computer Congress, Stream 8, IIP-2002*, 177–189. Kluwer Academic Publisher.