# Graphoids and Natural Conditional Functions

**Daniel Hunter**
*Northrop Research and Technology Center, Palos Verdes Peninsula, California*

## ABSTRACT

*A natural conditional function (NCF) is a function from possible states of the world to natural numbers, representing the degrees of implausibility of the different states. This paper shows that the relation of conditional independence, when defined in terms of NCFs, has a natural representation in terms of influence diagrams. First it is shown that the relation of conditional independence relative to an NCF satisfies certain axioms for conditional independence (the graphoid axioms). Then it is proved that the conditional independencies deducible from the graphoid axioms together with a set of conditional independence statements structured in a certain way (those forming a causal input list) are exactly the conditional independencies semantically implied by that input list and are also identical with the set of conditional independencies that can be read off the corresponding influence diagram using the graphical criterion of d-separation. The computational implications of these results are discussed.*

KEYWORDS: ***belief revision, influence diagrams, nonmonotonic reasoning, belief networks, graphoids***

## INTRODUCTION

Geiger and Pearl [1] and Verma and Pearl [2] prove important results concerning the logic of influence diagrams. They give an axiomatic characterization of the relation of conditional independence and show that the entailments of a certain class of conditional independence statements, those expressible in a *causal input list*, are captured by those independencies that can be read off the corresponding influence diagram using the criterion of d-separation.

This paper proves corresponding results for the relation of conditional independence as defined by natural conditional functions (NCFs). NCFs were introduced by Wolfgang Spohn [3] as a means of analyzing deterministic or

---

*Address correspondence to Daniel Hunter, Northrop Research Center, One Research Park, Palos Verdes Peninsula, CA 90274.*
Received June 1, 1990; accepted December 27, 1990.

International Journal of Approximate Reasoning 1991; 5:489–504

655 Avenue of the Americas, New York, NY 10010 0888-613X/91/$3.50

categorical belief. It is shown that conditional independence so defined satisfies the axioms of graphoids and that the graphical criterion of d-separation captures implications about NCF conditional independence for causal input lists.

First some notational matters. I use letters from the front of the alphabet—$A$, $B$, $C$, etc., with or without numerical subscripts—for multivalued variables; the letters $W$, $X$, $Y$, and $Z$ are used to denote *sets* of multivalued variables; and a lowercase letter is used to denote a value of the corresponding capital letter. I use $\mathscr{V}(\phi)$ denote the set of values of $\phi$, where $\phi$ is either a variable or a set of variables. A value for a set of variables is a set of values of the variables in the set. It is assumed that each variable has more than one value and that, unless otherwise indicated, distinct variables have disjoint sets of values. This latter condition can be made to hold by identifying a value $a$ of a variable $A$ with the ordered pair $(A, a)$. "iff" is used as an abbreviation for "if and only if."

## SPOHN'S THEORY OF BELIEF REVISION

This section describes the motivation for, and the essential details of, Spohn's theory of belief revision. There are many interesting aspects of Spohn's theory that cannot be covered here. The reader is urged to consult Ref. 3 for a fuller presentation of the theory.

Probability theory provides a widely accepted framework for understanding the revision of *degrees* of belief. The use of Bayesian networks (Pearl [4]) as a tool for probabilistic updating is a burgeoning research area. Moreover, probability theory has been found useful in analyzing philosophical issues related to causation and explanation (Harper and Skyrms [5]).

It is striking that before Spohn's work on belief revision, no formalism of comparable power was available for the case of deterministic, or categorical, belief. Categorical belief is an all-or-nothing affair; one either believes, disbelieves, or does neither, concerning some proposition. Categorical belief seems to be the normal kind of belief that people have. People generally do not represent their beliefs numerically except in the context of a game of chance. Nonetheless, if pressed, most people will admit that their beliefs come in varying strengths—that, for example, they believe more strongly that the sun will rise tomorrow than that their car will be working tomorrow. This might lead some to think that ordinary belief could be represented in terms of perhaps somewhat vague probabilities, with definite belief in a proposition amounting to that proposition's having a (subjective) probability greater than some threshold. This suggestion fails for the following reason. Categorical belief is closed under implication: If a person believes each one of the propositions $A_1, \ldots, A_n$ and is led to see (e.g., via a proof) that $A_1, \ldots, A_n$ logically imply a proposition $B$, then that person should believe $B$. So a logically omniscient believer's beliefs would be closed under logical implication. However, probabilistic belief, as defined above, is not closed under logical implications. For any $\epsilon > 0$, we may have the probability of each of $A_1, \ldots, A_n$ being greater than $1 - \epsilon$ while some implication of the $A_i$ has probability less than $1 - \epsilon$. A much discussed example of this sort is the lottery paradox (Kyburg [6]): Let $0 < \epsilon < 1$. Consider a fair lottery with $N$ participants, where $N > 1/\epsilon$. If there is only one possible winner, the probability of any one participant's not winning is greater than $1 - \epsilon$. However, the conjunction of the statements "$x$ does not win," for each participant $x$, has probability zero, since (we may suppose) it is known that someone will win.

The conclusion from the above is that plain belief cannot be analyzed in probabilistic terms. If we accept this conclusion, the challenge is to produce a theory as well suited to analyzing plain belief and the revision of plain belief as is probability theory to probabilistic belief and probabilistic belief change. Spohn has done just that.

## NATURAL CONDITIONAL FUNCTIONS

The first step in analyzing plain belief and plain belief change (hereafter just "belief" and "belief change") is to formally represent the total content of someone's belief. Then belief in a particular proposition can be defined as entailment by the total content of one's belief (under the idealization of logical omniscience).

Spohn chose to work within the possible worlds framework in his analysis. We may think of a possible world as a complete and consistent description of the state of the world relative to some set of atomic propositions. The idea is that we represent a person's belief state as an assignment of a degree of implausibility to the possible worlds. Let $\Theta = \{s_1, s_2, \ldots, s_n\}$ be a set of pairwise mutually exclusive and jointly exhaustive states of the world. In Spohn's theory, a state of belief is captured by a *natural conditional function* (NCF), which is a function from $\Theta$ into the natural numbers $(0, 1, 2, \ldots)$ that assigns zero to at least one member of $\Theta$.[1] Intuitively, an NCF is a grading of states of the world in terms of their degree of *implausibility*. That is, the greater the value of the NCF for a given state, the more implausible that state is.

[1] In fact, Spohn considered a more general function known as an *ordinal condition function* (OCF). An OCF is like an NCF except that its range is the set of ordinal numbers, including transfinite ordinals. To avoid technical trivialities having to do with the arithmetic of transfinite ordinals, we restrict the range to natural numbers.

A *proposition* is a statement whose truth value depends upon which member of $\Theta$ is the true state of the world. Propositions are true or false in states. We follow the convention of identifying a proposition with the set of states in which it is true. Formally, then, a proposition is simply a subset of $\Theta$. An NCF $\kappa$ can be extended to consistent (i.e., nonempty) propositions by defining for each nonempty subset $A$ of $\Theta$

$$\kappa(A) = \min\{\kappa(s_i) \mid s_i \in A\}$$

In this paper we will think of the possible states as complete specifications of the values of some finite set of multivalued variables. In this framework a proposition is an assertion about the values of some subset of the variables. Let us assume in what follows that each value of a variable is unique to that variable, so that we can write $\kappa(a)$ for the degree of implausibility that variable $A$ has value $a$, and similarly $\kappa(x)$ for the degree of implausibility that the values in $x$ belong to the variables in $X$.

A notion of conditional independence for the Spohn system can be defined. Let $X, Y, Z$ be sets of variables and $\kappa$ an NCF. $\kappa(x \mid y)$ is defined to be $\kappa(x, y) - \kappa(y)$. Then we may define "$X$ is independent of $Y$ given $Z$ with respect to $\kappa$" by

For all values $x$ of $X$, $y$ of $Y$, and $z$ of $Z$, $\kappa(x \mid y, z) = \kappa(x \mid z)$.

It is easy to show that this definition is equivalent to

$$\kappa(x, y \mid z) = \kappa(x \mid z) + \kappa(y \mid z).$$

An NCF $\kappa$ induces a *strength of belief* function $\beta$ over proper subsets $A$ of $\Theta$ as follows:

$$\beta(A) =_{\text{df}} \begin{cases} -\kappa(A) & \text{if } \kappa(A) > 0 \\ \kappa(\neg A) & \text{otherwise} \end{cases}$$

Now we can say how beliefs are revised when new information is obtained. Suppose that proposition $P$ is learned with strength $\alpha$. Let $\kappa$ be the NCF before $P$ is learned. Then we define $\kappa'$, the NCF that results from learning $P$, by its value for each state $s$ as follows:

$$\kappa'(s) = \begin{cases} \kappa(s) - \kappa(P) & \text{if } s \in P \\ \kappa(s) - \kappa(\neg P) + \alpha & \textit{otherwise} \end{cases}$$

## AN EXAMPLE OF BELIEF REVISION

In order to make the foregoing belief representation framework more intuitive and to introduce the issues of concern, let us consider an example of a belief system in which the relation of conditional independence plays an important role in organizing the beliefs. Suppose we consider the effects of a fire breaking out in a certain building. We may believe fairly strongly that given that there is a fire in the building, there will be smoke in the building, and that given smoke in the building, a smoke alarm will go off. We may also believe, perhaps somewhat less strongly, that there will be an explosion if a fire breaks out, due, say, to a suspected gas leak. Suppose we also believe very strongly that in the absence of a fire, none of the other mentioned events would occur. Finally, suppose we have no belief one way or the other about whether a fire will occur.

The causal relations between the foregoing events are indicated in the influence diagram shown in Figure 1. (We consider the nodes in this network to be binary variables taking the values *true* and *false*, or 1 and 0.) Intuitively, the indicated causal relations between events will imply certain conditional independencies. For example, once we know the value of the variable SMOKE, the value of the variable ALARM becomes independent of the variable FIRE. Similarly, if we were to learn the true value of the variable FIRE, then the variables SMOKE and EXPLOSION would become independent. In general, knowledge of the direct cause or causes of a given variable makes belief about the variable independent of beliefs about any other variables that are connected to the given variable only through the direct causes.

To formalize the above belief system in the Spohnian framework, first note that the strength of belief in a proposition $A$ conditional on a proposition $B$ is equal to a positive integer $x$ [in symbols: $\beta(A \mid B) = x$] if and only if

$$\kappa(\neg A \mid B) = x,$$

where $\kappa$ is the corresponding NCF. Also note that $\kappa(\neg A \mid B) > 0$ implies $\kappa(A \mid B) = 0$.

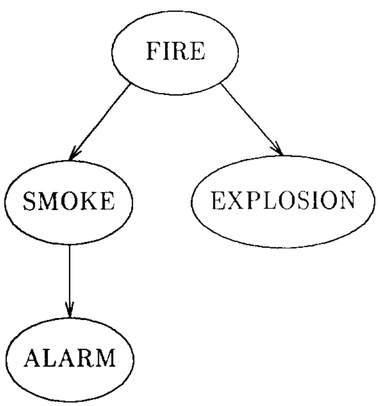


**Figure 1.** Causal relations between events.

Abbreviating the name of each event by the first letter in the name, the beliefs about the causal relations between events can be captured by conditions on the NCF $\kappa$ such as the following:

$$\kappa(\neg E \mid F) = 1, \qquad \kappa(E \mid \neg F) = 4$$
$$\kappa(\neg S \mid F) = 2, \qquad \kappa(S \mid \neg F) = 4$$
$$\kappa(\neg A \mid S) = 2, \qquad \kappa(A \mid \neg S) = 4$$

The other cases follow from the fact that $\kappa(\neg A \mid B) > 0$ implies $\kappa(A \mid B) = 0$. The absence of any belief regarding FIRE can be captured by

$$\kappa(F) = \kappa(\neg F) = 0.$$

Call the set of above conditions $\Gamma$. By itself $\Gamma$ does not determine a joint NCF over the set of all variables. However, if we add the conditional independencies implied by the influence diagram, we do uniquely determine an NCF. In fact, the following statements of conditional independency suffice to determine the joint (a lowercase letter ranges over values of the corresponding capital letter):

$$\kappa(s \mid f, e) = \kappa(s \mid f)$$
$$\kappa(a \mid s, f, e) = \kappa(a \mid s)$$

These two conditions imply that a joint $\kappa(f, s, e, a)$ may be written

$$\kappa(f) + \kappa(s \mid f) + \kappa(e \mid f) + \kappa(a \mid s).$$

All the terms in the above expression are determined by $\Gamma$, so this proves the assertion just made, that the stated conditions on the NCF, together with the conditional independencies, uniquely determine the NCF.

To revise beliefs in the light of new information, we use the decomposition of the joint NCF together with the revision rule to update each proposition. For example, if ALARM comes to be believed with strength, say, 3, we may compute the new NCF $\kappa'$ for other propositions as follows. First we have

$$\begin{aligned}
\kappa'(S) &= \min\{\kappa'(S \& A), \kappa'(S \& \neg A)\} \\
&= \min\{\kappa(S \& A) - \kappa(A), \kappa(S \& \neg A) - \kappa(\neg A) + 3\} \\
&= \min\{0, 2\} = 0
\end{aligned}$$

while

$$\begin{aligned}
\kappa'(\neg S) &= \min\{\kappa(\neg S \& A) - \kappa(A), \kappa(\neg S \& \neg A) - \kappa(\neg A) + 3\} \\
&= \min\{4, 3\} = 3.
\end{aligned}$$

Hence in the new NCF, SMOKE is believed with strength 3. The new strength of belief in FIRE can be computed in a similar manner from the new strength of belief in SMOKE, and from there the strength of belief in EXPLOSION can be computed.

The above example suggests a method for constructing and revising NCFs over a set of variables. Construct an influence diagram capturing the dependencies between the variables. Then for each variable $V$ specify a marginal NCF for $V$ conditional on each combination of values of $V$'s parents (if $V$ has no parents, this amounts to specifying an unconditional marginal NCF for $V$). The conditional independencies implied by the influence diagram will then uniquely determine the NCF. The influence diagram is updated by applying the revision rule locally to directly connected components.

This procedure makes a number of assumptions that need to be made explicit. First, we are assuming that our intuitive notion of conditional independency can be captured by the graphical technique of an influence diagram. Second, we are assuming that the notion of conditional independence as defined in terms of NCFs corresponds in an appropriate way to both our intuitive notion and the conditional independencies indicated by the influence diagram. Neither of these assumptions can be taken for granted. The first assumption has been proved true in Refs. 1 and 2 for an important class of conditional independence statements. The burden of the rest of this paper is to prove the second assumption for the same class of statements.

## GRAPHOIDS AND SEMIGRAPHOIDS

What axioms should a relation of conditional independence, whether defined in terms of NCF's, probabilities, or some other way, satisfy? Let $I(X, Z, Y)$ mean that the set of variables $X$ is independent of the set of variables $Y$ given the set of variables $Z$. In their work on probabilistic conditional independence, Geiger and Pearl [1] list the following axioms, taken from Dawid [7]:

**(1a)** $I(X, Z, Y) \Leftrightarrow I(Y, Z, X)$

**(1b)** $I(X, Z, Y \cup W) \Rightarrow I(X, Z, Y)$

**(1c)** $I(X, Z, Y \cup W) \Rightarrow I(X, Z \cup Y, W)$

**(1d)** $I(X, Z \cup Y, W) \wedge I(X, Z, Y) \Rightarrow I(X, Z, Y \cup W)$

We denote the set consisting of axioms (1a)–(1d) by SG. A relation $I$ that satisfies all the members of SG is a *semigraphoid*. If the relation additionally satisfies the following axiom, then it is a *graphoid*:

$$(1e) \quad I(X, Z \cup Y, W) \& I(X, Z \cup W, Y) \Rightarrow I(X, Z, Y \cup W)$$

where $W$ is disjoint from $X$, $Y$, and $Z$.

We have the following theorem.

THEOREM 1 *The relation of conditional independence relative to an NCF is a graphoid.*

Proof First we need to note that since the sets $W$, $X$, $Y$, and $Z$ are pairwise disjoint, any combination of values of these variables is consistent, so that the NCF value for any such combination will be defined.

(1a) Immediate.

(1b) Assume $\kappa(X, Y \cup W \mid Z) = \kappa(X \mid Z) + \kappa(Y \cup W \mid Z)$. Then

$$\begin{aligned}\kappa(x, y \mid z) &= \min\{\kappa(x, y, w \mid z) : w \in \mathscr{V}(W)\}\\ &= \min\{\kappa(x \mid z) + \kappa(y, w \mid z) : w \in \mathscr{V}(W)\}\\ &= \kappa(x, y \mid z) + \min\{\kappa(y, w \mid z) : w \in \mathscr{V}(W)\}\\ &= \kappa(x \mid z) + \kappa(y \mid z).\end{aligned}$$

(1c) Assume $\kappa(X, Y \cup W \mid Z) = \kappa(X \mid Z) + \kappa(Y \cup W \mid Z)$. It easily follows that $\kappa(x, w \mid z, y) = \kappa(x \mid z) + \kappa(w \mid z, y)$. From (1b) we have $\kappa(x \mid z) = \kappa(x, y \mid z) - \kappa(y \mid z) = \kappa(x \mid y, z)$. Hence $\kappa(x, w \mid z, y) = \kappa(x \mid z, y) + \kappa(w \mid z, y)$.

(1d) Assume $\kappa(X, W \mid Z, Y) = \kappa(X \mid Z, Y) + \kappa(W \mid Z, Y)$ and $\kappa(X, Y \mid Z) = \kappa(X \mid Z) + \kappa(Y \mid Z)$. Then

$$\begin{aligned}\kappa(x, y, w \mid z) &= \kappa(y \mid z) + \kappa(x \mid z, y) + \kappa(w \mid x, z, y)\\ &= \kappa(x \mid z) + \kappa(y \mid z) + \kappa(w \mid z, y)\\ &= \kappa(x \mid z) + \kappa(w, y \mid z).\end{aligned}$$

(1e) Assume $\kappa(x \mid z, y, w) = \kappa(x \mid z, y)$ and $\kappa(x \mid z, y, w) = \kappa(x \mid z, w)$ for all $x \in \mathscr{V}(X)$, $y \in \mathscr{V}(Y)$, $z \in \mathscr{V}(Z)$, and $w \in \mathscr{V}(W)$. Consider an arbitrary value $z$ of $Z$. We have $\kappa(x \mid z, y) = \kappa(x \mid z, w) = \kappa(x \mid z, y, w)$ for every value $y$ of $Y$ and every value $w$ of $W$. This implies that for a particular choice of $z$, $\kappa(x \mid z, y, w)$ (with $y$ and $w$ varying) is a constant, call it $\alpha_z$. Thus

$$\begin{aligned}\kappa(x \mid z) &= \min\{\kappa(y, w \mid z) + \kappa(x \mid y, w, z) : y \in \mathscr{V}(Y), w \in \mathscr{V}(W)\}\\ &= \alpha_z + \min\{\kappa(x \mid y, w, z) : y \in \mathscr{V}(Y), w \in \mathscr{V}(W)\}\\ &= \kappa(x \mid z, y, w).\end{aligned}$$ ■

Note that a probabilistic version of the proof of (1e) would require an assumption of strict positivity (no combination of values has zero probability) to ensure that the needed conditional probabilities are defined.

Let $\theta$ be a total ordering of the variables. A *causal list* (*relative to* $\theta$) is a set of statements such that for each variable $A$ there occurs in the list exactly

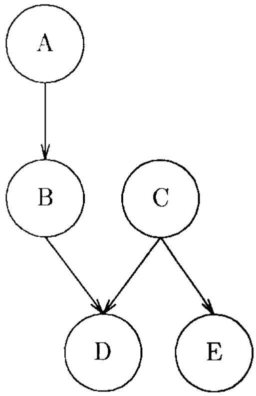


**Figure 2.** An influence diagram.

one statement of the form

$$I(A, X, Y - X)$$

where $Y$ is the set of ancestors of $A$ and $X$ is a subset of $Y$. The statement $I(A, X, Y - X)$ therefore says that given values for the $X$'s, $A$ is independent of all other predecessors in the ordering $\theta$. A causal list can be used to produce an influence diagram by drawing arrows, for each statement $I(A, X, Y - X)$ in the list, from the variables in $X$ to the variable $A$.

A causal list $L$ entails independence statements other than those in $L$. For example, $L$ will entail that, given its parents, a variable $A$ is independent, not just of all other ancestors, but of all nondescendant variables. We say that a causal list $L$ entails an independence statement $I$, written $L \models I$, if and only if every NCF that satisfies every statement in $L$ also satisfies $I$.

Given an influence diagram $D$, there is a graphical criterion for conditional independence (Pearl [4, p. 117]). To state this criterion we need some definitions. We say that a path from variable $\alpha$ to variable $\beta$ is *viable* if it contains no node with two incoming arrows. Thus the path $ABD$ from $A$ to $D$ in Figure 2 is viable because it contains no nodes with arrows meeting head to head. However, the path $BDC$ from $B$ to $C$ is not viable because at node $D$ two arrows meet head to head. A viable path is *blocked* by a set of variables if some member of the set occurs in the path. For example, the set $\{B\}$ blocks the path $ABD$. A nonviable path is *unblocked* by a set of variables if every node at which arrows meet head to head either is in the set or has a descendant in the set. Hence the path $BDC$ is unblocked by the set $\{D\}$ (and by any superset of that set). The d-separation criterion then says that $X$ and $Y$ are independent given $Z$ in influence diagram ID, written $I_{\text{ID}}(X, Y, Z)$, if every *viable* path in ID from a variable in $X$ to a variable in $Y$ is *blocked* by $Z$ and no nonviable path in ID from a variable in $X$ to a variable in $Y$ is *unblocked* by $Z$.

So far we have seen three approaches to conditional independence: the definition in terms of NCFs, the axiomatic approach, and the graphical d-separation criterion. What, if any, is the connection among these three? The following theorem answers this question.

THEOREM 2 *Let $L$ be a causal list and $D$ the corresponding influence diagram. Then the following statements are equivalent:*

(i) $L \models I(X, Z, Y)$

(ii) $L \cup \mathrm{SG} \vdash I(X, Z, Y)$

(iii) $I_D(X, Z, Y)$

The proof of Theorem 2 requires the following lemmas.

LEMMA 1 *Let $L$ be a causal list with respect to the ordering of variables $(A_1, A_2, \ldots, A_n)$, and let $D$ be an influence diagram for $L$. Then an NCF $\kappa$ satisfies $L$ iff for all values $a_1, a_2, \ldots, a_n$ of $A_1, A_2, \ldots, A_n$, respectively,*

$$\kappa(a_1, a_2, \ldots, a_n) = \sum_{i=1}^{n} \kappa(a_i \mid b_{i1}, \ldots, b_{il_i})$$

*where $b_{i1}, \ldots, b_{il_i}$ are the values of $A_i$'s parents in $D$ that are in $\{a_1, \ldots, a_n\}$.*

Proof $\kappa(a_1, a_2, \ldots, a_n)$ can be written as $\kappa(a_1) + \kappa(a_2 \mid a_1) + \kappa(a_n \mid a_1, \ldots, a_{n-1})$. By the choice of the parents of a variable, we see that in the $i$th term of this sum, values of variables that are not parents of $A_i$ may be removed without affecting the sum. ■

For the next lemma, we need to define two notions. First, where $L$ is a causal list for a set of variables $V$ and $X$ is a subset of $V$, we define $L(X)$, the *restriction of $L$ to $X$*, to be the causal list gotten from $L$ by deleting, for each variable $A$ not in $X$, the entry $I(A, Z, Y)$ and by replacing, for each variable $B$ in $X$, the entry $I(B, Z, Y)$ by $I(B, Z \cap X, Y \cap X)$. Second, if NCF $\kappa$ is defined over a set $X$ of variables, we define the *vacuous extension* of $\kappa$ to a set $V \supseteq X$ to be the NCF $\kappa_V^+$ such that for any set of values $x$ of members of $X$ and values $v$ of members of $V - X$, $\kappa_V^+(x, v) = \kappa(x)$.

LEMMA 2 *Let $L$ be a causal list for a set of variables $V$, and let $X$ be a subset of $V$. If an NCF $\kappa$ satisfies $L(X)$, then the vacuous extension of $\kappa$ to $V$ satisfies $L$.*

Proof Suppose NCF $\kappa$, over $X$, satisfies $L(X)$, and suppose $I(A, Z, Y) \in L$. If $A \notin X$, then for any value $a$ of $A$ and sets of values $z$ and $y$ of $Z$ and

$Y$, respectively, $\kappa_V^+(a \mid z) = \kappa_V^+(a \mid z, y) = 0$, and so $\kappa_V^+$ satisfies $I(A, Z, Y)$. Assume, then, that $A \in X$. Then $I(A, Z \cap X, Y \cap X) \in L(X)$. Let $Z_1 = Z \cap X$ and $Z_2 = Z - X$, and similarly let $Y_1 = Y \cap X$ and $Y_2 = Y - X$. Then any set $z$ of values of the members of $Z$ can be written as the union of a set $z_1$ of values of members of $Z_1$ and a disjoint set $z_2$ of members of $Z_2$, and a set $y$ of values of members of $Y$ can be likewise be decomposed into the union of disjoint sets $y_1$ of values of members of $Y_1$ and $y_2$ of values of members of $Y_2$. Then

$$\begin{aligned}\kappa_V^+(a \mid z, y) &= \kappa_V^+(a \mid z_1, z_2, y_1, y_2)\\ &= \kappa(a \mid z_1, y_1)\\ &= \kappa(a \mid z_1)\\ &= \kappa_V^+(a \mid z_1, z_2)\\ &= \kappa_V^+(a \mid z).\end{aligned}$$

Now we can return to the proof of Theorem 2.

Proof We show (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) $\Rightarrow$ (iii).

(iii) $\Rightarrow$ (ii). Let $M = \{I(X, Z, Y) : L \cup \mathrm{SG} \vdash I(X, Y, Z)\}$. It is clear that $M$ is a dependency model as defined in Ref. 2 and that $L$ is a causal input list for $M$. Furthermore, $M$ of necessity is a semigraphoid. By Theorem 2 of Ref. 2, it follows that every independency revealed by $D$ via the d-separation criterion is a member of $M$, that is, if $I_D(X, Z, Y)$, then $L \cup \mathrm{SG} \vdash I(X, Z, Y)$.

(ii) $\Rightarrow$ (i). By Theorem 1 of the present paper, every NCF $\kappa$ satisfies SG when $I(X, Z, Y)$ is taken to mean that $X$ and $Y$ are conditionally independent given $Z$ relative to $\kappa$. Hence if an NCF $\kappa$ satisfies $L$ and $L \cup \mathrm{SG} \vdash I(X, Z, Y)$, then $\kappa$ also satisfies $I(X, Z, Y)$.

(i) $\Rightarrow$ (iii). Assume that $\neg\, I_D(X, Z, Y)$. Since $X$ and $Y$ are not d-separated by $Z$, there must exist a path from a node $A$ in $X$ to a node $B$ in $Y$ such that every node with converging arrows is either in $Z$ or has a descendant in $Z$ and no node with an outgoing arrow is in $Z$. Let $P$ be such a path having the least number of converging arrows. The result will be proved if we can construct an NCF $\kappa$ with respect to which $A$ and $B$ are not independent given $Z$. Let $W$ be the set of variables in path $P$ together with all variables along a path from a variable in $P$ to a variable in $Z$ (if there is more than one such path, we pick the shortest one). Let $U$ be $Z \cap W$. By choosing $P$ to be a path from $A$ to $B$ with the smallest number of converging arrows, we ensure that each member of $U$ has exactly one ancestor in $P$; otherwise, as Geiger and Pearl point out [1], there would be a path from $A$ to $B$ with fewer converging arrows than in $P$, contrary to the definition of $P$.

By Lemma 2, if we can find an NCF $\kappa$ over $W$ such that $\kappa$ satisfies $L(W)$ and such that $A$ and $B$ are not independent given $U$, then the vacuous extension of $\kappa$ to $V$, the set of all variables, will satisfy $L$ and will not make $A$ and $B$ independent given $Z$.

In defining the required NCF $\kappa$, we wish to treat nonbinary variables as if they were binary. This may be done as follows. First, pick a designated value of each variable and call it 1. Next, label the set of nondesignated values of a variable by 0. Taking the domain of values of each variable to be $\{0, 1\}$, we will then define an NCF $\kappa$ over the cross-product space of the variables. $\kappa$ could be extended to the actual domain of each variable, but nothing in the proof depends upon how the extension is done.

If $C$ is a root variable, we define

$$\kappa(C = 1) = 0, \qquad \kappa(C = 0) = 1$$

If $C$ has exactly one parent $D$, we define

$$\kappa(C = c \mid D = d) = \begin{cases} 0 & \text{if } c = d \\ 1 & \text{otherwise} \end{cases}$$

If $C$ has parents $D_1, \ldots, D_k$, $k > 1$, we define

$$\kappa(C = c \mid D_1 = d_1, \ldots, D_k = d_k) = \begin{cases} 0 & \text{if } c = 1 \text{ and } d_1 = \ldots = d_k \\ 0 & \text{if } c = 0 \text{ and the } d_i \text{ are not all equal} \\ 1 & \text{otherwise} \end{cases}$$

By defining $\kappa$ over all variables in accordance with the equation in Lemma 1, the independencies of $L(W)$ will be satisfied.

We state without proof the following lemmas.

LEMMA 3 *For all variables* $C_1, \ldots, C_n$, $\kappa(C_1 = C_2 = \cdots = C_n = 1) = 0$.

LEMMA 4 *For every variable* $C$, $\kappa(C = 0) = 1$.

We now show by induction on the length of the path from $A$ to $B$ that $\kappa(B = 0 \mid A = 0, U \equiv 1) = 0$. The result will follow since by the above lemmas, $\kappa(B = 0 \mid U \equiv 1) > 0$.

For the basis case, suppose that the length of the path from $A$ to $B$ is 1. Then either $A$ is the sole parent of $B$ or $B$ is the sole parent of $A$. In either case, it is trivial to show $\kappa(B = 0 \mid A = 0) = 0$.

Suppose, then, that the path from $A$ to $B$ is of length greater than 1. Then $B$'s immediate neighbor on path $P$ is a node $C$ distinct from $A$. There are two cases to consider. Either $C$ is a head-to-head node or it is not. First assume it is not. Then $C$ d-separates $A$ and $U$ from $B$, and we may write $\kappa(B = 0 \mid A = 0, U \equiv 1)$ as

$$\min_c \{\kappa(C = c \mid A = 0, U \equiv 1) + \kappa(B = 0 \mid C = c)\}.$$

By induction, $\kappa(C = 0 \mid A = 0, U \equiv 1) = \kappa(B = 0 \mid C = 0) = 0$. Hence $\kappa(B = 0 \mid A = 0, U \equiv 1) = 0$.

Now suppose $C$ is a head-to-head node. $B$ is one parent of $C$. Let $D$ be the other parent (by the choice of $W$, no variable has more than two parents in $W$). Let $E$ stand for the first descendent of $C$ (possibly $C$ itself) in $U$. Let $U_1$ stand for $U - \{E\}$. Since $E$'s only ancestor in $P$ is $C$, the set $\{D, E\}$ d-separates $C$ from $\{A\} \cup U_1$ and the set $\{C, D\}$ d-separates $B$ from $\{A\} \cup \{E\} \cup U_1$. Since we have already established that conditional independence under the d-separation criterion implies conditional independence with respect to any NCF satisfying the causal list, we may write $\kappa(B = 0 \mid A = 0, U \equiv 1)$ as

$$\min_{d,c} \{\kappa(D = d \mid A = 0, U_1 \equiv 1) + \kappa(C = c \mid D = d, E = 1) + \kappa(B = 0 \mid D = d, C = c)\}.$$

By induction, $\kappa(D = 0 \mid A = 0, U_1 \equiv 1) = 0$. We have

$$\begin{aligned}
\kappa(C = 1 \mid D = 0, E = 1) &= \kappa(D = 0) + \kappa(C = 1 \mid D = 0) + \kappa(E = 1 \mid C = 1) \\
&\quad - \min_c \{\kappa(D = 0) + \kappa(C = c \mid D = 0) + \kappa(E = 1 \mid C = c)\} \\
&= 1 - \min_c \{\kappa(C = c \mid D = 0) + \kappa(E = 1 \mid C = c)\} \\
&= 0.
\end{aligned}$$

And a similar argument shows that $\kappa(B = 0 \mid D = 0, C = 1) = 0$. Hence $\kappa(B = 0 \mid A = 0, U \equiv 1) = 0$.

That completes the proof of Theorem 2. ■

## COMPARISONS

Spohn's theory of belief revision lends itself to a computationally efficient implementation of iterative nonmonotonic reasoning. It is of interest to compare Spohn's theory with other theories of nonmonotonic reasoning. A detailed comparison is out of place here, but certain recent developments in the field of nonmonotonic reasoning are worth examining in the light of Spohn's work on belief revision.

Kraus et al. [8] and Makinson [9] give axioms for a relation of nonmonotonic inference and show that the best-known systems for nonmonotonic reasoning-default logic, circumscription, McDermott and Doyle's modal systems, and autoepistemic logic—fail to satisfy one or more of these axioms.

Hanks and McDermottt [10] describe anomalies in the application of default rules and circumscription to an intuitive case of nonmonotonic reasoning. In general, there is a growing awareness of the inadequacy of traditional approaches to nonmonotonic reasoning.

Recent work on preferential and ranked preferential models (Kraus et al. [8], Makinson [9], Shoham [11]) has seemed to some to promise relief from some of these difficulties. In a preferential model, the worlds or states are related by a binary preference relation $<$. An inference relation $\mid\sim$ is defined by saying that $A \mid\sim B$ if and only if $B$ is true in all the most preferred worlds in which $A$ is true. Depending on the properties of $<$, various nonmonotonic logics result from this definition. In a ranked preferential model, the preference relation may be thought of as stemming from a total ordering of some partition of the worlds (so that where $w$ and $v$ are worlds, $w < v$ if and only if $w$ occurs in a partition element preceding the partition element of which $\upsilon$ is a member).

It is clear that an NCF is a ranked preferential model. But it is more than that. There is also a notion of *distance* between worlds, not just of rank. For reasons explained by Spohn [3], this additional expressive power turns out to be important in iterated belief change. In particular, it allows us to say what the new ranking is when a belief change is made. It also allows the recovery of the old ranking when a belief change is "taken back" (e.g., when it is learned that information previously received is incorrect).

Another important difference between Spohn's theory of belief revision and other systems for nonmonotonic reasoning is that the former comes equipped with a powerful notion of conditional independence. In the present paper I have attempted to demonstrate that this notion of conditional independence is sufficiently rich to form the basis for efficient network-structured inference techniques. No comparable inference mechanism has, to my knowledge, been worked out for other systems of nonmonotonic reasoning.

## CONCLUSIONS

The equivalence of the graphical d-separation criterion for conditional independence and the criterion in terms of NCFs has important computational implications. With the theorems proved in this paper, it can be shown that an influence diagram can serve as a computational structure for Spohnian belief updating when each node in the influence diagram stores a marginal NCF over the node and its parents. Updating is done *locally*: Each node need only query its immediate neighbors in order to update its NCF values. This allows updating to be done in parallel. Thus the computational advantages possessed by Bayesian networks (Pearl [4]) are also available for Spohnian updating of deterministic beliefs. I describe an algorithm for performing Spohnian belief revision using influence diagrams in Ref. 12.

In closely related work, Shenoy [13] has shown that the Spohn system for belief revision satisfies the axiomatic framework for local computation developed by Shenoy and Shafer in Ref. 14, and Geiger [15] has proved for the case of probabilistic belief results corresponding to Theorem 2 of the present paper.

## References

1. Geiger, D., and Pearl, J., On the logic of causal models, in *Uncertainty in Artificial Intelligence*, Vol. 4 (R. D. Shachter, T. S. Levitt, L. N. Kanal, and J. F. Lemmer, Eds.), North-Holland, Amsterdam, 3-14, 1990.

2. Verma, T., and Pearl, J., Causal networks: semantics and expressiveness, in *Uncertainty in Artificial Intelligence*, Vol. 4 (R. D. Shachter, T. S. Levitt, L. N. Kanal, and J. F. Lemmer, Eds.), North-Holland, Amsterdam, 69-76, 1990.

3. Spohn, W., Ordinal conditional functions: a dynamic theory of epistemic states, in *Causation in Decision, Belief Change, and Statistics*, Vol. II (W. L. Harper and B. Skyrms, Eds.), Kluwer, Dordrecht, The Netherlands, 105-134, 1988.

4. Pearl, J., *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann, San Mateo, Calif., 1988.

5. Harper, W. L., and Skyrms, B., Eds., *Causation in Decision, Belief Change, and Statistics*, Kluwer, Dordrecht, The Netherlands, 1988.

6. Kyburg, H., *Probability and the Logic of Rational Belief*, Wesleyan Univ. Press, Middleton, Conn., 1961.

7. Dawid, A. P., Conditional independence in statistical theory, *J. Roy. Stat. Soc. Ser. B* **41**, 1-31, 1979.

8. Kraus, S., Lehmann, D., and Magidor, M., Nonmonotonic reasoning, preferential models and cumulative logics, *AI* **44**(1/2), 167-207, 1990.

9. Makinson, D., General theory of non-monotonic reasoning, in *Lecture Notes in Artificial Intelligence*, Vol. 346, *Non-Monotonic Reasoning* (M. Reinfrank, J. de Kleer, M. L. Ginsberg, and E. Sandewall, Eds.), Springer-Verlag, Berlin, 1-18, 1989.

10. Hanks, S., and McDermott, D., Nonmonotonic logic and temporal projection, *AI* **33**(3), 379-412, 1987.

11. Shoham, Y., *Reasoning About Change*, MIT Press, Cambridge, Mass., 1988.

12. Hunter, D., Parallel belief revision, in *Uncertainty in Artificial Intelligence*, Vol. 4 (R. D. Shachter, T. S. Levitt, L. N. Kanal, and J. F. Lemmer, Eds.), North-Holland, Amsterdam, 241-251, 1990.

13. Shenoy, P., On Spohn's rule for revision of beliefs, *Int. J. Approx. Reasoning*,to appear.

14. Shenoy, P., and Shafer, G., Axioms for probability and belief-function propagation, in *Uncertainty in Artificial Intelligence*, Vol. 4 (R. D. Shachter, T. S. Levitt, L. N. Kanal, and J. F. Lemmer, Eds.), North-Holland, Amsterdam, 169-198, 1990.

15. Geiger, D., Graphoids: A qualitative framework for probabilistic inference, Ph.D. dissertation, Univ. Calif., Los Angeles, 1990.