On Axiomatizing Probabilistic Conditional Independencies in Bayesian Networks

C.J. Butz

School of Information Technology & Engineering
University of Ottawa, Ottawa, ON, K1N 6N5, Canada

Abstract. Several researchers have suggested that *Bayesian networks* (BNs) should be used to manage the inherent uncertainty in information retrieval. However, it has been argued that manually constructing a large BN is a difficult process. In this paper, we obtain the only *minimal complete* subset of the semi-graphoid axiomatization governing the independency information in a BN. This result may be useful in developing an automated BN construction procedure for information retrieval purposes.

1 Introduction

Probability theory provides a rigorous foundation for the management of uncertain knowledge [4]. We may assume that knowledge is represented as a joint probability distribution. The probability of an event can be obtained (in principle) by an appropriate marginalization of the joint distribution. Obviously, it may be impractical to obtain the joint distribution directly: for example, one would have to specify 2^n entries for a distribution over n binary variables. However, Bayesian networks utilize probabilistic conditional independencies that are assumed to hold in the problem domain to indirectly obtain the required joint probabilities. A *Bayesian network* (BN) [4] consists of a *directed acyclic graph* (DAG) and a corresponding set of conditional probability distributions. The DAG encodes all of the probabilistic conditional independencies satisfied by a particular joint distribution. Thus, BNs provide a semantic modeling tool which facilitate the acquisition of probabilistic knowledge. Several researchers have suggested that BNs should be used to manage the inherent uncertainty in information retrieval. For instance, it was recently shown [5] how retrieval performance can be improved by using BNs to combine document content with the link structure of the Web. However, it has been argued that manually constructing a large BN is a difficult process [3]. Thereby, it would be desirable if an automated procedure could be developed for constructing a BN for information retrieval.

The *semi-graphoid axiomatization* is a set $\{(SG1), (SG2), (SG3), (SG4)\}$ of four inference axioms respectively called symmetry, decomposition, weak union, and contraction. Dawid [2] originally showed that these axioms were *sound* for the implication of probabilistic conditional independence. Pearl [4] realized that these axioms were in fact *complete* for inferring all conditional independencies in a BN. However, Pearl [4] incorrectly conjectured that the semi-graphoid axiomatization was complete for probabilistic conditional independence. This conjecture sparked a flurry of studies into the completeness problem, including [6, 7, 11] to name but a few. To the best of our knowledge, however, no investigation has ever questioned the claim that the semi-graphoid inference axioms are *independent*.

In this paper, we obtain the only minimal complete subset of the semi-graphoid axiomatization. The symmetry axiom (SG1) is stated as an iff inference axiom, while decomposition (SG2), weak union (SG3), and contraction (SG4) are all stated as if-then inference axioms. Pearl's claim that the semi-graphoid axioms are independent (see Theorem 1 in [4]) needs to be somewhat qualified. The key observation in this exposition is that the contraction axiom (SG4), like the symmetry axiom (SG1), is *not* an if-then axiom but is in fact an iff inference axiom $(SG4)'$. This means that the semi-graphoid axiomatization, traditionally written as $\{(SG1), (SG2), (SG3), (SG4)\}$, can be written as $\{(SG1), (SG2), (SG3), (SG4)'\}$. It immediately follows that the decomposition axiom (SG2) and the weak union axiom (SG3) can be removed from $\{(SG1), (SG2), (SG3), (SG4)'\}$, since any probabilistic conditional independence obtained by an application of $(SG2)$ or $(SG3)$ can be obtained by an application of $(SG4)'$. In other words, $\{(SG1), (SG4)'\}$ is a minimal complete subset of the semi-graphoid axiomatization. We also show that $\{(SG1), (SG4)'\}$ is the *only* subset enjoying this property. The important point to remember, however, is that the set $\{(SG1), (SG4)'\}$ *completely* characterizes all of the probabilistic conditional independencies holding in a BN.

This paper is organized as follows. The complete minimal subset of the semi-graphoid axiomatization is given in Section 2. The conclusion is presented in Section 3.

2 A Minimal Complete Subset of the Semi-Graphoid Axiomatization

We first introduce the fundamental notion of *probabilistic conditional independency*. Let X, Y and Z be disjoint subsets of variables in R. Let x, y, and z denote arbitrary values of X, Y and Z, respectively. We say Y and Z are *conditionally independent* given X under the joint probability distribution p, denoted $I_p(Y, X, Z)$, if

$$p(y \mid x, z) = p(y \mid x), \tag{1}$$

whenever $p(x, z) > 0$. This conditional independency $I_p(Y, X, Z)$ can be equivalently written as

$$p(y, x, z) = \frac{p(y, x) \cdot p(x, z)}{p(x)}. \tag{2}$$

We write $I_p(Y, X, Z)$ as $I(Y, X, Z)$ if the joint probability distribution p is understood.

Let $\sum$ be a set of probabilistic conditional independencies and σ be a single independency. We say $\sum$ *logically implies* σ, written $\sum \models \sigma$, if every distribution which satisfies $\sum$ also satisfies σ. That is, there is no counter-example distribution such that all of the independencies in $\sum$ are satisfied but σ is not. The *implication problem* [9] is to test whether a given set $\sum$ of independencies logically implies another independency σ, namely, $\sum \models \sigma$.

An *inference axiom* gives conditions as to when certain probabilistic conditional independencies must be satisfied by a distribution, provided that the distribution satisfies other given independencies. Given a set $\sum$ of independencies and a set of inference axioms, the *closure* of $\sum$, written $\sum^+$, is the smallest set containing $\sum$ such that the inference axioms cannot be applied to the set to yield a independency not in the set. More specifically, the set $\sum$ *derives* a independency σ, written $\sum \vdash \sigma$, if σ is in $\sum^+$. A set of inference axioms is *sound* if whenever $\sum \vdash \sigma$, then $\sum \models \sigma$. A set of inference axioms is *complete* if the converse holds, that is, if $\sum \models \sigma$, then $\sum \vdash \sigma$. In other words, saying a set of axioms are complete means that if $\sum$ logically implies the independency σ, then $\sum$ derives σ.

The semi-graphoid inference axioms are [4]:

$$
\begin{array}{ll}
(SG1): I(Y,X,Z) \text{ if and only if } I(Z,X,Y); & \text{[symmetry]} \\
(SG2): \text{If } I(Y,X,ZW), \text{ then } I(Y,X,Z); & \text{[decomposition]} \\
(SG3): \text{If } I(Y,X,ZW), \text{ then } I(Y,XW,Z); & \text{[weak union]} \\
(SG4): \text{If } I(Y,X,Z) \text{ and } I(Y,XZ,W), \text{ then } I(Y,X,ZW). & \text{[contraction]}
\end{array}
$$

These inference axioms were shown *sound* by Dawid [2]. This means that if a conditional independency σ is derived from an input set $\sum$ of conditional independencies using the semi-graphoid axiomatization, then $\sum$ logically implies σ. Pearl [4] realized that the converse was true, provided that $\sum$ is a causal input list. In other words, if the conditional independencies in $\sum$ define a BN, then the semi-graphoid axioms will derive every conditional independency logically implied by $\sum$. It should be emphasized that if $\sum$ is an arbitrary set of independencies, then the semi-graphoid axioms may not derive *every* independency logically implied by $\sum$. An example of such a case can be found in [6].

Pearl has stated that the semi-graphoid inference axioms are independent (see Theorem 1 in [4]). In other words, no inference axiom in {($SG1$), ($SG2$), ($SG3$), ($SG4$)} is redundant. This remark needs to be somewhat qualified.

Consider again the semi-graphoid axiomatization. Notice that the symmetry (SG1) is an iff inference axiom, while decomposition (SG2), weak union (SG3), and contraction (SG4) are all if-then inference axioms. The following result indicates that the contraction axiom (SG4), like the symmetry axiom (SG1), is in fact an iff inference axiom.

Theorem 1. *The contraction axiom (SG4) in Pearl's semi-graphoid axioms [4] is bidirectional, namely:*

$$(SG4)': \ I(Y,X,Z) \text{ and } I(Y,XZ,W) \textit{ if and only if } I(Y,X,ZW). \ \text{[contraction]}$$

Proof. In [2], Dawid showed that $\{I(Y,X,Z),\ I(Y,XZ,W)\} \models I(Y,X,ZW)\}$, $\{I(Y,X,ZW)\} \models I(Y,X,Z)$, and $\{I(Y,X,ZW)\} \models I(Y,XZ,W)$ all hold. $\square$

Theorem 1 is significant since it means that Pearl's semi-graphoid axioms $(SG1)$,$(SG2)$,$(SG3)$,$(SG4)$ can be written as $\{(SG1), (SG2), (SG3), (SG4)'\}$:

$(SG1):$ $I(Y,X,Z)$ if and only if $I(Z,X,Y)$; [symmetry]
$(SG2):$ If $I(Y,X,ZW),$ then $I(Y,X,Z)$; [decomposition]
$(SG3):$ If $I(Y,X,ZW),$ then $I(Y,XW,Z)$; [weak union]
$(SG4)':$ $I(Y,X,Z)$ and $I(Y,XZ,W)$ if and only if $I(Y,X,ZW)$. [contraction]

Our objective now is to determine whether we can obtain a set with *fewer* inference axioms without changing the closure, i.e., obtain an equivalent set of inference axioms by removing any redundant axioms. Clearly, axioms (SG2) and (SG3) are redundant, since any conditional independency obtained using (SG2) and (SG3) can be obtained by the new contraction axiom $(SG4)'$. Thus, we obtain the following result.

Theorem 2. *The following two inference axioms* $(SG1)$ *and* $(SG4)'$ *are a minimal complete subset of Pearl's semi-graphoid axiomatization:*

$(SG1):$ $I(Y,X,Z)$ if and only if $I(Z,X,Y)$; [symmetry]
$(SG4)':$ $I(Y,X,Z)$ and $I(Y,XZ,W)$ if and only if $I(Y,X,ZW)$. [contraction]

Proof. Clearly, $\{(SG4)'\}$ is equivalent to $\{(SG2),(SG3),(SG4)\}$. It immediately follows that $\{(SG1),(SG4)'\}$ is *equivalent* to $\{(SG1),(SG2),(SG3),(SG4)\}$.

We now show that $\{(SG1),(SG4)'\}$ is minimal. Again using the fact that $\{(SG4)'\}$ is equivalent to $\{SG2,SG3,SG4\}$, it immediately follows that $(SG1)$ is not redundant in $\{(SG1),(SG4)'\}$ since $(SG1)$ is not redundant in the semi-graphoid axioms $\{(SG1),(SG2),(SG3),(SG4)\}$ [4]. Obviously $(SG4)'\}$ is not redundant in $\{(SG1),(SG4)'\}$. Thus, $\{(SG1),(SG4)'\}$ is *minimal.* □

Theorem 2 is important since it indicates that a conditional independency σ can be obtained from a given set $\sum$ using $\{(SG1),(SG2),(SG3),(SG4)\}$ if and only if σ can be obtained from $\sum$ using $\{(SG1),(SG4)'\}$. This means that the set $\{(SG1),(SG4)'\}$ of inference axioms *completely* characterizes all of the probabilistic conditional independencies holding in a BN. It can be easily verified that $(SG2)$ and $(SG3)$ cannot be written as iff axioms.

Corollary 3. $\{(SG1),(SG4)'\}$ *is the only minimal complete subset of the semi-graphoid axiomatization* $\{(SG1),(SG2),(SG3),(SG4)'\}$.

3 Conclusion

The key observation in this paper is that the contraction axiom $(SG4)$, like the symmetry axiom $(SG1)$, is in fact an iff inference axiom $(SG4)'$. This means that the Pearl's [4] semi-graphoid axiomatization $\{(SG1),(SG2),(SG3),(SG4)\}$, can be succinctly written as $\{SG1,SG4'\}$. In fact, we have shown that this is the *only* minimal complete subset of the semi-graphoid axiomatization. This is important since it means that the set $\{(SG1), (SG4)'\}$ *completely* characterizes all probabilistic conditional independencies holding in a BN.

A study of axiomatization is not only important from a theoretical point of view. On the contrary, a complete axiomatization provides an algorithmic approach to the design of relational databases [1], BNs [8], and multi-agent BNs [10]. Thus, the results in this paper may be of some help in developing an automated method for constructing BNs for information retrieval purposes.

References

1. Beeri, C., Fagin, R., and Howard, J.H.: A complete axiomatization for functional and multivalued dependencies in database relations. Proc. of ACM-SIGMOD International Conference on Management of Data (1977) 47-61
2. Dawid, A.P.: Conditional independence in statistical theory. Journal of the Royal Statistical Society **41B** (1979) 1-31
3. Koller, D. and Pfeffer, A. Object-oriented bayesian networks. Proc. of the Thirteenth Conference on Uncertainty in Artificial Intelligence (1997) 302-313
4. Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers, San Francisco (1988)
5. Silva, I., Ribeiro-Neto, B., Calado, P., Moura, E., and Ziviani, N.: Link-based and content-based evidential information in a belief network model. Proc. of the Twenty-third Annual ACM SIGIR Conference on Research and Development in Information Retrieval (2000) 96-103
6. Studeny, M.: Multiinformation and the problem of characterization of conditional-independence relations. Problems of Control and Information Theory **18**(1) (1989) 3-16
7. Studeny, M.: Conditional independence relations have no finite complete characterization. Proc. of the Eleventh Prague Conference on Information Theory, Statistical Decision Foundation and Random Processes (1990) 377-396
8. Verma, T. and Pearl, J.: An algorithm for deciding if a set of observed independencies has a causal explanation. Proc. of the Eighth Conference on Uncertainty in Artificial Intelligence (1992) 323-330
9. Wong, S.K.M., Butz, C.J., and Wu, D.: On the implication problem for probabilistic conditional independency. IEEE Trans. Syst. Man Cybern. SMC-A **30**(6) (2000) 785-805
10. Wong, S.K.M. and Butz, C.J.: Constructing the dependency structure of a multi-agent probabilistic network. IEEE Trans. Knowl. Data Eng. **13**(3) (2001) (to appear)
11. Wong, S.K.M. and Wang, Z.W.: On axiomatization of probabilistic conditional independence. Proc. of the Tenth Conference on Uncertainty in Artificial Intelligence (1994) 591-597

This article was processed using the LaTeX macro package with LLNCS style