A Note on Equivalence Classes of Directed Acyclic Independence Graphs

David Madigan

Probability in the Engineering and Informational Sciences / Volume 7 / Issue 03 / July 1993, pp 409 - 412
DOI: 10.1017/S0269964800003004, Published online: 27 July 2009

Probability in the Engineering and Informational Sciences, 7, 1993, 409-412. Printed in the U.S.A.

A NOTE ON EQUIVALENCE CLASSES OF DIRECTED ACYCLIC INDEPENDENCE GRAPHS

DAVID MADIGAN

Department of Statistics, GN-22
University of Washington
Seattle, Washington 98195

1. INTRODUCTION AND PRELIMINARIES

Directed acyclic independence graphs (DAIGs) play an important role in recent developments in probabilistic expert systems and influence diagrams (Chyu [1]). The purpose of this note is to show that DAIGs can usefully be grouped into equivalence classes where the members of a single class share identical Markov properties. These equivalence classes can be identified via a simple graphical criterion. This result is particularly relevant to model selection procedures for DAIGs (see, e.g., Cooper and Herskovits [2] and Madigan and Raftery [4]) because it reduces the problem of searching among possible orientations of a given graph to that of searching among the equivalence classes.

Following Lauritzen, Dawid, Larsen, and Leimer [3], we consider a directed acyclic graph $G = (V,E)$ with v nodes representing discrete random variables X_i, $i = 1, \ldots, v$, and a set of directed edges E. The *undirected graph* associated with G is defined as $G^{\sim} = (V,E^{\sim})$ with the same vertex set and an undirected edge replacing each directed edge. We define the *morality* of G, denoted $\mathcal{M}(G)$ to be the set of triples (i,j,k) where $i,j,k \in V$, $j,k \in \text{pa}(i)$, and there is no edge linking j and k (j and k are "immoral" parents of i). A directed acyclic graph satisfies the Wermuth condition if no subgraph has the configuration shown in Figure 1, i.e., if no node has two or more "immoral parents." The *moral graph* associated with G is the undirected graph $G^M = (V,E^M)$ on the same vertex set

 0269-9648/93 $5.00 + .00

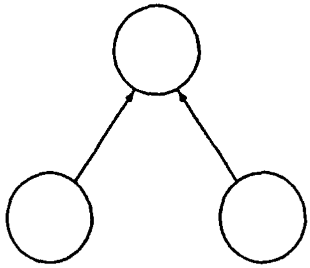

Figure 1. A forbidden Wermuth configuration.

and with an edge set obtained by including all edges in E together with all edges necessary to eliminate forbidden Wermuth configurations from G. If two DAIGs G_1 and G_2 have identical Markov properties, then we have $G_1 \sim G_2$.

2. EQUIVALENCE CLASSES AND MORALITY

It is well known that for a DAIG G, if $G^M = G^{\sim}$, then the Markov properties of G are identical to those of G^M (see, e.g., Whittaker [6, Chapter 3]). Let $G_1 = (V, E_1)$ be a DAIG that is derived from G by reorienting one or more edges. It follows immediately that if $G_1^M = G^M = G^{\sim}$, then $G_1 \sim G$. For example, consider the DAIGs of Figure 2. Here, $G_A^M = G_B^M = G_A^{\sim} = G_B^{\sim}$ and thus $G_A \sim G_B$.

Now consider a DAIG G for which $G^M \neq G^{\sim}$. We will show that any orientation of the edges of G that does not alter the morality of G does not alter the Markov properties of G. Before formally establishing this result, we consider the example in Figure 3. We have that $\mathcal{M}(G_A) = \mathcal{M}(G_B) = \mathcal{M}(G_C) = \{(3,2,4)\}$ and hence $G_A \sim G_B \sim G_C$ (for all three we have $3 \perp 1 | 2,4$ and $2 \perp 4 | 1$).

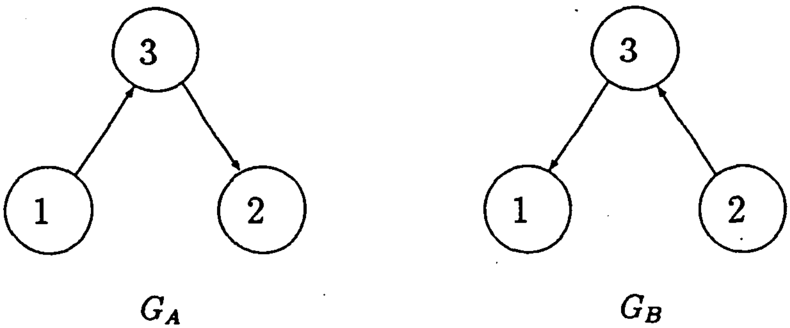

Figure 2.

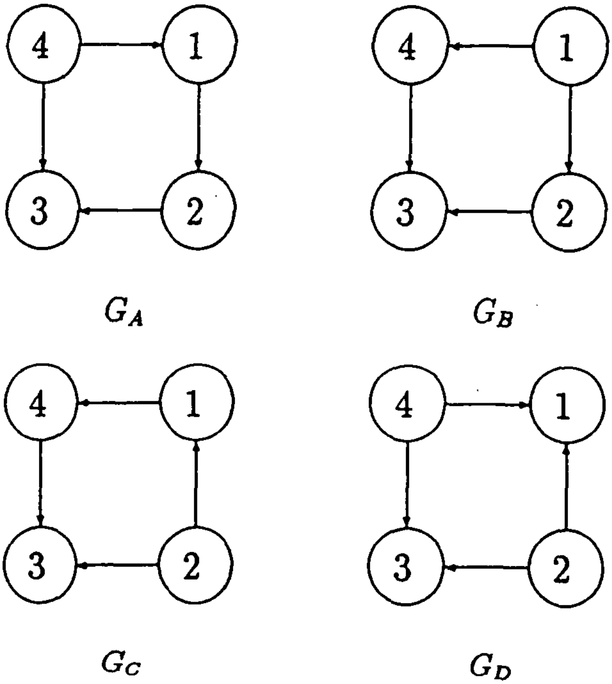

FIGURE 3.

However, $\mathcal{M}(G_D) = \{(3,2,4),(1,2,4)\}$ and the Markov properties of G_D are different from the other three ($3 \perp 1 \mid 2,4$ and $2 \perp 4$).

Here we present a result of Shachter [5] required in Proposition 2.2.

THEOREM 2.1: *Let $G_{\text{old}} = (V,E)$ be a DAIG and let $i,j \in V$ where $i \in \text{pa}(j)^{\text{old}}$. Let G_{new} be a DAIG that is obtained from G_{old} by reorienting the edge from i to j and setting:*

$$\text{pa}(i)^{\text{new}} = H \cup \{j\} \quad \text{and} \quad \text{pa}(j)^{\text{new}} = H,$$

where $H = \text{pa}(i)^{\text{old}} \cup \text{pa}(j)^{\text{old}} \setminus \{i\}$. Then $G_{\text{old}} \sim G_{\text{new}}$.

PROOF: This follows from the arc reversal theorem of Shachter [5]. ■

PROPOSITION 2.2: *Let $G = (V,E)$ and $G_J = (V,E_J)$ where $G^{\sim} = G_J^{\sim}$ and G_J is derived from G by reorienting J edges, $J \geq 1$. If $\mathcal{M}(G) = \mathcal{M}(G_J)$, then $G \sim G_J$.*

PROOF: First consider the case where $J = 1$. Suppose that $i \in \text{pa}(j)^G$ and $j \in \text{pa}(i)^{G_1}$ for some (unique) $i,j \in V$. If $\text{pa}(j)^G \cup \text{pa}(i)^G \setminus \{i\} = \varnothing$, then from Theorem 2.1 we have $G \sim G_1$.

If there is some $k \in \text{pa}(j)^G$, $k \neq i$, then it follows that $k \in \text{pa}(i)^G$ since $\mathcal{M}(G) = \mathcal{M}(G_1)$ and cycles are forbidden. Similarly, for any $l \in \text{pa}(i)^G$ we have $l \in \text{pa}(j)^G$. Therefore, no new new links are required for the arc reversal procedure of Theorem 2.1 and $G \sim G_1$.

It remains to show that G_J can be derived from G by reorienting one edge at a time. Suppose that a set $F \subset E$ of $F \leq J$ edges remains to be reoriented and that the reversal of any one of these edges would result in a graph G_{F-1} with $\mathcal{M}(G_{F-1}) \neq (\mathcal{M}(G) = \mathcal{M}(G_J))$. Then for every edge $(i,j) \in F$, reversal of (i,j) will add (remove) an element to (from) $\mathcal{M}(G)$, which will subsequently be removed (added) with later reversals. For each such edge $(i,j) \in F$ there exists a $k \in \text{pa}(i)/\text{pa}(j)$ with $(k,i) \in F$. But for this to be true *for all* $(i,j) \in F$, the edges of F must form one or more directed cycles, which is forbidden. ■

Acknowledgments

This paper has benefited from discussions with Russell Almond, Danika Lew, Adrian Raftery, and Jeremy York.

References

1. Chyu, C.C. (1991). Decomposable probabilistic influence diagrams. *Probability in the Engineering and Informational Sciences* 5: 229–243.
2. Cooper, G.F. & Herskovits, E. (1991). A Bayesian method for constructing Bayesian belief networks from databases. In B.D. D'Ambrosio, P. Smets, & P.P. Bonissone (eds.), *Proceedings of the Seventh Conference on Uncertainty in Artificial Intelligence*. San Mateo, CA: Morgan Kaufman, pp. 86–94.
3. Lauritzen, S.L., Dawid, A.P., Larsen, B.N., & Leimer, H.-G. (1990). Independence properties of directed Markov fields. *Networks* 20: 491–505.
4. Madigan, D. & Raftery, A.E. (1991). Model selection and accounting for model uncertainty in graphical models using Occam's window. Technical Report 213, Department of Statistics, University of Washington, Seattle.
5. Shachter, R.D. (1990). An ordered examination of influence diagrams. *Networks* 20: 535–563.
6. Whittaker, J. (1990). *Graphical models in applied multivariate statistics.* Chichester: Wiley.

Addendum

I have recently been informed that similar results in the context of chain graphs have been presented in Frydenberg, Morten (1990). The chain graph Markov property. *Scandinavian Journal of Statistics* 17: 333–353.