TECHNICAL REPORT
R-114

CHAPTER 4

The Logic of Influence Diagrams

Judea Pearl, Dan Geiger and Thomas Verma, *University of California at Los Angeles, USA*

ABSTRACT

This chapter explores the role of directed acyclic graphs (DAGs) as a representation of conditional independence relationships. We show that DAGs offer polynomially sound and complete inference mechanisms for inferring conditional independence relationships from a given causal set of such relationships. As a consequence, *d-separation*, a graphical criterion for identifying independencies in a DAG, is shown to be both correct and optimal. This criterion can be used, in lieu of arithmetic manipulations, to test the legitimacy of graphical transformations on influence diagrams, to identify relevant sources of information in the diagram and, thus, to guide the control of inferencing and information-gathering strategies.

4.1 INTRODUCTION AND SUMMARY OF RESULTS

Networks employing directed acyclic graphs (DAGs) have a long and rich tradition, starting with the geneticist Sewal Wright (1921). He developed a method called *path analysis* (Wright, 1934) which later became an established representation of causal models in economics (Wold, 1964), sociology (Blalock, 1971; Kenny, 1979) and psychology (Duncan, 1975). Good (1961) used DAGs to represent causal hierarchies of binary variables with disjunctive causes. *Influence diagrams* represent another application of DAG representation (Howard and Matheson, 1981; Olmsted, 1983; Shachter, 1986) developed for decision analysis, they contain both event nodes and decision nodes. *Recursive models* is the name given to such networks by statisticians seeking meaningful and effective decompositions of contingency tables (Lauritzen, 1982; Wermuth and Lauritzen, 1983; Kiiveri *et al.*, 1984). *Bayesian belief networks* (or *causal networks*) is the name adopted for describing networks that perform evidential

Influence Diagrams, Belief Nets and Decision Analysis
Edited by R. M. Oliver and J. Q. Smith

reasoning (Pearl, 1985, 1986, 1988). This chapter establishes a clear semantics for such networks that might explain their wide usage as models for forecasting, decision analysis and evidential reasoning.

An influence diagram* can be viewed as an economical scheme for representing conditional independence relationships and for deducing new independencies from those used in the construction of the diagram. The nodes in the diagram represent variables in some domain and its topology is specified by a list of qualitative judgements elicited from an expert in this domain. The specification list designates to each variable v a set of parents judged to have direct influence on v, and this amounts to asserting that, given its parents-set, each variable is independent of its other predecessors in some total order of the variables. This stratified list of independencies, henceforth called input list or causal list, implies many additional independence relationships that can be read off the diagram. For example, it is well known that, given its parents, each variable is also independent of all its non-descendants (Howard and Matheson, 1981). Additionally, if S is a set of nodes containing v's parents, v's children and the parents of those children, then v is independent of all other variables in the system, given those in S (Pearl, 1986). These assertions are examples of valid consequences of the input list, i.e. independencies that hold for every probability distribution that satisfies the conditional independencies specified by the input. If one ventures to perform topological transformations on the diagram (e.g. arc reversal or node removal, Shachter, 1986), caution must be exercised to ensure that each transformation does not introduce extraneous, invalid independencies, and/or that the number of valid independencies which become obscured by the transformation is kept at a minimum. Thus, in order to decide which transformations are admissible, one should have simple criteria for deciding which conditional independence statement is valid and which is not. The development of such criteria is the central theme of this chapter.

This chapter deals with the following questions:

1. What are the valid consequences of the input list?
2. What are the valid consequences of the input list that can be read off the diagram?
3. Are the two sets identical?

The answers obtained are as follows:

1. A statement is a valid consequence of the input set if and only if it can be derived from it using the following four axioms. Letting X, Y, and Z stand for three disjoint subsets of variables, and denoting by $I(X, Z, Y)$ the statement: 'the variables in X are conditionally independent of those in Y, given those in Z', the axioms state:

*Our analysis will focus on the so-called knowledge part of influence diagrams, thus excluding decision nodes.

Symmetry

$$I(X, Z, Y) \Rightarrow I(Y, Z, X) \tag{4.1a}$$

Decomposition

$$I(X, Z, Y \cup W) \Rightarrow I(X, Z, Y) \quad \text{and} \quad I(X, Z, W) \tag{4.1b}$$

Weak union

$$I(X, Z, Y \cup W) \Rightarrow I(X, Z \cup W, Y) \tag{4.1c}$$

Contraction

$$I(X, Z \cup Y, W) \quad \text{and} \quad I(X, Z, Y) \Rightarrow I(X, Z, Y \cup W). \tag{4.1d}$$

These axioms form a system called semi-graphoid (Pearl and Paz, 1985; Pearl and Verma, 1987) and were first proposed as heuristic properties of conditional independence by Dawid (1979).

2. Every statement that can be read off the DAG using the *d*-separation criterion represents a valid consequence of the input list (Verma, 1986).

   The *d*-separation criterion is defined as follows (Pearl, 1985): for any three disjoint subsets X, Y, Z of nodes in a DAG D, Z is said to *d*-separate X from Y, denote $I(X, Z, Y)_D$, if there is no path from a node in X to a node in Y along which the following two conditions hold: (a) every node with converging arrows is in Z or has a descendant in Z, and (b) every other node is outside Z. (The definition is elaborated in Section 4.2.)
3. The two sets are identical, namely, a statement is valid if and only if it is graphically validated under *d*-separation in the diagram (Geiger and Pearl, 1988a).

The first result establishes the decidability of verifying whether an arbitrary statement is a valid consequence of the input set, i.e. applying axioms (4.1a)–(4.1d) on a causal input list is guaranteed to generate all its valid consequences and none other. The second result renders the *d*-separation criterion a polynomially sound inference rule, i.e. it runs in polynomial time and certifies only valid statements. The third renders the *d*-separation criterion a complete inference rule. In summary, the diagram constitutes a sound and complete inference mechanism that identifies, in polynomial time, each and every valid consequence in the system. Interestingly, result 2 holds for any semi-graphoid system, not necessarily probabilistic conditional independencies. Thus, influence diagrams can serve as effective inference instruments for a variety of dependence relationships, e.g. partial correlations (Crámer, 1946) and qualitative database dependencies (Fagin, 1977).

The results above are true only for causal input sets, i.e. those that recursively specify the dependence of each variable on its predecessors in some total order, e.g. chronological. The general problem of verifying whether a given conditional independence statement logically follows from an arbitrary set of such statements, may be undecidable. Its decidability would be resolved upon finding a complete set of axioms for conditional independence, i.e. axioms that are powerful enough to derive all valid consequences of an arbitrary input set. Until very recently, all indications were that axioms (4.1a)–(4.1d) are complete, as conjectured in Pearl and Paz (1985), but a new result of Studeny (1988) has refuted that hypothesis. Currently, it appears rather unlikely that there exists a finite set of axioms which is complete for conditional independence, thus, the general decidability problem remains unsettled. The completeness problem is treated in Geiger and Pearl (1988b) and completeness results for specialized subsets of probabilistic dependencies are summarized in Section 4.2.6.

4.2 SOUNDNESS AND COMPLETENESS

4.2.1 The *d*-separation criterion

The definition of *d*-separation is best motivated by regarding DAGs as a representation of causal relationships. Designating a node for every variable and assigning a link between every cause to each of its direct consequences defines a graphical representation of a causal hierarchy. For example, the propositions 'It is raining' (α), 'the pavement is wet' (β) and 'John slipped on the pavement' (γ) are well represented by a three-node chain, from α through β to γ. The chain indicates that either rain or wet pavement could cause slipping, yet wet pavement is designated as the direct cause; rain could cause someone to slip only by wetting the pavement, not if the pavement is covered. Moreover, knowing the condition of the pavement renders 'slipping' and 'raining' independent, and this is represented graphically by a *d*-separation condition, $I(\alpha, \beta, \gamma)_D$, showing node α and γ separated from each other by node β.

Now assume that 'broken pipe' (δ) is considered another direct cause for wet pavement, as in Figure 4.1. An informational dependency may be induced between the two causes of wet pavement: 'rain' and 'broken pipe'. Although they appear connected in Figure 4.1, these propositions are marginally independent and become dependent once we learn that the pavement is wet or that someone broke his leg. An increase in our belief in either cause would

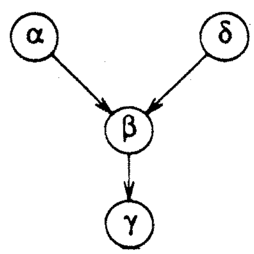

Figure 4.1

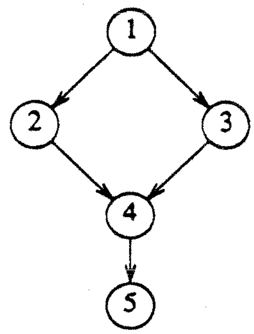

Figure 4.2

decrease our belief in the other as it would 'explain away' the observation. The following definition of d-separation permits us to graphically identify such induced dependencies from the DAG (d connoted 'directional').

Definition (*d-separation*). If X, Y, and Z are three disjoint subsets of nodes in a DAG D, then Z is said to d-separate X from Y, denoted $I(X, Z, Y)_D$, if and only if there is no path* from a node in X to a node in Y along which the following two conditions hold: (1) every node with converging arrows either is or has a descendant in Z, and (2) every other node is outside Z. A path satisfying the conditions above is said to be active otherwise it is said to be blocked (by Z). Whenever a statement $I(X, Z, Y)_D$ holds in a DAG D, the predicate $I(X, Z, Y)$ is said to be graphically verified (or an independency), otherwise it is graphically unverified by D (or a dependency). □

In Figure 4.2, for example, $X = \{2\}$ and $Y = \{3\}$ are d-separated by $Z = \{1\}$; the path $2 \leftarrow 1 \rightarrow 3$ is blocked by $1 \in Z$ while the path $2 \rightarrow 4 \leftarrow 3$ is blocked because 4 and all its descendants are outside Z. Thus $I(2, 1, 3)$ is graphically verified by D. However, X and Y are not d-separated by $Z' = \{1, 5\}$ because the path $2 \rightarrow 4 \leftarrow 3$ is rendered active. Consequently, $I(2, \{1, 5\}, 3)$ is graphically unverified by D; by virtue of 5, a descendant of 4, being in Z. Learning the value of the consequence 5, renders its causes 2 and 3 dependent, like opening a pathway along the converging arrows at 4.

4.2.2 Conditional independence and its graphical representations

Definition. If X, Y and Z are three disjoint subsets of variables of a distribution P, then X and Y are said to be *conditionally independent* given Z, denoted

$$I(X, Z, Y)_P, \quad \text{iff} \quad P(x, y \mid z) = P(x \mid z) P(y \mid z)$$

*By path we mean a sequence of edges in the underlying undirected graph, i.e. ignoring the directionality of the arcs.

for all possible assignments $X = x$, $Y = y$ and $Z = z$ for which $P(Z = z) > 0$. $I(X, Z, Y)_P$ is called a (*conditional independence*) *statement*. A conditional independence statement σ logically *follows* from a set Σ of such statements if σ holds in every distribution that obeys Σ. In such case we also say that σ is a *valid consequence* of Σ. □

It is easy to verify that the set of statements $I(X, Z, Y)_P$ generated by any probability distribution P must obey axioms (4.1a)–(4.1d). More generally, any system of statements $I(X, Z, Y)$ that obey axioms (4.1a)–(4.1d) is called a semi-graphoid (Pearl and Paz, 1985; Pearl and Verma, 1987). Intuitively, the essence of these axioms lies in equations (4.1c) and (4.1d), asserting that when we learn an irrelevant fact, all relevance relationships among other variables in the system should remain unaltered; any information that was relevant remains relevant and that which was irrelevant remains irrelevant. These axioms, common to almost every formalization of informational dependencies, are very similar to those assembled by Dawid (1979) for probabilistic conditional independence and those proposed by Smith (1987b) for generalized conditional independence. The difference is only that Dawid and Smith lumped equations (4.1b)–(4.1d) into one, and added an axiom to handle some cases of overlapping sets X, Y, Z. We shall henceforth call axioms (4.1a)–(4.1d) *Dawid's axioms*, or semi-graphoid axioms, interchangeably. Interestingly, both undirected graphs and DAGs conform to the semi-graphoid axioms (hence the name) if we associate the statement $I(X, Z, Y)$ with the graphical condition 'every path from X to Y is blocked by the set of nodes corresponding to Z'. (In DAGs, blocking is defined by d-separation.) It is this commonality, we speculate, that renders graphs such a popular scheme for representing informational dependencies.

Ideally, to employ a DAG D as a graphical representation for dependencies of some distribution P we would like to require that for every three disjoint sets of variables in P (and nodes in D) the following equivalence should hold:

$$I(X, Z, Y)_D \quad \text{iff} \quad I(X, Z, Y)_P. \tag{4.2}$$

This would provide a clear graphical representation of all variables that are conditionally independent. When equation (4.2) holds, D is said to be a perfect map of P. Unfortunately, this requirement is often too strong because there are many distributions that have no perfect map in DAGs. The spectrum of probabilistic dependencies is in fact so rich that it cannot be cast into any representation scheme that uses polynomial amount of storage (Verma, 1987). Geiger (1987) provides a graphical representation based on a collection of graphs (multi-DAGs) that is powerful enough to perfectly represent an arbitrary distribution; however, as shown by Verma, it requires, on the average, an exponential number of DAGs.

Being unable to provide perfect maps at a reasonable cost, we compromise the requirement that the graphs represent each and every dependency of P, and allow some independencies to escape representation. We will require through that the set of undisplayed independencies be minimal.

Definition. A DAG D is said to be an I-map of P if for every three disjoint subsets X, Y and Z of variables the following holds:

$$I(X, Z, Y)_D \Rightarrow I(X, Z, Y)_P. \tag{4.3}$$

D is said to be a minimal I-map of P if no edge can be deleted from D without destroying its I-mapness. □

The task of finding a DAG which is a minimal I-map of a given distribution P can be accomplished by the traditional chain-rule decomposition of probability distributions. The procedure consists of the following steps: assign a total ordering d to the variables of P. For each variable i of P, identify a minimal set of predecessors S_i that renders i independent of all its other predecessors (in the ordering of the first step). Assign a direct link from every variable in S_i to i. The analysis of Verma (1986) and Pearl and Verma (1987) ensures that the resulting DAG is an I-map of P, and is minimal in the sense that no edge can be deleted without destroying its I-mapness. The input list L for this construction consists of n conditional independence statements, one for each variable, all of the form $I(i, S_i, U_{(i)} - S_i)$, where $U_{(i)}$ is the set of predecessors of i and S_i is a subset of $U_{(i)}$ that renders i conditionally independent of all its other predecessors. This set of conditional independence statements is called a causal input list and is said to define the DAG D. The term 'causal' input list stems from the following analogy: suppose we order the variables chronologically, such that a cause always precedes its effect. Then, from all potential causes of an effect i, a causal input list selects a minimal subset that is sufficient to explain i, thus rendering all other preceding events superfluous. This selected subset of variables are considered direct causes of i and therefore each is connected to it by a direct link.

4.2.3 The main results

Clearly, the constructed DAG represents more independencies than those listed in the input, namely, all those that are graphically verified by the d-separation criterion. The results reported in the previous subsection guarantee that all graphically verified statements are indeed valid in P, i.e. the DAG is an I-map of P. It turns out that the constructed DAG has another useful property: it graphically verifies every conditional independence statement that logically follows from L (i.e. holds in every distribution that obeys L). Hence, we cannot hope to improve the d-separation criterion to display more independencies, because all valid consequences of L (which defines D) are already captured by d-separation.

The three theorems below formalize these results. Proofs can be found in the references cited.

Theorem 4.1 (soundness) (Verma, 1986). Let D be a DAG defined by a causal input list L, of some dependency model obeying axioms (4.1a)–(4.1d) (e.g. probabilistic dependence). Then, every graphically verified statement is a valid consequence of L. □

Theorem 4.2 (closure) (Verma, 1986). Let D be a DAG defined by a causal input list L. Then, the set of graphically verified statements is exactly the closure of L under axioms (4.1a)–(4.1d). □

Theorem 4.3 (completeness) (Geiger and Pearl, 1988a). Let D be a DAG defined by a causal input list L. Then, every valid consequence of L is graphically verified by D (equivalently, every graphically unverified statement in D is not a valid consequence of L). □

Theorem 4.1 guarantees that the DAG displays only valid statements. Theorem 4.2 guarantees that the DAG displays all statements that are derivable from L via axioms (4.1). Theorem 4.3 assures that the DAG displays all statements that logically follow from L, i.e. the axioms in (4.1) are complete, capable of deriving all valid consequences of a causal input list. Moreover, since a statement in a DAG can be verified in polynomial time, Theorems 4.1–4.3 provide a complete polynomial inference mechanism for deriving all independency statements that are implied by a causal input list.

The first two theorems are more general than the third in the sense that they hold for every dependence relationship that obeys axioms (4.1a)–(4.1d), not necessarily those based on probabilistic conditional independence (proofs can be found in Verma, 1986 and Verma and Pearl, 1988). Among these dependence relationships are partial correlations (Crámer, 1946; Pearl and Paz, 1985) and qualitative dependencies (Fagin, 1977; Shafer *et al.*, 1987) which can readily be shown to obey axioms (4.1). The completeness of d-separation (Theorem 4.3) relative partial correlations has been established in Geiger and Pearl (1988) while completeness relative to qualitative dependencies has not yet been examined.

4.2.4 Deterministic nodes and *D*-separation

Theorem 1–3 assume that L contains only statements of the form $I(i, S_i, U_{(i)} - S_i)$. Occasionally, however, we are in possession of stronger forms of independence relationships, in which case additional statements should be read of the DAG. A common example is the case of a deterministic variable (Shachter, 1988), i.e. a variable that is functionally dependent on its corresponding parents in the DAG. The existence of each such variable i could be encoded in L by a statement of global independence $I(i, S_i, U - S_i - i)$ asserting that conditioned on S_i, i is independent of all other variables, not merely of its predecessors. The independencies that are implied by the modified input list can be read from the DAG using an enhanced version of d-separation, named *D-separation.*

Definition. If X, Y, and Z are three disjoint subsets of nodes in a DAG D, then Z is said to D-separate X from Y, if and only if there is no path from a node in X to a node in Y along which the following three conditions hold: (1) every node with converging arrows either is or has a descendant in Z; (2) every other node is outside Z; and (3) no node is functionally determined by Z. □

The new criterion certifies all independencies that are revealed by d-separation plus additional ones due to the enhancement of the input list. For example, if the arc $2 \rightarrow 4$ in Figure 4.2 were deterministic, then 2 D-separates 5 from 3 and, indeed, 5 and 3 must be conditionally independent given 2 because 2 determines the value of 4. In fact, the link $3 \rightarrow 4$ is redundant because $I(4, 2, 135)$ implies $I(4, 25, 1)$. It is for this reason that deterministic nodes can be presumed to receive only deterministic incoming arrows. The soundness and completeness of D-separation is stated in Theorem 4.4.

Theorem 4.4 (Geiger and Verma, 1988). Let D be a DAG defined by a causal input list L, possibly containing functional dependencies. Then a statement is a valid consequence of L if and only if it is graphically verified in D by the D-separation criterion. □

These graphical criteria provide easy means of recognizing conditional independence in influence diagrams as well as identifying the set of parameters needed for any given computation (Section 4.4.3). We now show how these theorems can be employed as an inference mechanism. Assume an expert has identified the following conditional independencies between variables denoted 1–5:

$$L = \{I(2, 1, \varnothing), I(3, 1, 2), I(4, 23, 1), I(5, 4, 123)\}$$

(the first statement in L is trivial). We address two questions. First, what is the set of all valid consequences of L? Second, in particular, is $I(3, 124, 5)$ a valid consequence of L? For general input lists the answer for such questions may be undecidable but, since L is a causal list, it defines a DAG that graphically verifies each and every valid consequences of L. The DAG D is the one shown in Figure 4.2, which constitutes a compact representation of all valid consequences of L. To answer the second question, we simply observe that $I(3, 124, 5)$ is graphically verified in D. A graph-based algorithm for another subclass of statements, called fixed context statements, is given in Geiger and Pearl (1988b). In that paper, results analogous to Theorem 4.1–4.3 are proven for Markov fields; a representation scheme based on undirected graphs (Isham, 1981, Lauritzen, 1982).

4.2.5 Strong completeness

Theorem 4.3 can be restated to assert that for every DAG D and any statement σ graphically unverified by D there exists a probability distribution P_σ that embodies D's causal input set L and the dependency σ. By Theorem 4.2, P_σ must embody all graphically verified statements as well because they are all derivable from L by Dawid's axioms. Thus, Theorems 4.2 and 4.3 guarantee the existence of a distribution P_σ that satisfies all graphically verified statements and a single, arbitrarily chosen, graphically unverified statement (i.e. a dependency). The question answered by Theorem 4.5 is the existence of a distribution P that embodies all independencies of D and *all* its dependencies, not merely a single dependency. Such a set of axioms is said to be strongly complete (Beeri *et al.*, 1977).

Theorem 4.5 (strong completeness) (Geiger and Pearl, 1988a). For every DAG D there exists a distribution P such that for every three disjoint sets of variables X, Y and Z the following holds:

$$I(X,Z,Y)_D \quad \text{iff} \quad I(X,Z,Y)_P. \qquad \square$$

Theorem 4.5 legitimizes the use of DAGs as a representation scheme for probabilistic dependencies; a model builder who uses the language of DAGs to express dependencies is guarded from inconsistencies.

4.2.6 Other completeness results

We conclude this section with a summary of completeness results for specialized subsets of independence statements. Proofs can be found in Geiger and Pearl (1988b) and Geiger *et al.* (1988). The first result establishes an axiomatic characterization of marginal statements, i.e. statements of the form $I(X, Z_0, Y)$ where the middle argument Z_0 is fixed. The second result provides a complete axiomatic characterization of fixed-context statements, i.e. statements of the form $I(X, Z, Y)$ where $X \cup Z \cup Y$ sum to a fixed set of variables U.

Theorem 4.6 (completeness for marginal independence). Let Σ be a set of marginal statements closed under the following axioms:

Symmetry	$I(X, Z_0, Y) \rightarrow I(Y, Z_0, X)$
Decomposition	$I(X, Z_0, YW) \rightarrow I(X, Z_0, Y)$
Mixing	$I(X, Z_0, Y)$ and $I(XY, Z_0, W) \rightarrow I(X, Z_0, YW)$.

There exists a probability model P that obeys all statements in Σ and none other. $\square$

Theorem 4.7 (completeness for fixed-context). Let Σ be a set of fixed-context statements closed under the axioms:

Symmetry	$I(Y, Z, Y) \rightarrow i(Y, Z, X)$
Weak union	$I(X, Z, YW) \rightarrow I(X, ZY, W)$
Weak contraction	$I(XY, Z, W)$ and $I(X, ZW, Y) \rightarrow I(X, Z, YW)$.

There exists a probability model P that obeys all statements in Σ and none other. Moreover, if Σ is closed also under

$$\text{Intersection} \qquad I(X,ZW,Y) \text{ and } I(X,ZY,W) \rightarrow I(X,Z,YW)$$

then P can be selected to be strictly positive. □

The membership question, whether a given marginal statement follows from an arbitrary set of such statements, can be answered in linear time (Geiger *et al.*, 1988). For fixed-context statements, the membership question can be answered in quadratic time (Beeri *et al.*, 1987).

4.3 COROLLARIES

Theorem 4.1 leads to four corollaries which are the key to the construction of influence diagrams from a given distribution P. First we define influence diagrams in terms of the independencies they portray, then we justify their method of construction.

Definition. Given a probability distribution P on a set U of variables, $U = \{X_1, X_2, \ldots, X_n\}$, a DAG $D = (U, E)$ is called an influence diagram* of P if and only if D is a minimal I-map of P. □

Corollary 4.1. Given a probability distribution $P(x_1, x_2, \ldots, x_n)$ and any ordering d of the variables. The DAG created by designating as parents of X_i any minimal set S_i of predecessors satisfying

$$I(X_i, S_i, U_{(i)} - S_i), \quad U_{(i)} = \{X_1, X_2, \ldots, X_{i-1}\} \tag{4.4}$$

is an influence diagram of P. Conversely, every influence diagram of P can be constructed by identifying the parent sets S_i defined in (4.4) along some ordering d. If P is strictly positive, then all the parent sets are unique (Pearl and Paz, 1985) and the influence diagram is unique as well (given d). □

Although the structure of the diagram depends strongly on the node ordering used in the construction, each diagram is nevertheless an I-map of the underlying distribution P. This means that all conditional independencies portrayed in the diagram (via d-separation) are valid in P and hence, are order independent. An immediate corollary of this observation yields an order-independent test for minimal I-mapness.

Corollary 4.2. Given a DAG D and a probability distribution P, a necessary and sufficient condition for D to be a minimal I-map (hence an influence diagram) of P is that each variable X_i be conditionally independent of all its non-descendants, given its parents S_i, and no proper subset of S_i satisfies this condition.

The necessary part follows from the fact that every parent-set S_i d-separates X_i from all its non-descendants. The sufficient part holds because X_i's independence of all its non-descendants entails X_i's independence of its predecessors in a particular ordering d (as required by Corollary 4.1). □

Corollary 4.3. If an influence diagram D is constructed from P (by the method of Corollary 4.1) in some ordering d, then any ordering d' consistent with the direction of arrows in D would give rise to an identical diagram.

The validity of Corollary 4.3 follows from that of Corollary 4.2, which ensures that the set S_i will satisfy equation (4.4) in any new ordering, as long as the new set of X_i's predecessors does not contain any of X_i's old descendants. Thus, once the network is constructed, the original order can be forgotten; only the partial order displayed in the diagram matters.

Another interesting corollary of Theorem 4.1 is a generalization of the celebrated *Markov-chain* property which is used extensively in the probabilistic analysis of random walks, time-series data and other stochastic processes (Feller, 1968; Meditch, 1969). The property states that if in a sequence of n trials $X_1, X_2, \ldots, X_n$ the outcome of any trial $X_k, k = 2, 3 \ldots n$ depends only on the outcome of its directly preceding trial X_{k-1} then, given the entire past and future trials $X_1, X_2, \ldots, X_{k-1}, X_{k+1}, \ldots, X_n$, the outcome of X_k depends only on its two nearest neighbors, X_{k-1} and X_{k+1}. Formally,

$$I(X_k, X_{k-1}, X_{k-2} \ldots X_1), \quad 2 \leqslant k \leqslant n$$
$$\Rightarrow I(X_k, X_{k-1} X_{k+1}, X_n \ldots X_{k+2} X_{k-2} \ldots X_1), \quad 2 \leqslant k \leqslant n-1$$

(The converse holds only for full graphoids, e.g. strictly positive distributions.) Theorem 1 generalizes the Markov-chain property to dependencies other than probabilistic and to structures other than chains. The d-separation criterion uniquely determines a *Markov blanket* for any given node X_i in an influence diagram, namely, a set $\mathbf{BL}(X_i)$ of variables that renders X_i independent of all variables not in $\mathbf{BL}(X_i)$, i.e. $I(X_i, \mathbf{BL}(X_i), U - \mathbf{BL}(X_i) - X_i)$.

Corollary 4.4. In any influence diagram, the union of the following three types of neighbors is sufficient for forming a Markov blanket of a node X_i: the direct parents of X_i, the direct successors of X_i and all direct parents of the latter. □

Thus, if the diagram consists of a single path (i.e. a Markov chain), the Markov blanket of any nonterminal node consists of its two immediate neighbors, as expected. In trees, the Markov blanket consists of the (unique) parent and the immediate successors. In Figure 4.2, however, the Markov blanket of node 3 is $\{1, 4, 3\}$. Note that in general, these Markov blankets are not minimal; alternative ordering might display X_i with a smaller set of neighbors.

The necessary part of Corollary 4.2 was stated without proof in Howard and Matheson (1981) and was later used in the derivations of Olmsted (1983) and Shachter (1986). Corollaries 4.2 and 4.3 are proven in Smith (1987) using the axioms of equation (4.1). Since Theorem 4.1 establishes d-separation a sound procedure relative to Dawid's axioms, the validity of such corollaries can now be verified by purely graphical means.

4.4 APPLICATIONS

4.4.1 Entailment and equivalence

One problem concerning influence diagrams that has received much attention is that of *entailment*, i.e. establishing a criterion for testing whether one influence diagram entails all the conditional independencies portrayed by another (Smith, 1987). We shall now establish such a criterion using the notions of d-separation.

Definition. Given two diagrams D_1 and D_2, we say that D_1 entails D_2 (equivalently, D_2 is implied by D_1) if every independency verified in D_2 is implied by D_1, i.e. if for all disjoint subsets of nodes X, Y and Z, we have

$$I(X,Z,Y)_{D_2} \Rightarrow I(X,Z,Y)_{D_1}. \tag{4.5}$$

Remark. Although the two diagrams may contain different sets of nodes, we require only that equation (4.5) holds for all triplets (X,Y,Z) that appear in both diagrams.

Based on Theorem 4.1 and its corollaries, condition (4.5) is equivalent to stating that every independency displayed in D_2 must hold in any probability distribution P_1 that underlies D_1, i.e. D_2 is an I-map of any distribution P of which D_1 is an I-map. Moreover, Theorem 4.1 also yields a simple procedure for verifying (4.5) without testing all triplets (X,Y,Z) in D_2. According to Corollary 4.2, it is sufficient to verify equation (4.5) for just one causal list of D_2.

Corollary 4.5. A necessary and sufficient condition for a diagram D_1 to entail D_2 is that, for every node X_i having parents S_i^2 in D_2 we have: $I(X_i, S_i^2, U_{(i)}^2 - X_i - S_i^2)_{D_1}$, where $U_{(i)}^2 = \{X_1, X_2, \ldots, X_{i-1}\}$ is a set of predecessors of X_i in some ordering that is consistent with the arrows of D_2. □

As an example, consider the four DAGs shown in Figure 4.3. Here D_1 does not entail D_2 because $I(5,1,24)$ is not verified in D_1. Similarly D_1 does not entail D_3; $I(5,14,2)$ is not verified in D_1. However, D_1 entails D_4 because the following three statements:

$$I(2,1,\varnothing), \qquad I(5,14,2), \qquad I(4,125,\varnothing)$$

are verified in D_1. These statements correspond to the causal list of D_4 in the ordering $(1,2,5,4)$. In general, the question of entailment can be decided in $O(n^2)$ time, because each d-separation statement can be verified in linear time (Geiger and Verma, 1988).

Definition. Two DAGs, D_1 and D_2, are said to be equivalent, written $D_1 \equiv D_2$, if each entails the other, namely, if they portray the same set of conditional independence relationships over the set of nodes common to both.

The equivalence of two DAGs can be established either by entailment both ways or, more transparently, by invoking the notions of adjacency and conditional adjacency. □

Definition. Two nodes a and b in an influence diagram are adjacent, written $a - b$, if there is a direct are connecting them. They are conditionally adjacent

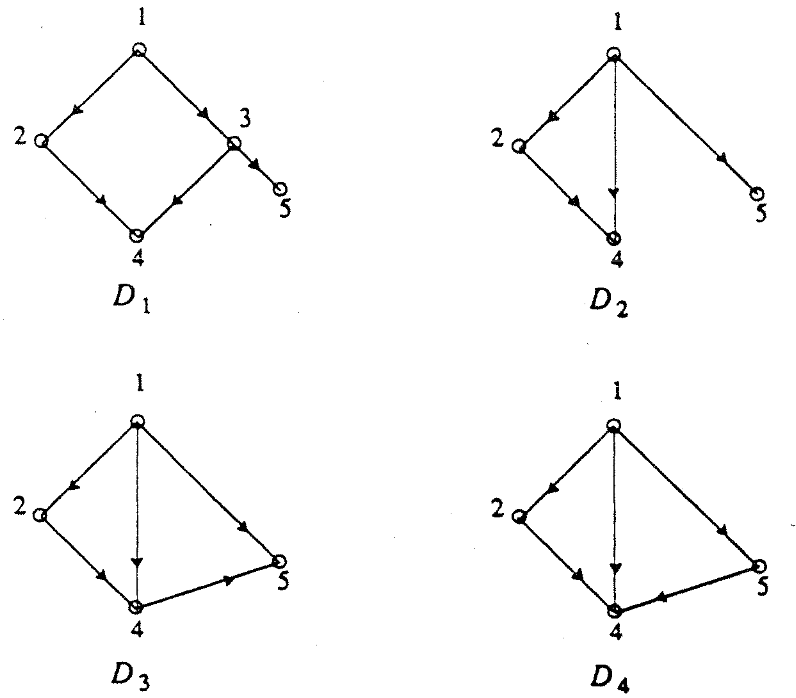

Figure 4.3

given c, written $a - b|c$, if a and b are adjacent or, alternatively, if a and b are the common parents of c or some ancestor of c. □

Adjacency is equivalent to non-d-separability, namely, $a - b$ if and only if there is no set S (not containing a or b) which d-separates a from b. Using these notions, the criterion for entailment can be stated concisely:

Corollary 4.6 (equivalence). For any two influence diagrams, D and E, over the set of nodes $N, D \equiv E$ if and only if both of the following hold:

1. $\forall_{a,b \in N} \; a \underset{D}{-} b \Leftrightarrow a \underset{E}{-} b$

2. $\forall_{a,b,c \in N} \; a \underset{D}{-} b|c \Leftrightarrow a \underset{E}{-} b|c.$

Thus, $D \equiv E$ if and only if D and E share the same set of undirected edges and the same set of converging arrows emanating form non-adjacent sources; non-converging arrows may have arbitrary orientations (as long as they do not form directed cycles). Typical examples of equivalent networks are those created by different orientations of a tree or, more generally, different orientations of a chordal graph. Chordal graphs can be oriented in such a way that all converging arrows emanate from adjacent sources (Pearl, 1988), thus satisfying the conditions of Corollary 4.6.

4.4.2 Arc reversal and node removal

A transformation on an influence diagram is said to be sound if it does not introduce extraneous independencies, namely, if the resulting diagram is entailed by the original one. The importance of guarding against the introduction of new independencies lies in the requirement that the numerical parameters which label the arcs of the transformed diagram be sufficient to reconstruct the distribution that underlies the original diagram. It can be shown that as long as the transformed diagram does not display new independencies, the original distribution can be reconstructed by taking the product of the conditional probabilities of all child–parents families in the transformed diagram (e.g. see Pearl, 1988). Thus, no information is lost in the transformation.

Corollary 4.5 defines the conditions for sound transformations of influence diagrams. For example, the conditions forbid the removal of any arc from the diagram; only reorientation and addition of arcs are permitted. Additionally, if a reorientation destroys a pair of non-adjacent common parents, then the pair must be joined by a link; otherwise a conditional independence between such parents (given the child and the grandparents) would be created by the transformation. Finally, if a reorientation creats a new pair of common parents, then the pair must be joined by a link.

From this corollary, most results concerning manipulations of influence diagrams are immediate. Consider, for example, the rule for *arc reversal* (Olmsted, 1983): if an arc from node X to node Y is reversed, X and Y should inherit each other's parents. The reason for this rule is shown in Figure 4.4, where S_X, S_Y and S_{XY} represent sets of parent nodes of X, Y and X and Y respectively. The transformed diagram without parent inheritance (Figure 4.4b) displays two independencies which are not verified by the original diagram (Figure 4.4a), namely, $I(X, YS_{XY}S_X, S_Y)$ and $I(Y, S_{XY}S_Y, S_X)$. To destroy these two independencies it is necessary (and sufficient) to add the two arcs $S_Y \rightarrow X$

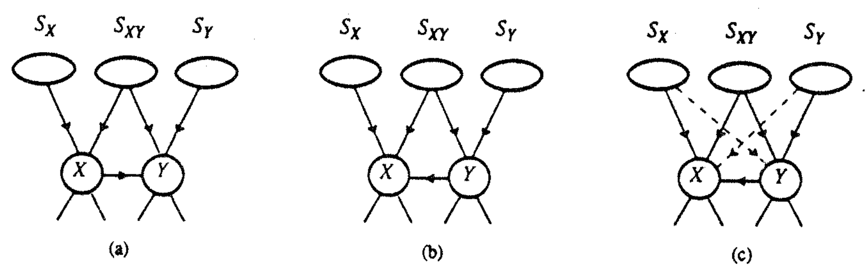

Figure 4.4

and $S_X \rightarrow Y$, as shown in Figure 4.4(c). Figure 4.4(c) satisfies the condition of Corollary 4.5, because the final parent set $YS_XS_{XY}S_Y$ of X is shown in Figure 4.4(a) to d-separate X from all its non-descendants (as defined by Figure 4.4(c)). A similar d-separation condition is verified for the final parent set of Y. Note, that the added arc $S_X \rightarrow Y$ can be reversed and still satisfy the condition for X but, then, additional arcs might be required between Y and the ancestors of S_X.

The parent-inheritance rule for arc reversal was alluded to in Howard and Matheson (1981) and later was proven for probability models by Olmsted (1983), assuming (without proof) that Corollary 4.2 holds. This arc-reversal rule was also used by Shachter (1986) as the basic mechanism for evaluating influence diagrams, and was proven by Smith (1987) using axioms (4.1a)–(4.1d). The soundness of the d-separation criterion (Theorem 4.2) offers a graphical proof for the soundness of the arc-reversal rule, once we verify that every parents–child statement verified in the transformed diagram is graphically verified in the original diagram. Moreover, the completeness of the d-separation criterion implies that the parent-inheritance rule for arc reversal is also minimal, i.e. every arc added in this process in necessary, when only one arc is to be reversed; it may not be minimal when several arcs are reversed.

The rule for *node removal* also follows from Corollary 4.5. Traditionally, node removal was executed using a two-step process: to remove an arbitrary node, simply reorient any links directed out of it (using arc reversal), then remove the node and all its links. If several arcs are to be reversed, then sequentially applying the arc-reversal method to the individual arcs may result in the addition of unnecessary links, as is shown in Figure 4.3; eliminating node 3 by first reversing arc $3 \rightarrow 4$ then reversing $3 \rightarrow 5$ requires the addition of an arc between 2 and 5. Yet, diagram D_4 shows that the arc $2 \rightarrow 5$ is superfluous.

The rule for direct node removal is stated in the following corollary.

Corollary 4.7 (node removal). A node X can be eliminated from the diagram if every child of X inherits X's parents, every pair of X's children is connected by an arc and every child that receives an arrow from another child inherits all parents of latter. These added arcs are both necessary and sufficient. Accordingly, a leaf node can be removed without adding any arcs. □

4.4.3 Information required

When an influence diagram is constructed from an expert or from other representations of knowledge, it is important to determine the information needed to answer a given query of the form 'find $P(x_J|x_K)$', where J and K are two arbitrary sets of nodes in the diagram (Shachter, 1985, 1988). Assuming that each node X_i stores the conditional distribution $P(x_i|S_i)$, where S_i are the parents of X_i, the task is to identify the set of nodes that must be consulted in the process of computing $P(x_J|x_K)$ or, alternatively, the set of nodes that can be assigned arbitrary conditional distributions without affecting the quantity $P(x_J|x_K)$.

The required set can be identified by the d-separation criterion if we represent the parameters of the distribution $P(x_i|S_i)$ as a dummy parent of node X_i. From Theorem 4.1, all dummy nodes that are d-separated from J by K represent variables that are conditional independent of J given K and so, the information stored in these nodes can be ignored. Thus, the information required to compute $P(x_J|x_K)$ resides in the set of dummy nodes which are not d-separated from J given K.

Shachter (1985, 1988) has devised an algorithm for finding a set of nodes M guaranteed to contain sufficient information for computing $P(x_J|x_K)$. The outcome of Shachter's algorithm can now be stated declaratively: M contains every ancestor of $J \cup K$ that is not d-separated from J given K and none other. The completeness of d-separation further implies that M is minimal; no node in M can be excluded on purely topological grounds (i.e. without considering the numerical values of the probabilities involved). If the diagram contains deterministic nodes, then the D-separation criterion should be used instead of d-separation.

Identifying the set of nodes that are d-separated from J by K is also important for limiting the scope of the subnetwork that should undergo restructuring (e.g. clustering or conditioning, Pearl, 1988) as well as for controlling evidence-gathering strategies. Knowing the precise values of variables represented by such nodes will have no effect on the outcome $P(x_J|x_K)$ and, therefore, these variables need not be tested or examined. A linear time algorithm for identifying the set of relevant nodes is reported in Geiger and Verma (1988).

ACKNOWLEDGEMENTS

We thank N. Dalkey, R. Fagin, S. Lauritzen and A. Paz for many valuable discussions. This work was partially supported by the National Science Foundation Grant No. IRI-8610155. 'Graphoids: a computer representation for dependencies and relevance in automated reasoning (computer information science)'.

4.5 DISCUSSION

4.5.1 Discussion by Herman Rubin

The chapter seems to be a well-written mathematical paper, providing methods for deducing whether independence relations and conditional independence relations hold from the graph structure alone of influence diagrams. I have not had the time to look over the details, but this looks like a contribution to the mathematical treatment of the objects involved.

On the other hand, is it anything more? The author seems obsessed in reducing everything to those independence relations. He even wishes to define causality in those terms. I find this difficult to understand. In thinking over his remarks to my discussion at the meeting, I was pushed to derive a definition of causality. It is not definable from the joint distributions.

> A is a cause for B if an extranatural agent interfering only with A can affect B.

This is not symmetric in A and B; we have what the philosophers and physicists call time's arrow to contend with. I maintain that the sun shining is a cause of the grass growing but not vice versa. I will continue to maintain this asymmetry until the barely conceivable possibility of infants returning to the womb to render their mothers pregnant occurs with reasonable frequency.

Furthermore, independence relationships are not all that important. Sewall Wright essentially expressed his ideas in a manner very similar to influence diagrams. However, he was interested not so much in which variables affected which, but in how. He had no hesitation in introducing unobservable variables, and while his statistical knowledge was necessarily limited, he made good use of it. Of course he used British statistics; what else was there?

The mathematical economists, who were attempting to devise a method for extraeconomic agents, in particular governments, to take actions which would have 'beneficial' effects on the economy, were essentially using influence diagrams and decision trees, although these formal concepts had not even been invented. I do not think the terminology would have been particularly helpful to them. In any case, the problem was not that of which variables could be ignored, but of the quantitative nature of the relations involved.

4.5.2 Reply

The thrust of Dr Rubin's comment is that, while every causal relationship implies a certain pattern of independencies, the converse is not necessarily true; one cannot infer causal relationships from dependence information or even from a complete joint distribution function.

There is no dispute over this observation. However, this does not mean that we should stop exploring the relations between probability and causality. On the contrary, the observation gives rise to two interesting questions:

1. What, precisely, is the pattern of independencies that accompanies causal assertions?
2. If the bulk of human knowledge is derived from empirical observations, and if empirical observations are summarized by statistical averages, how is it that people manage to cast this knowledge in standard causal schemata whose structures are uniformly accepted and rarely disputed?

In our chapter we have attempted to address question 1. We have treated causal assertions as a shorthand notation for certain patterns of conditional independencies, and we have provided graphical and logical machinery for extracting the sum total of other independencies that follow from such assertions.

One can argue (as Jim Matheson did in one cocktail party) that our results bear no relation whatsoever to causation, since the theorems and corollaries retain their validity for any stratified list of independencies, not necessarily those ordered chronologically*. Our answer is as follows. Of course, if in some peculiar domain we find experts conversing with weirdly stratified dependencies that are ordered contrary to the flow of causation—so be it. Our contention is only that whenever people do choose to communicate in causal sentences, the independencies portrayed in our so-called 'causal lists' or 'input lists' offer a reasonable probabilistic interpretation of such sentences, and that the graphs that ensue offer an efficient mechanism for representing the ramifications of these sentences.

The mathematical properties of the graphical representation are verifiable analytically, while the reasonableness of our interpretation is an empirical claim about the psychology of causation, and should be verified by behavioral experiments. The claim is that, if we apply our proposed probabilistic interpretation to sentences about causation, and draw their probabilistic consequences, then these consequences will never clash with the intuition or intent of the sentence provider.

Let us now address question 2, regarding the empirical basis of causal knowledge. Temporal information undoubtedly provides important clues for causal directionality and causal organization, but, are these clues essential or merely an expedient convention? In the social sciences, for example, we often seek causal models for events that cannot be temporally ordered. We say, for example, that the attitude of a person is a cause for a certain behavior, though it is impossible to determine which comes first. It is important, therefore, ti give a nontemporal probabilistic interpretation of causation, based on the notion of dependence, viewing the temporal component of causation merely as a convenient indexing standard chosen to facilitate communication and prediction. The first author has attempted to develop such an interpretation in Pearl (1988) and it seems appropriate to quote some related passages:

> The asymmetry conveyed by causal directionality is viewed as a notational device for encoding still more intricate patterns of relevance relationships, such as *nontransitive* and *induced* dependencies... . Two events do not become relevant to each other merely by virtue of predicting a common consequence, but they do become relevant when the consequence is actually observed. The opposite is true for two consequences of a common cause (pages 18–19).
>
> Note that the topology of a Bayesian network can be extremely sensitive to the node ordering d. What is a tree in one ordering might become a complete

*Similar objections can be raised against the use of the term 'influence', and no doubt were raised centuries ago when the word 'dependence' was first introduced into probability theory.

graph if the ordering is reversed.... This sensitivity to order may seem paradoxical at first; d can be chosen arbitrarily, whereas people have fairly uniform conceptual structures, e.g. they agree on whether two propositions are directly or indirectly related. This consensus about the structure of dependencies shows the dominant role causality plays in the formation of these structures. In other words, the standard ordering imposed by the direction of causation indirectly induces identical topologies on the networks that people adopt to encode experiential knowledge. Were it not for the social convention of adopting a standard ordering of events that conforms to the flow of time and causation, human communication as we know it might be impossible. Why, then, do we use temporal ordering to organize our memory? It may be because information about temporal precedence is more readily available than other indexing information, or it may be that networks constructed with temporal ordering are inherently more parsimonious (i.e. they display more independencies). Experience with expert systems applications does not entirely rule out the second possibility (Shachter and Heckerman, 1987) (pages 125–126).

REFERENCES

Beeri, C., Fagin, R., and Howard, R. A. (1977) A complete axiomatization of functional dependencies and multi-valued dependencies in database relations, *Proceedings, 1977 ACM SIGMOD Int. Conf. on Mgmt of Data* Toronto, Canada, pp. 47–61.

Blalock, H. M. (1971) *Causal Models in the Social Sciences*, Macmillan, London.

Cramér, H. (1946) *Mathematical Methods of Statistics*, Princeton University Press, Princeton, N.J.

Dawid, A. P. (1979) Conditional independence in statistical theory, *J.R. Statist. Soc.*, **B41**(1), 1–31.

Duncan, O. D. (1975) *Introduction to Structural Equation Models*, Academic Press, New York.

Fagin, R. (1977) Multivalued dependencies and a new form for relational databases, *ACM Transactions on Database Systems*, **2**(3), 262–78, September.

Feller, W. (1968) *An Introduction to Probability Theory and its Applications*, 3rd edn, Wiley, New York, Chapter XV.

Geiger, D. (1987) *Towards the Formalization of Informational Dependencies*, UCLA Cognitive Systems Laboratory, Technical Report R-102, December.

Geiger, D., Paz, A., and Pearl, J. (1988) *Axioms and Algorithms for Inferences Involving Probabilistic Independence*, Technical Report (R-119), in preparation.

Geiger, D., and Pearl, J. (1988a) On the logic of causal models, *Proc. of the 4th Workshop on Uncertainty in AI*, St Paul, Minn., August, pp. 136–47.

Geiger, D., and Pearl, J. (1988b) *Logical and Algorithmic Properties of Conditional Independence*, UCLA Cognitive Systems Laboratory, Technical Report 870056 (R-97), February. To appear in *Proc. of Workshop on Statistics and AI*, Ft Lauderdale, Fla, January 1989.

Geiger, D., and Verma, T. (1988) *A Linear Algorithm for Recognizing Independence in Bayesian Networks*, UCLA Cognitive Systems Laboratory, Technical Report R-116, in preparation.

Good, I. J. (1961) A causal calculus, *Philosophy of Science*, **11**, 305–18. Also in I. J. Good, *Good Thinking: The Foundations of Probability and its Applications*, University of Minnesota Press, Minneapolis (1983).

Howard, R. A., and Matheson, J. E. (1981) Influence diagrams. In *Principles and Applications of Decision Analysis*, Strategic Decisions Group, Menlo Park, Calif.

Isham, V. (1981) An introduction to spatial point processes and Markov random fields, *International Statistical Review*, **49**, 21–43.

Kenny, D. A. (1979) *Correlation and Causality*, Wiley, New York.

Kiiveri, H., Speed, T. P., and Carlin, J. B. (1984) Recursive causal models, *Journal of Australian Math Society*, **36**, 30–52.

Lauritzen, S. L. (1982) *Lectures on Contingency Tables*, 2nd edn, University of Aalborg Press, Aalborg, Denmark.

Meditch, J. S. (1969) *Stochastic Optimal Linear Estimation and Control*, McGraw-Hill, New York.

Olmsted, S. M. (1983) On representing and solving decision problems, Ph.D. thesis, EES Dept, Stanford University.

Pearl, J. (1985) A constraint propagation approach to probabilistic reasoning, *Proc. of the First Workshop on Uncertainty in AI*, Los Angeles, August, pp. 31–42.

Pearl, J. (1986) Fusion, propagation and structuring in belief networks, *Artificial Intelligence*, **29**(3) September, 241–88.

Pearl, J. (1988) *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, San Mateo, Calif.

Pearl, J., and Paz, A. (1985) *GRAPHOIDS: a graph-based logic for reasoning about relevance relations*, UCLA Computer Science Department Technical Report 8500038 (R-53), October; also, *Proceedings, ECAI-86*, Brighton, UK, June 1986.

Pearl, J., and Verma, T. (1987) The logic of representing dependencies by directed acyclic graphs, *Proc. AAA-I*, Seattle, Wash., July, pp. 374–79.

Shachter, R. D. (1985) Intelligent probabilistic inference, *Proc. of the First Workshop on Uncertainty in AI*, Los Angeles, August, pp. 237–44.

Shachter, R. D. (1986) Evaluating influence diagrams, *Operations Research*, **34**(6), 871–82.

Shachter, R. D., and Heckerman, D. (1987) Thinking backward for knowledge acquisition, *AI Magazine*, **8**, 55–62.

Shachter, R. D. (1988) Probabilistic inference and influence diagrams, to appear in *Operations Research*.

Shafer, G., Shenoy, P. P., and Mellouli, K. (1987) Propagating belief functions in qualitative Markov trees, The University of Kansas, Lawrence School of Business, Working Paper No. 190, June. To appear in *International Journal of Approximate Reasoning*.

Smith, J. Q. (1987) *Influence Diagrams for Statistical Modelling*, Technical Report No. 117, Department of Statistics, University of Warwick, Coventry, England, June.

Studeny, M. (1988) Attempts at axiomatic description of conditional independence, *Workshop on Uncertainty processing in Expert Systems*, Alsovice, Czechoslovakia, 20–23, June.

Verma, T. S. (1986) *Causal Networks: Semantics and Expressiveness*, UCLA Cognitive Systems Laboratory Technical Report R-65.

Verma, T. S. (1987) *Some Mathematical Properties of Dependency Models*, UCLA Cognitive Systems Laboratory Technical Report R-103.

Verma, T., and Pearl, J. (1987) The logic of representing dependencies by directed acyclic graphs, *Proc. AAAI*, Seattle, Wash., July, pp. 374–79.

Verma, T. and Pearl, J. (1988) Causal networks: semantics and expressiveness, *Proceedings of the 4th Workshop on Uncertainty in AI*, St Paul, Minn., pp. 352–9.

Wermuth, N. and Lauritzen, S. L. (1983) Graphical and recursive models for contingency tables, *Biometrika*, **70**, 537–52.

Wold, H. (1964) *Econometric Model Building*, North-Holland, Amsterdam.

Wright, S. (1921) Correlation and causation, *Journal of Agricultural Research*, **20**, 557–85.

Wright, S. (1934) The method of path coefficients, *Ann. Math. Statist.*, **5**, 161–215.