On Spohn's Rule for Revision of Beliefs

Prakash P. Shenoy

School of Business, University of Kansas,
Lawrence, Kansas

ABSTRACT

The main ingredients of Spohn's theory of epistemic beliefs are (1) a functional representation of an epistemic state called a disbelief function and (2) a rule for revising this function in light of new information. The main contribution of this paper is as follows. First, we provide a new axiomatic definition of an epistemic state and study some of its properties. Second, we study some properties of an alternative functional representation of an epistemic state called a Spohnian belief function. Third, we state a rule for combining disbelief functions that is mathematically equivalent to Spohn's belief revision rule. Whereas Spohn's rule is defined in terms of the initial epistemic state and some features of the final epistemic state, the rule of combination is defined in terms of the initial epistemic state and the incremental epistemic state representing the information gained. Fourth, we state a rule of subtraction that allows one to recover the addendum epistemic state from the initial and final epistemic states. Fifth, we study some properties of our rule of combination. One distinct advantage of our rule of combination is that besides belief revision, it can be used to describe an initial epistemic state for many variables when this information is given as several independent epistemic states each involving few variables. Another advantage of our reformulation is that we can show that Spohn's theory of epistemic beliefs shares the essential abstract features of probability theory and the Dempster–Shafer theory of belief functions. One implication of this is that we have a ready-made algorithm for propagating disbelief functions using only local computation.

KEYWORDS: *Spohn's theory, consistent epistemic state, content of an epistemic state, disbelief function, Spohnian belief function, Spohn's rules for belief revision, A, α-conditionalization, λ-conditionalization, rule of combination for disbelief functions, rule of subtraction for disbelief functions, axioms for local computation of marginals*

Address correspondence to Prakash P. Shenoy, School of Business, Summerfield Hall, University of Kansas, Lawrence, KS 66045.

1. INTRODUCTION

This paper is about Spohn's theory of epistemic beliefs [1, 2]. Spohn's theory is an elegant, simple, and powerful calculus designed to represent and reason with plain human beliefs.

The motivation behind Spohn's theory is the need for (1) a formalism to represent plain epistemic belief and (2) procedures for revising beliefs when new information is obtained. As Spohn [2, p. 315] notes, probability theory is inadequate for this purpose for several reasons. *I believe A is true* cannot be represented by $P(A) = 1$ because a probability of 1 is incorrigible, that is, $P(A \mid B) = 1$ for all B such that $P(A \mid B)$ is well defined. However, plain belief is clearly corrigible. I may believe it is snowing outside but when I look out the window and observe that it has stopped snowing, I now believe that it is not snowing outside. Nor can we represent *I believe A is true* with $P(A) \geq 1 - \epsilon$, say, where $\epsilon \geq 0$ is some small number, because, according to the notion of plain belief, if I believe A is true and I believe B is true, then I believe A and B is true. However, if $P(A) \geq 1 - \epsilon$ and $P(B) \geq 1 - \epsilon$, then it is not necessary that $P(A$ and $B) \geq 1 - \epsilon$. The Dempster–Shafer theory of belief functions [3, 4] suffers from the same problems. Once we assume $\text{Bel}(A) = 1$, no further evidence will reduce belief in A to less than 1.

The main ingredients of Spohn's theory are (1) a functional representation of an epistemic state called a natural (or ordinal) conditional function and (2) a rule for revising this function in light of new information. Since the values of a natural conditional function represent degrees of disbeliefs, we call such a function a *disbelief function*. Like a probability distribution function, a disbelief function for a variable is completely specified by its values for the singleton subsets of configurations of the variable. Spohn [2, pp. 318–320] has interpreted the values of a disbelief function as infinitesimal probabilities (see also Pearl [5]). Smets (private communication) and Dubois and Prade [6] have pointed out that a disbelief function can be interpreted as the negative of the logarithm of a possibility function as studied by, for example, Zadeh [7] and Dubois and Prade [8].

The main contribution of this paper is as follows. First, we provide an axiomatic definition of a consistent epistemic state. Some of the axioms we propose are different from the ones proposed by Spohn. Our axioms are a little easier to understand, but we show that the two sets of axioms are mathematically equivalent. These axioms are also found in Gardenfors [9].

Second, we study some properties of an alternative functional representation of a consistent epistemic state called a Spohnian belief function. Although the definition is stated in Ref. 1, there is not much else in there about this function. We state several properties of belief functions. When compared to disbelief functions, belief functions are easier to interpret. But since they contain redundant information, they are harder to manipulate mathematically.

Third, we state a rule of combination for disbelief functions that is mathematically equivalent to Spohn's belief revision rule. Whereas Spohn's rule is defined in terms of the initial epistemic state and some features of the final epistemic state, the rule of combination is defined in terms of the initial epistemic state and the incremental epistemic state representing the information gained. The rule of combination for disbelief functions is pointwise addition.

Fourth, we state a rule of subtraction that always allows one to recover the addendum epistemic state from the initial and final epistemic states. This rule is useful in cases where information gained is expressed in terms of the final epistemic state, as it allows us to recover the addendum epistemic state that is combined with the initial to obtain the final. The subtraction rule is also useful for nonmonotonic reasoning when it becomes necessary to retract the conclusion of an earlier inference without influencing conclusions drawn using other means. The rule of subtraction for disbelief functions is pointwise subtraction.

Fifth, we study some properties of our rule of combination. Spohn's belief revision rule is formulated to revise an epistemic state in light of new information. On the other hand, the rule of combination described in this paper is more flexible. It also can be used to describe an initial epistemic state for many variables when this information is provided as several independent epistemic states each involving few variables. The initial epistemic state for all variables is then obtained by combining these independent epistemic states using the rule of combination. This is a distinct advantage of our reformulation of Spohn's belief revision rule.

Another advantage of our reformulation is that we can show that Spohn's theory of epistemic beliefs shares the essential abstract features of probability theory and the Dempster–Shafer theory of belief functions as described by Shenoy and Shafer [10, 11]. These features are (1) a functional representation of knowledge (or beliefs), (2) a rule of marginalization, and (3) a rule of combination. In Shenoy and Shafer [10, 11], we also state three axioms for the marginalization and combination rules that enable one use local computation in the calculation of the marginals of a joint function without explicitly having to compute the joint function. In this paper we show that rules of marginalization and combination for disbelief functions satisfy the required three axioms. One implication of this is that we have a ready-made algorithm for propagating disbelief functions using only local computation as described in detail in Shenoy and Shafer [11].

An outline of the remainder of this paper is as follows. In Section 2, we provide an axiomatic definition of a consistent epistemic state and study some of its properties.

In Section 3, we introduce the notation. We use the multivariate framework and restrict ourselves to the finite case: the number of variables is finite, and each variable has a finite frame. Our motivation here is easier comprehension.

In Section 4, we define disbelief functions and the marginalization operation. Most of the material in this section is due to Spohn.

In Section 5, we define and study some properties of an alternative representation of a consistent epistemic state that we call a Spohnian belief function.

In Section 6, we state and describe Spohn's rules for revision of beliefs. Spohn has described two rules for belief revision called A, α-conditionalization and λ-conditionalization. A, α-conditionalization is a special case of λ-conditionalization. Most of the material in this section is due to Spohn.

In Section 7, we state a rule of combination for disbelief functions and then show that Spohn's A, α-conditionalization rule for belief revision can be expressed in terms of the rule of combination.

In Section 8, first, we state a rule of subtraction for disbelief functions. Second, given the final and initial disbelief functions, we show that we can always subtract the latter from the former to recover the disbelief function that was added to the initial to obtain the final. Third, we show that Spohn's λ-conditionalization rule for belief revision and our rule of combination are mathematically equivalent.

In Section 9, first, we state some elementary properties of the rule of combination. Second, we illustrate how an initial disbelief function for many variables can be constructed by combining independent disbelief functions each of which involves only a few variables. We sketch what we mean by independent disbelief functions. Third, we show that the rules for marginalizing and combining disbelief functions satisfy the three axioms stated by Shenoy and Shafer [10, 11] for computing marginals using local computation.

Finally, in Section 10, we conclude with a short discussion of what we have accomplished in this paper.

Some terminology we use in this paper is different from that used by Spohn [1, 2]. In reference to an epistemic state, what Spohn calls ''consistent and deductively closed'' we simply call ''consistent,'' and what Spohn calls ''net content'' we simply call ''content.'' Regarding functional representations of an epistemic state, what Spohn calls ''a natural conditional function'' we simply call ''a disbelief function,'' and the function whose values are referred to by Spohn as ''degrees of firmness of belief'' we simply call ''a Spohnian belief function.''

Finally, we would like to mention that there is more to Spohn's theory than that discussed here. Analogous to the concept of conditional independence in probability theory, Spohn defines conditional independence with respect to disbelief functions and states many results regarding this concept (see also Hunter [12]). We strongly recommend that the reader read Spohn [1] to appreciate first-hand the elegance, simplicity, and power of Spohn's theory.

2. CONSISTENT EPISTEMIC STATES

In this section, we will define axiomatically a consistent epistemic state. Some of our axioms are different from those used by Spohn, but the two sets of

axioms are equivalent. Next, we describe a characterization of a consistent epistemic state due to Spohn.

Consider a variable X. Let $\mathcal{W}_X$ denote a finite set of possible values of X such that exactly one is true. We shall call $\mathcal{W}_X$ a *frame* for X. Assume further that $\mathcal{W}_X$ is defined such that the propositions regarding X that are of interest are precisely those of the form "The true value of X is in A" where A is a subset of $\mathcal{W}_X$. Thus the propositions regarding X that are of interest are in a one-to-one correspondence with the subsets of $\mathcal{W}_X$ [4, p. 36].

The correspondence between subsets and propositions is useful, because it translates the logical notions of conjunction, disjunction, implication, and negation into the set-theoretic notions of intersection, union, inclusion, and complementation [4, pp. 36–37]. Thus, if A and B are two subsets of $\mathcal{W}_X$, and A' and B' are the corresponding propositions, then $A \cap B$ corresponds to the conjunction of A' and B', $A \cup B$ corresponds to the disjunction of A' and B', $A \subseteq B$ if and only if A' implies B', and A is the set-theoretic complement of B with respect to $\mathcal{W}_X$ (written as $A = \sim B$) if and only if A' is the negation of B'. Notice also that the proposition that corresponds to $\emptyset$ is false, and the proposition that corresponds to $\mathcal{W}_X$ is true. If A is a proper, nonempty subset of $\mathcal{W}_X$, we shall refer to the proposition that corresponds to A as *contingent*. Henceforth, we will simply refer to a proposition by its corresponding subset. The set of subsets of $\mathcal{W}_X$ will be denoted by $2^{\mathcal{W}_X}$.

In an epistemic state for X, some propositions are believed to be true (or simply, *believed*), some are believed to be false (or simply, *disbelieved*), and the remainder are neither believed nor disbelieved. Logical consistency requires that these beliefs satisfy certain conditions (axioms). A definition of a *consistent* epistemic state is as follows.

DEFINITION 1   *An epistemic state is said to be consistent if the following five axioms are satisfied:*

> **A1.** For any propositions A, exactly one of the following conditions holds:
>> (i) A is believed;
>> (ii) A is disbelieved;
>> (iii) A is neither believed nor disbelieved.
>
> **A2.** $\mathcal{W}_X$ is (always) believed.
> **A3.** A is believed if and only if $\sim A$ is disbelieved.
> **A4.** If A is believed and $B \supseteq A$, then B is believed.
> **A5.** If A and B are believed, then $A \cap B$ is believed.

Some simple consequences of Definition 1 are as follows.

PROPOSITION 1   *The following conditions always hold in any consistent epistemic state:*

> **A6.** $\emptyset$ is (always) disbelieved.
> **A7.** If A is disbelieved and $B \subseteq A$, then B is disbelieved.
> **A8.** If A and B are disbelieved, then $A \cup B$ is disbelieved.

A9. If A is not believed, then this does not necessarily imply that $\sim A$ must be believed; it is possible that $\sim A$ is also not believed. And if A is not disbelieved, then this does not necessarily imply that $\sim A$ must be disbelieved; it is possible that $\sim A$ is also not disbelieved.

A10. If $\mathcal{B}$ denotes the set of all believed propositions, then $\cap \mathcal{B} \neq \varnothing$.

A11. If $\mathcal{B}$ denotes the set of all believed propositions, then $A \in \mathcal{B}$ whenever $A \supseteq (\cap \mathcal{B}')$ for some $\mathcal{B}' \subseteq \mathcal{B}$.

Proof A6 follows from A2 and A3. A7 follows from A4 and A3. A8 follows from A5 and A3.

To show that A9 is true, consider an epistemic state in which the only proposition that is believed is $\mathcal{W}$, the only proposition that is disbelieved is $\varnothing$, and all other propositions are neither believed nor disbelieved. Clearly, this epistemic state satisfies axioms A1–A5.

To show A10, note that by repeated application of A5, $(\cap \mathcal{B}) \in \mathcal{B}$. And from A6, it follows that $\cap \mathcal{B} \neq \varnothing$.

Finally, to show A11, note that by repeated application of A5, $(\cap \mathcal{B}') \in \mathcal{B}$, and since $A \supseteq (\cap \mathcal{B}')$, it follows from A4 that $A \in \mathcal{B}$.

It is clear from axioms A1 and A3 that we can specify a consistent epistemic state simply by listing all propositions that are believed. Then axiom A3 tells us exactly which propositions are disbelieved. And from axiom A1, the remaining propositions are neither believed nor disbelieved.

Let $\mathcal{B}$ denote the set of all propositions (subsets) that are believed, and let $\mathcal{D}$ denote the set of all propositions that are disbelieved in some consistent epistemic state. Theorem 1 gives a characterization of $\mathcal{B}$, and its corollary gives a characterization of $\mathcal{D}$.

THEOREM 1 *Suppose $\mathcal{B}$ denotes the set of all propositions that are believed in some epistemic state for X. Then the epistemic state is consistent if and only if there exists a unique nonempty subset C of $\mathcal{W}_X$ such that*

$$\mathcal{B} = \left\{ A \in 2^{\mathcal{W}_X} \mid A \supseteq C \right\}$$

Proof (*Sufficiency*) Assume that the epistemic state is consistent. Consider the proposition $\cap \mathcal{B}$. It follows from A10 that $\cap \mathcal{B} \neq \varnothing$. The sufficiency part of the proof follows by letting $C = \cap \mathcal{B}$.

(*Necessity*). Suppose $\mathcal{B} = \{ A \in 2^{\mathcal{W}_X} \mid A \supseteq C \}$ for some nonempty subset C of $\mathcal{W}_X$. We will show that the epistemic state satisfies axioms A1–A5. First note that A1 and A3 are satisfied by definition. To show A2, note that $\mathcal{W}_X \supseteq C$. Therefore, $\mathcal{W}_X \in \mathcal{B}$, that is, $\mathcal{W}_X$ is believed. To show A4, suppose that A is believed and $B \supseteq A$. Since A is believed, $A \in \mathcal{B}$. Therefore, $A \supseteq C$. Since $B \supseteq A$, $B \supseteq C$. Therefore, $B \in \mathcal{B}$, that is, B is believed. To

show A5, suppose that A and B are believed, that is, A and B belong to $\mathscr{B}$. By hypothesis, $A \supseteq C$ and $B \supseteq C$. Hence, $A \cap B \supseteq C$. Therefore, $(A \cap B) \in \mathscr{B}$, that is, $A \cap B$ is believed.

COROLLARY TO THEOREM 1   *Suppose $\mathscr{D}$ denotes the set of all propositions that are disbelieved in some epistemic state for X. Then the epistemic state is consistent if and only if there exists a unique proper subset D of $\mathscr{W}_X$ such that*

$$\mathscr{D} = \left\{ A \in 2^{*X} \mid A \subseteq D \right\}$$

Proof   The proof follows from Theorem 1 and axiom A3. Note that $D = \sim C$.

The characterization of $\mathscr{B}$ in Theorem 1 (but not Theorem 1 itself) is due to Spohn [1, p. 108]. Spohn defines a consistent epistemic state as one that satisfies A1, A3, A10, and A11, and consequently Theorem 1 follows trivially from A10 and A11. It follows from Theorem 1 that our definition of a consistent epistemic state (axioms A1–A5) is equivalent to Spohn's (A1, A3, A10, and A11).

Subset C in Theorem 1 is called the *content* of the consistent epistemic state. Note that the content constitutes a complete specification of an epistemic state. Thus another simple corollary of Theorem 1 is that in a frame $\mathscr{W}_X$ consisting of n elements, there are exactly $2^n - 1$ distinct possible consistent epistemic states (corresponding to each nonempty subset of $\mathscr{W}$ as the content). The content C represents the smallest believed proposition, and $D = \sim C$ represents the largest disbelieved proposition.

EXAMPLE 1 (Consistent epistemic states)   Consider a frame $\mathscr{W} = \{ x, y, z \}$. Table 1 lists all seven possible consistent epistemic states.

Epistemic states 1, 2, and 3 represent states of *complete* beliefs; each proposition is either believed or disbelieved. States 4–7 represent incomplete beliefs. Epistemic state 7 represents a state of complete ignorance in which nothing is believed or disbelieved (except, of course, the frame, which is always believed, and the empty set, which is always disbelieved).

3. VARIABLES, CONFIGURATIONS, AND PROPOSITIONS

In this section, we introduce the notation that will be used in the rest of the paper. We use the multivariate framework because, even though it does not generalize readily to the continuous case, it is more intuitive and easier to comprehend than the measurable-subset framework used by Spohn [1, 2].

Consider a variable X. The symbol $\mathscr{W}_X$ denotes the set of possible values of

Table 1. The Seven Possible Consistent Epistemic States for a Variable with Frame $\mathscr{W} = \{x, y, z\}$

State i	Content $\mathscr{C}$	Believed $\mathscr{B}$	Disbelieved $\mathscr{D}$	Neither believed nor disbelieved
1	$\{x\}$	$\{x\}, \{x,y\}, \{x,z\}, \mathscr{W}$	$\varnothing, \{y\}, \{z\}, \{y,z\}$	None
2	$\{y\}$	$\{y\}, \{x,y\}, \{y,z\}, \mathscr{W}$	$\varnothing, \{x\}, \{z\}, \{x,z\}$	None
3	$\{z\}$	$\{z\}, \{x,z\}, \{y,z\}, \mathscr{W}$	$\varnothing, \{x\}, \{y\}, \{x,y\}$	None
4	$\{x,y\}$	$\{x,y\}, \mathscr{W}$	$\varnothing, \{z\}$	$\{x\}, \{y\}, \{x,z\}, \{y,z\}$
5	$\{x,z\}$	$\{x,z\}, \mathscr{W}$	$\varnothing, \{y\}$	$\{x\}, \{z\}, \{x,y\}, \{y,z\}$
6	$\{y,z\}$	$\{y,z\}, \mathscr{W}$	$\varnothing, \{x\}$	$\{y\}, \{z\}, \{x,y\}, \{x,z\}$
7	$\mathscr{W}$	$\mathscr{W}$	$\varnothing$	$\{x\}, \{y\}, \{z\}, \{x,y\}, \{x,z\}, \{y,z\}\ \{x,z\}$

X. We assume that one and only one of the elements of $\mathscr{W}_X$ can be the true value of X. We call $\mathscr{W}_X$ the *frame for X*. For example, suppose we are interested in determining whether a person is a pacifist or not. We can construct a variable P whose frame has two elements: p (for pacifist) and $\sim p$ (not pacifist).

Let $\mathscr{X}$ denote the set of all variables. In this paper we will be concerned only with the case where $\mathscr{X}$ is finite. We will also assume that all the variables in $\mathscr{X}$ have finite frames.

We will often deal with nonempty subsets of variables in $\mathscr{X}$. Given a nonempty subset g of $\mathscr{X}$, let $\mathscr{W}_g$ denote the Cartesian product of $\mathscr{W}_X$ for X in g, that is,

$$\mathscr{W}_g = \times \{ \mathscr{W}_X \mid X \in g \}$$

We can think of the set $\mathscr{W}_g$ as the set of possible values of the joint variable g. Accordingly, we call $\mathscr{W}_g$ the *frame for g*. Also, we will refer to elements of $\mathscr{W}_g$ as *configurations of g*. We will use this terminology even when g consists of a single variable. Thus we will refer to elements of $\mathscr{W}_X$ as *configurations of X*. We will use lower case, boldface letters such as $\mathbf{x}$, $\mathbf{y}$ to denote configurations. Also, if $\mathbf{x}$ is a configuration of g and $\mathbf{y}$ is a configuration of h, and $g \cap h = \varnothing$, then $(\mathbf{x},\mathbf{y})$ will denote a configuration of $g \cup h$.

Projection of configurations simply means dropping extra coordinates; if $(r, \sim q, \sim p)$ is a configuration of $\{R,Q,P\}$, for example, then the projection of $(r, \sim q, \sim p)$ to $\{R,P\}$ is simply $(r, \sim p)$, which is a configuration of $\{R,P\}$. If g and h are sets of variables, $h \subseteq g$, and $\mathbf{x}$ is a configuration of g, then we will let $\mathbf{x}^{\downarrow h}$ denote the projection of $\mathbf{x}$ to h.

By extension of a subset of a frame to a subset of a larger frame, we mean a cylinder set extension. If g and h are sets of variables, $h \subseteq g$, and A is a subset of $\mathscr{W}_h$, then the *extension of A to g* is $A \times \mathscr{W}_{g-h}$. We will let $A^{\uparrow g}$ denote the extension of A to g. For example, consider three variables R,P,Q with frames $\mathscr{W}_R = \{r, \sim r\}$, $\mathscr{W}_P = \{p, \sim p\}$, and $\mathscr{W}_Q = \{q, \sim q\}$, respectively. Then the extension of $\{(r, \sim p), (\sim r, p)\}$ (which is a subset of $\mathscr{W}_{\{R,P\}}$) to $\{R,P,Q\}$ is

$$\{ (r, \sim p, q), (r, \sim p, \sim q), (\sim r, p, q), (\sim r, p, \sim q) \}$$

Note that the propositions corresponding to A and $A^{\uparrow g}$ are logically equivalent.

We will denote the set of all natural numbers by $\mathbb{N}$ and the set of integers by $\mathbb{Z}$; $\mathbb{N} = \{0, 1, 2, \cdots\}$ and $\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$. The extended set of natural numbers consists of the set of all natural numbers to which a symbol, $+\infty$, has been added with the following properties. If $x \in \mathbb{N}$, then (i), $x < +\infty$ and (ii) $x + +\infty = +\infty$. The extended natural number set is denoted by $\mathbb{N}^+$. Similarly, we will let $\mathbb{Z}^+$ denote the extended set of integers

consisting of set of all integers to which two symbols $+\infty$ and $-\infty$ have been added with the following properties. If $x \in \mathbb{Z}$, then (i) $-\infty < x < +\infty$, (ii) $x + +\infty = +\infty$, and (iii) $x + -\infty = -\infty$.

4. DISBELIEF FUNCTIONS

The basic functional representation of an epistemic state in Spohn's theory is called an ordinal conditional function in Spohn [1, p. 115] and a natural conditional function in Spohn [2, p. 316]. We simply call this function a disbelief function.

DEFINITION 2 [2, P. 316]   *A disbelief function for g is a function* $\delta : 2^{*g} \rightarrow \mathbb{N}^+$ *such that*
   D0. $\delta(\{w\}) \in \mathbb{N}$ for all $w \in \mathcal{W}_g$;
   D1. There exists a configuration $w \in \mathcal{W}_g$ for which $\delta(\{w\}) = 0$;
   D2. For any $A \in (2^{*g} - \{\varnothing\})$, $\delta(A) = \text{MIN}\{\delta(\{w\}) \,|\, w \in A\}$; and
   D3. $\delta(\varnothing) = +\infty$.

Note that although a disbelief function is defined for the set of all subsets of $\mathcal{W}_g$, it is completely specified by its values for each singleton subset of $\mathcal{W}_g$.

A disbelief function is a complete representation of a consistent epistemic state. To see what propositions are believed in state δ, consider the subset $C = \{w \in \mathcal{W}_g \,|\, \delta(\{w\}) = 0\}$. By condition D1 in Definition 2, C is always nonempty. C represents the content of the epistemic state; that is, A is believed in state δ iff $A \supseteq C$. Thus A is believed in state δ iff $\delta(\sim A) > 0$; A is disbelieved in state δ iff $\delta(A) > 0$; and A is neither believed nor disbelieved in state δ iff $\delta(A) = \delta(\sim A) = 0$.

A disbelief function consists of more than a representation of a consistent epistemic state. It also includes degrees of belief and disbelief. If $\delta(A) > 0$, then $\delta(A)$ can be interpreted as the degree of disbelief in proposition A; that is, A is more disbelieved than B if $\delta(A) > \delta(B) > 0$. And if $\delta(\sim A) > 0$, then $\delta(\sim A)$ can be interpreted as the degree of belief for A, that is, A is more believed than B if $\delta(\sim A) > \delta(\sim B) > 0$.

Consider the following disbelief function for g: $\delta(\{w\}) = 0$ for all $w \in \mathcal{W}_g$. This means that the only proposition that is disbelieved is $\varnothing$ and the only proposition that is believed is $\mathcal{W}_g$. We shall call such a disbelief function *vacuous*. It represents a state of complete ignorance.

PROPOSITION 2 [1, P. 115]   *Suppose δ is a disbelief function for g. Then*
   D4. For each $A \in 2^{*g}$, either $\delta(A) = 0$ or $\delta(\sim A) = 0$ or both.
   D5. For each $A, B \in 2^{*g}$, $\delta(A \cup B) = \text{MIN}\{\delta(A), \delta(B)\}$.

Proof   D4 follows from condition D1 in Definition 2, and D5 follows from condition D2 in Definition 2.

Marginalization

Suppose δ is a disbelief function for g. Suppose $h \subseteq g$. We may be interested only in propositions regarding variables in h. In this case, we would like to marginalize δ to h. The following definition of marginalization is motivated by the fact that each proposition $A \in 2^{\mathscr{W}h}$ about variables in h can be regarded as a proposition $A^{\uparrow g} \in 2^{\mathscr{W}g}$ about variables in g.

DEFINITION 3   *Suppose δ is a disbelief function for g, and suppose $h \subseteq g$. The marginal of δ for h, denoted by $\delta^{\downarrow h}$, is a disbelief function for h given as follows*:

$$\delta^{\downarrow h}(A) = \delta(A^{\uparrow g}) = \mathrm{MIN}\left\{\delta(\{(x,y)\}) \,|\, x \in A, y \in \mathscr{W}_{g-h}\right\} \qquad (1)$$

for all $A \in 2^{\mathscr{W}h}$.

In particular, if A is a singleton subset, that is, $A = \{x\}$ for some $x \in \mathscr{W}_h$, then (1) simplifies to

$$\delta^{\downarrow h}(\{x\}) = \mathrm{MIN}\left\{\delta(\{(x,y)\}) \,|\, y \in \mathscr{W}_{g-h}\right\}.$$

Note that if δ is a disbelief function for g and $k \subseteq h \subseteq g$, then $(\delta^{\downarrow h})^{\downarrow k} = \delta^{\downarrow k}$. In words, if we regard marginalization as reduction of δ by deletion of variables, then the order in which the variables are deleted makes no difference in the final answer.

EXAMPLE 2 (Disbelief function and marginalization)   We would like to determine whether a stranger (about whom we know nothing) is a pacifist or not depending on whether or not he is a Republican and whether or not he is a Quaker. Consider three variables R, Q, and P. R has two possible values: r (for Republican), $\sim r$ (not Republican); Q has two values: q (Quaker) and $\sim q$ (not Quaker); and P has two possible values: p (pacifist) and $\sim p$ (not pacifist). Our belief that most Republicans are not pacifists and that most Quakers are pacifists is represented by the disbelief function δ for $\{R, Q, P\}$ shown in Table 2 (the construction of this disbelief function will be explained in Section 9).

Note that the marginal of δ for R is the vacuous disbelief function for R,

$$\delta^{\downarrow\{R\}}(\{r\}) = \delta^{\downarrow\{R\}}(\{\sim r\}) = 0.$$

Thus in epistemic state δ, we neither believe he is a Republican nor that he is not a Republican. Similarly, note that the marginals of δ for $\{Q\}$ and $\{P\}$ are also vacuous. The marginal of δ for $\{R, P\}$ is as follows:

$$\delta^{\downarrow\{R,P\}}(\{r,p\}) = 1$$
$$\delta^{\downarrow\{R,P\}}(\{r,\sim p\}) = \delta^{\downarrow\{R,P\}}(\{\sim r,p\}) = \delta^{\downarrow\{R,P\}}(\{\sim r,\sim p\}) = 0$$

Table 2.   A Disbelief Function for $(R, Q, P\}$

w	$\delta(\{w\})$
(r, q, p)	1
$(r, q, \sim p)$	2
$(r, \sim q, p)$	1
$(r, \sim q, \sim p)$	0
$(\sim r, q, p)$	0
$(\sim r, q, \sim p)$	2
$(\sim r, \sim q, p)$	0
$(\sim r, \sim q, \sim p)$	0

Thus we disbelieve pacifist Republicans (to degree 1). Similarly, note that the marginal of δ for $\{Q, P\}$ is as follows:

$$\delta^{\downarrow\{Q,P\}}(\{q, \sim p\}) = 2$$

$$\delta^{\downarrow\{Q,P\}}(\{q, p\}) = \delta^{\downarrow\{Q,P\}}(\{\sim q, p\}) = \delta^{\downarrow\{Q,P\}}(\{\sim q, \sim p\}) = 0$$

Thus we disbelieve pacifist Republicans (to degree 1). Similarly, note that the marginal of δ for $\{R, Q\}$ is as follows:

$$\delta^{\downarrow\{R,Q\}}(\{r, q\}) = 1$$

$$\delta^{\downarrow\{R,Q\}}(\{r, \sim q\}) = \delta^{\downarrow\{R,Q\}}(\{\sim r, q\}) = \delta^{\downarrow\{R,Q\}}(\{\sim r, \sim q\}) = 0$$

Thus we disbelieve republican Quakers (to degree 1).

5. SPOHNIAN BELIEF FUNCTIONS

As we saw in the previous section, a disbelief function models degrees of disbelief for disbelieved propositions directly whereas it models degrees of belief for believed propositions only indirectly; if A is a believed proposition, then the degree of belief for A is $\delta(\sim A)$. Can we model both beliefs and disbeliefs directly? The answer is yes. We call such a representation a (Spohnian) belief function. However, as we shall see, belief functions are not as easy to manipulate mathematically as disbelief functions. This is because a belief function has redundant information: If we believe proposition A to degree α, then we must disbelieve $\sim A$ to degree α. Including both these statements in the definition of a disbelief function is the main cause of the difficulty. Nevertheless, the ease of interpretation of a belief function makes its study worthwhile.

DEFINITION 4 [1, p. 116]   *A (Spohnian) belief function for g is a*

Table 3. The Belief Function β Corresponding to Disbelief Function δ

w	$\delta(\{w\})$	$\beta(\{w\})$
(r,q,p)	1	-1
$(r,q,\sim p)$	2	-2
$(r,\sim q,p)$	1	-1
$(r,\sim q,\sim p)$	0	0
$(\sim r,q,p)$	0	0
$(\sim r,q,\sim p)$	2	-2
$(\sim r,\sim q,p)$	0	0
$(\sim r,\sim q,\sim p)$	0	0

function $\beta:2^{*_g} \to \mathbb{Z}^+$ such that

$$\beta(A) = \begin{cases} -\delta(A) & \text{if } \delta(A) > 0 \\ \delta(\sim A) & \text{if } \delta(A) = 0 \end{cases}$$

for all $A \in 2^{*_g}$ where δ is some disbelief function for g.

The number $\beta(A)$ is interpreted as the degree of belief in proposition A. A is believed iff $\beta(A) > 0$, A is disbelieved iff $\beta(A) < 0$, and A is neither believed nor disbelieved iff $\beta(A) = 0$. Furthermore, if $\beta(A) > \beta(B) > 0$, then A is more believed than B, and if $\beta(A) < \beta(B) < 0$, then A is more disbelieved than B. Propositions $\varnothing$ and $\mathcal{W}_g$ are the extreme cases. $\varnothing$ is the most disbelieved proposition $[\beta(\varnothing) = -\infty]$, and $\mathcal{W}_g$ is the most believed proposition $[\beta(\mathcal{W}_g) = +\infty]$.

EXAMPLE 3 (Spohnian belief function)   Consider the disbelief function δ for $\{R,Q,P\}$ from Example 2. Table 3 shows the corresponding belief function β.

The disbelief function that produces a given belief function is unique and can be recovered from the belief function.

THEOREM 2   *Suppose β is the belief function for g given by the disbelief function δ. Then*

$$\delta(A) = \begin{cases} 0 & \text{if } \beta(A) \geq 0 \\ -\beta(A) & \text{if } \beta(A) < 0 \end{cases}$$

for all $A \in 2^{*_g}$.

Proof   Suppose A is such that $\beta(A) > 0$. Then from the definition of β, it must be true that $\beta(A) = \delta(\sim A)$. Therefore from condition D4, $\delta(A) = 0$. If A is such that $\beta(A) < 0$, then from the definition of β, $\delta(A) > 0$ and

$\beta(A) = -\delta(A)$, that is, $\delta(A) = -\beta(A)$. If A is such that $\beta(A) = 0$, then from the definition of β, $\delta(A) = 0$.

Theorems 3, 4, and 5 describe some properties of belief functions. Theorem 3 confirms that a belief function does contain redundant information.

THEOREM 3 *Suppose β is a belief function for g. Then for any proposition $A \subseteq \mathcal{W}_g$, $\beta(A) = -\beta(\sim A)$.*

Proof Let δ be the disbelief function corresponding to β. First consider the case where $\delta(A) > 0$. In this case, it follows from condition D4 in Proposition 2 that $\delta(\sim A) = 0$. From the definition of a belief function, $\beta(A) = -\delta(A)$, and $\beta(\sim A) = \delta(\sim (\sim A)) = \delta(A)$. Therefore the result follows. The proof is similar for the case $\delta(A) = 0$.

In Definition 4, we defined a belief function in terms of a disbelief function. Theorem 4 shows that the class of belief functions can be characterized without reference to disbelief functions, and in doing so, one can see why belief functions are not as easy to work with as disbelief functions.

THEOREM 4 *A function $\beta : 2^{\mathcal{W}_g} \to \mathbb{Z}^+$ is a belief function if and only if there exists a proper subset D of $\mathcal{W}_g$ such that*
  B0. $\beta(\{w\}) \in \mathbb{Z}$ for all $w \in \mathcal{W}_g$.
  B1. $\beta(\{w\}) < 0$ for all $w \in D$
  B2. $\beta(A)$
$$
= \begin{cases} (-1)\,\text{MIN}\,\{-\beta(\{w\}) \mid w \in A\} & \text{if } A \subseteq D \\ \text{MIN}\,\{-\beta(\{w\}) \mid w \in \sim A\} & \text{if } A \supseteq \sim D, \text{i.e., } \sim A \subseteq D \\ 0 & \text{otherwise} \end{cases}
$$
   for all $A \in (2^{\mathcal{W}_g} - \{\varnothing, \mathcal{W}_g\})$.
  B3. $\beta(\varnothing) = -\infty$, and $\beta(\mathcal{W}_g) = +\infty$.

Proof (Sufficiency) Let β be a function satisfying conditions B0–B3. To show that β is a belief function, we need to construct a disbelief function δ such that $\beta(A) = -\delta(A)$ if $\delta(A) > 0$ and $\beta(A) = \delta(\sim A)$ if $\delta(A) = 0$. Let δ be a disbelief function defined as follows:

$$\delta(\{w\}) = -\beta(\{w\}) \text{ if } w \in D \qquad \text{and } \delta(\{w\}) = 0 \text{ if } w \in \sim D.$$

Since D is a proper subset of $\mathcal{W}_g$, from B0 and B1, clearly δ is a disbelief function.

Suppose $\varnothing \neq A \subseteq D$. This means that $\delta(\{w\}) = -\beta(\{w\}) > 0$ for all $w \in A$. Therefore,
$$\delta(A) = \text{MIN}\,\{\delta(\{w\}) \mid w \in A\} > 0.$$
By condition B2,
$$\beta(A) = (-1)\,\text{MIN}\,\{-\beta(\{w\}) \mid w \in A\}$$
$$= (-1)\,\text{MIN}\,\{\delta(\{w\}) \mid w \in A\} = -\delta(A).$$

Now suppose $A \neq \mathscr{W}_g$ and $A \supseteq \sim D$, that is, $\varnothing \neq \sim A \subseteq D$. In this case $\delta(A) = 0$, and from B1, $\delta(\{w\}) = -\beta(\{w\})$ for all $w \in \sim A$. From B2,

$$\beta(A) = \text{MIN} \{ -\beta(\{w\}) \mid w \in \sim A \} = \text{MIN} \{ \delta(\{w\}) \mid w \in \sim A \} = \delta(\sim A).$$

Finally, suppose that $A \in (2^{\mathscr{W}_g} - \{\varnothing, \mathscr{W}_g\})$ such that A is neither a subset of D nor contains $\sim D$. Since A is not a subset of D, $\delta(A) = 0$, and since A does not contain $\sim D$, $\delta(-A) = 0$. From B2, $\beta(A) = 0 = \delta(\sim A)$.

Since $\delta(\varnothing) = +\infty > 0$, we need to show that $\beta(\varnothing) = -\delta(\varnothing) = -\infty$. This is true from B3. Since $\delta(\mathscr{W}_g) = 0$, we need to show that $\beta(\mathscr{W}_g) = \delta(\sim \mathscr{W}_g) = \delta(\varnothing) = +\infty$. This is also true from B3. This completes the sufficiency part of the proof.

(Necessity) Suppose β is a belief function corresponding to some disbelief function δ. Condition B0 follows from condition D0 and Definition 4. Condition B1 is obvious from Definition 4 if we let $D = \{w \in \mathscr{W}_g \mid \delta(\{w\}) > 0\}$.

The proof of B2 is similar to the proof in the sufficiency part. And B3 is obvious from condition D3 and Definition 4.

Note that the proper subset D is the complement of the content of the consistent epistemic state represented by β. Thus a belief function is specified completely by its values for each configuration in D. Furthermore, from condition B2, it follows that a proposition is believed if and only if it contains the content, is disbelieved if and only if it is a subset of D, and is neither believed nor disbelieved otherwise. This is consistent with Theorem 1 and its corollary.

The following theorem describes marginalization for belief functions.

THEOREM 5   *Suppose δ is a disbelief function for g, and suppose $h \subseteq g$. Let β and $\beta^{\downarrow h}$ be belief functions for g and h, respectively, corresponding to disbelief functions δ and $\delta^{\downarrow h}$, respectively. Then*

$$\beta^{\downarrow h}(\{x\}) = \beta(\{x\}^{\uparrow g}) \text{ for all } x \in \mathscr{W}_h$$

Proof   First suppose $x \in \mathscr{W}_h$ is such that $\delta^{\downarrow h}(\{x\}) > 0$. Then $\beta^{\downarrow h}(\{x\}) = -\delta^{\downarrow h}(\{x\})$. Also, since $\delta(\{x\}^{\uparrow g}) = \delta^{\downarrow h}(\{x\})$ by definition of marginalization of disbelief functions,

$$\beta(\{x\}^{\uparrow g}) = -\delta(\{x\}^{\uparrow g}) = -\delta^{\downarrow h}(\{x\}) = \beta^{\downarrow h}(\{x\})$$

Next consider the case $x \in \mathscr{W}_h$ such that $\delta^{\downarrow h}(\{x\}) = 0$. Then,

$$\beta^{\downarrow h}(\{x\}) = \delta^{\downarrow h}(\sim \{x\}) = \text{MIN} \{ \delta^{\downarrow h}(\{y\}) \mid y \in \mathscr{W}_h - \{x\} \}$$

$$= \text{MIN} \{ \delta(\{y\}^{\uparrow g}) \mid y \in \mathscr{W}_h - \{x\} \} = \delta(\sim (\{x\}^{\uparrow g})) = \beta(\{x\}^{\uparrow g})$$

6. REVISION OF DISBELIEFS

In this section, first we state Spohn's A,α-conditionalization rule for modifying a disbelief function in light of new information. Then we describe four properties of this rule. Finally, we describe the general λ-conditionalization rule.

DEFINITION 5 [1, P. 117] *Suppose δ is a disbelief function for g representing our initial epistemic state. Suppose we learn something about contingent proposition A (or $\sim A$) that consequently leads us to believe A to degree α (or, equivalently, disbelieve $\sim A$ to degree α), where $\alpha \in \mathbb{N}$. The resulting epistemic state, called the A,α-conditionalization of δ and denoted by disbelief function $\delta_{A,\alpha}$, is defined as follows:*

$$\delta_{A,\alpha}(\{\mathbf{w}\}) = \begin{cases} \delta(\{\mathbf{w}\}) - \delta(A) & \text{if } \mathbf{w} \in A \\ \delta(\{\mathbf{w}\}) + \alpha - \delta(\sim A) & \text{if } \mathbf{w} \notin A \end{cases}$$

for all $\mathbf{w} \in \mathcal{W}_g$.

Spohn [1] describes four properties of this rule. Let β and $\beta_{A,\alpha}$ denote the belief functions corresponding to δ and $\delta_{A,\alpha}$, respectively. First note that $\delta_{A,\alpha}(A) = 0$ and $\delta_{A,\alpha}(\sim A) = \alpha$. Therefore, $\beta_{A,\alpha}(A) = \alpha$. This means that if $\beta(A) < \alpha$, then what we have learned about A (or $\sim A$) increases our belief in A. However, if $\beta(A) > \alpha$, then what we have learned about A decreases our belief in A. Finally, if $\beta(A) = \alpha$, then the beliefs remain unchanged after revision, that is, $\delta_{A,\alpha} = \delta$ [1, p. 118].

Second, since learning about A (or $\sim A$) does not discriminate between propositions contained in A, the relative degrees of disbelief of these propositions are unchanged; that is, $\delta(E) - \delta(E') = \delta_{A,\alpha}(E) - \delta_{A,\alpha}(E')$ for all E, $E' \subseteq A$. Also, since learning about A does not discriminate between propositions contained in $\sim A$, the relative degrees of disbelief of these propositions are also unchanged. What has changed is that the degrees of disbelief of propositions contained in $\sim A$ have shifted upwards relative to propositions contained in A [1, p. 117].

Before we continue with the properties of the belief revision rule, let us illustrate the rule with an example.

EXAMPLE 4 (A-α-conditionalization rule) Consider the situation in Example 2. Suppose our initial epistemic state is as given by δ in Table 2. Suppose that after a brief conversation with the stranger, we now believe that the person is a Republican to degree 3. In this case, $A = \{(r,q,p), (r,q,\sim p), (r,\sim q,p), (r,\sim q,\sim p)\}$, $\alpha = 3$, $\delta(A) = 0$, and $\delta(\sim A) = 0$. Then the revised disbelief

Table 4. Spohn's Rule for Belief Revision

w	$\delta(\{w\})$	$\delta'(\{w\})$	$\delta''(\{w\})$
(r,q,p)	1	1	0
$(r,q,\sim p)$	2	2	1
$(r,\sim q,p)$	1	1	4
$(r,\sim q,\sim p)$	0	0	3
$(\sim r,q,p)$	0	3	2
$(\sim r,q,\sim p)$	2	5	4
$(\sim r,\sim q,p)$	0	3	6
$(\sim r,\sim q,\sim p)$	0	3	6

function, denoted say by δ', is shown in Table 4. Note that according to our epistemic state δ', we now believe that the person is not a pacifist (to degree 1) and not a Quaker (to degree 1).

Suppose after further conversation with the stranger we now believe that the person is a Quaker to degree 3. In this case, $A = \{(r,q,p), (r,q,\sim p), \sim r,q,p), (\sim r,q,\sim p)\}$, $\alpha = 3$, $\delta'(A) = 1$, and $\delta'(\sim A) = 0$. The revised disbelief function is denoted by δ'' and is also shown in Table 4. According to the epistemic state δ'', we now believe that the person is a pacifist (to degree 1), a Republican (to degree 2), and a Quaker (to degree 3).

Third, A,α-conditionalization is reversible. Suppose δ is a disbelief function with corresponding belief function β, suppose A is a contingent proposition, and suppose $\beta(A) = \gamma$. Then $(\delta_{A,\alpha})_{A,\gamma} = (\delta_{\sim A,\alpha})_{A,\gamma} = \delta$ [1, p. 118]. In words, suppose the initial belief of proposition A is γ in epistemic state δ. Suppose we learn something about A that leads us to believe A to degree α. The resulting epistemic state is given by $\delta_{A,\alpha}$. Suppose that we learn something more about proposition A that has the effect of negating what we learned previously, that is, we now believe A to degree γ again. Then the resulting epistemic state $(\delta_{A,\alpha})_{A,\gamma}$ is the same as the epistemic state δ that we started with initially.

Fourth, A,α-conditionalization is partially commutative. If δ is a disbelief function, and A and B are contingent propositions such that $\delta(A \cap B) = \delta(A \cap \sim B) = \delta(\sim A \cap B) = 0$, then $(\delta_{A,\alpha})_{B,\gamma} = (\delta_{B,\gamma})_{A,\alpha}$ for all α, $\beta \in \mathbb{N}$ [1, p. 118]. Clearly, A,α-conditionalization is not always commutative. Example 5 illustrates a case where belief revision is not commutative, and Example 6 illustrates a case where belief revision is commutative.

EXAMPLE 5 (Noncommutative belief revision) Let S be a variable with frame $\{s, \sim s\}$ where $S = s$ represents that it is snowing outside and $S = \sim s$ represents that it is not snowing outside. Suppose my initial epistemic state is complete ignorance (it is 5:55 a.m., the radio-alarm hasn't turned on yet, and

the curtains are drawn across the window) represented by disbelief function δ such that $\delta(\{s\}) = \delta(\{\sim s\}) = 0$. At 6:00 a.m., the radio turns on and the announcer states that it is now snowing in the city. Here, $A = \{s\}$, and, say, $\alpha = 1$. Then $\delta_{A,\alpha}(\{s\}) = 0$ and $\delta_{A,\alpha}(\{\sim s\}) = 1$. Next, at 6:05, suppose I get up and open the curtains and observe that although there is a blanket of snow on the ground and the wind in blowing hard and moving the snow around, it does not seem like it is actually snowing outside. But it is still dark outside and hard to tell for sure. Here, $B = \{\sim s\}$ and, say, $\gamma = 1$. Then $(\delta_{A,\alpha})_{B,\gamma}(\{s\})$ $= 1$ and $(\delta_{A,\alpha})_{B,\gamma}(\{s\}) = 0$. Note that $\delta_{B,\gamma}(\{s\}) = 1$, $\delta_{B,\gamma}(\{\sim s\}) = 0$, and $(\delta_{B,\gamma})_{A,\alpha}(\{s\}) = 0$, $(\delta_{B,\gamma})_{A,\alpha}(\{\sim s\}) = 1$. Thus, $(\delta_{A,\alpha})_{B,\gamma} \neq (\delta_{B,\gamma})_{A,\alpha}$. Note that A and B do not satisfy the conditions stated above for revision to the commutative. Lack of commutativity in this case is not a problem. It is clear from the story that my belief about S at 6:05 a.m. is accurately represented by $(\delta_{A,\alpha})_{B,\gamma}$ and not by $(\delta_{B,\gamma})_{A,\alpha}$. The disbelief function $(\delta_{B,\gamma})_{A,\alpha}$ is meaningless in the context of the story.

EXAMPLE 6 (Commutative belief revision) Suppose G is a variable with frame $\{g_x, g_y, g_z\}$. A crime has been committed, and there are three suspects: x, y, and z. $G = g_x$ represent the proposition that suspect x is guilty, etc. Suppose my initial epistemic state is complete ignorance represented by disbelief function δ such that

$$\delta(\{g_x\}) = \delta(\{g_y\}) = \delta(\{g_z\}) = 0$$

First, after interviewing suspect x, who has a weak alibi, I disbelieve $\{g_x\}$ to degree 1, that is, $A = \{g_y, g_z\}$ and $\alpha = 1$. My disbelief function now is

$$\delta_{A,\alpha}(\{g_x\}) = 1 \qquad \delta_{A,\alpha}(\{g_y\}) = \delta_{A,\alpha}(\{g_z\}) = 0$$

Next, after interviewing suspect y, who has a stronger alibi, I now disbelieve $\{g_y\}$ to degree 2, that is, $B = \{g_x, g_z\}$, $\gamma = 2$. Then

$$\left(\delta_{A,\alpha}\right)_{B,\gamma}(\{g_x\}) = 1, \qquad \left(\delta_{A,\alpha}\right)_{B,\gamma}(\{g_y\}) = 2, \qquad \left(\delta_{A,\alpha}\right)_{B,\gamma}(\{g_z\}) = 0$$

Note that A and B satisfy the conditions required for commutativity, namely,

$$\delta(A \cap B) = \delta(A \cap \sim B) = \delta(\sim A \cap B) = 0.$$

Note that

$$\delta_{B,\gamma}(\{g_x\}) = 0, \qquad \delta_{B,\gamma}(\{g_x\}) = 2, \qquad \delta_{B,\gamma}(\{g_z\}) = 0$$

and

$$\left(\delta_{B,\gamma}\right)_{A,\alpha}(\{g_x\}) = 1, \quad \left(\delta_{B,\gamma}\right)_{A,\alpha}(\{g_y\}) = 2, \quad \left(\delta_{B,\gamma}\right)_{A,\alpha}(\{g_z\}) = 0.$$

Thus, $(\delta_{A,\alpha})_{B,\gamma} = (\delta_{B,\gamma})_{A,\alpha}$. Given what we learned from interviewing suspects x and y, the fact that we interviewed suspect x before suspect y has no bearing on the final epistemic state.

In Definition 5, a belief revision rule was stated in terms of a single proposition A that was believed to degree α. Spohn [2, p. 318] has generalized this definition to the case where the information gained may concern more than a single proposition. Spohn calls this general belief revision rule λ-*conditionalization*, where λ is the marginal of the resulting disbelief function for some subset h of variables. Here is a formal definition.

DEFINITION 6 [2, P. 318]   *Suppose δ is a disbelief function for g representing our initial epistemic state. Suppose we learn something about variables in set h that consequently leads us to an epistemic state represented by a disbelief function δ_λ for $g \cup h$ such that $\delta_\lambda^{\downarrow h} = \lambda$, where λ is a disbelief function for h. Then the epistemic state δ_λ, called the λ-conditionalization of δ, is defined as follows*:

$$\delta_\lambda(\{(w,u,v)\}) = \delta(\{(w,u)\}) + \lambda(\{(u,v)\}) - \delta(\{u\}^{\uparrow g})$$

for all $w \in \mathscr{W}_{g-h}$, $u \in \mathscr{W}_{g \cap h}$, $v \in \mathscr{W}_{h-g}$.

It is easy to show that δ_λ is indeed a disbelief function. Furthermore, it is also obvious that $\delta_\lambda^{\downarrow h} = \lambda$. Note that A,α-conditionalization is a special case of λ-conditionalization where λ is a disbelief function for g such that $\lambda(\{w\}) = 0$ if $w \in A$ and $\lambda(\{w\}) = \alpha$ if $w \notin A$. As noted by Spohn [2, p. 318], λ-conditionalization is the analog of Jeffrey's rule in probability theory [13, chap. 11].

7. BELIEF REVISION AS A RULE OF COMBINATION

In this section, first, we describe a rule of combination. Second, we show that Spohn's rule for modifying a disbelief function in light of new information can be expressed in terms of this rule of combination.

DEFINITION 7 (A RULE OF COMBINATION)   *Suppose δ_1 and δ_2 are disbelief functions for g_1 and g_2, respectively. The combination of δ_1 and δ_2, denoted by $\delta_1 \oplus \delta_2$, is a disbelief function for $g_1 \cup g_2$ defined as follows*:

$$(\delta_1 \oplus \delta_2)(\{w\}) = \delta_1(\{w^{\downarrow g_1}\}) + \delta_2(\{w^{\downarrow g_2}\}) - K \qquad (2)$$

for all $w \in \mathscr{W}_{g_1 \cup g_2}$, *where*

$$K = \text{MIN}\left\{\delta_1(\{w^{\downarrow g_1}\}) + \delta_2(\{w^{\downarrow g_2}\}) \mid w \in \mathscr{W}_{g_1 \cup g_2}\right\}$$

K is a normalization factor that ensures that $\delta_1 \oplus \delta_2$ is a disbelief function. Thus, combination consists of pointwise addition followed by normalization.

We will now show that Spohn's rule for belief revision as described in the previous section can be expressed in terms of the rule of combination.

THEOREM 6 *Suppose δ_i is an initial disbelief function for g, suppose β_i is the corresponding belief function, and suppose $h \subseteq g$. Suppose we learn something about some contingent proposition A of h. Let δ_f denote the revised disbelief function for g, and let β_f denote the corresponding belief function. Suppose $\beta_f^{\downarrow h}(A) = \alpha$, where $\alpha \in \mathbb{N}$; that is, after revising our beliefs, we believe proposition A to degree α. Then, depending on the value of $\beta_i^{\downarrow h}(A)$, there exists an appropriate disbelief function δ_Δ for h such that $\delta_f = \delta_i \oplus \delta_\Delta$.*

Proof   Let us consider the following cases.

Case (i): $0 \leq \beta_i^{\downarrow h}(A) \leq \alpha$ In this case, whatever we have learned has resulted in an increase or belief for A from degree $\beta_i^{\downarrow h}(A)$ to degree α. Define a disbelief function δ_Δ for h as follows:

$$\delta_\Delta(\{\mathbf{x}\}) = \begin{cases} 0 & \text{if } \mathbf{x} \in A \\ (\alpha - \beta_i^{\downarrow h}(A)) & \text{if } \mathbf{x} \notin A \end{cases}$$

δ_Δ represents the disbelief function representation of what is actually learned. Since $\beta_i^{\downarrow h}(A) \geq 0$, we have

$$\delta_i^{\downarrow h}(A) = \delta_i(A^{\uparrow g}) = 0$$

and

$$\delta_i(\sim A^{\uparrow g}) = \beta_i(A^{\uparrow g}) = \beta_i^{\downarrow h}(A)$$

Therefore, from Definition 5 we have

$$\delta_f(\{\mathbf{w}\}) = \begin{cases} \delta_i(\{\mathbf{w}\}) & \text{if } \mathbf{w} \in A^{\uparrow g} \\ \delta_i(\{\mathbf{w}\}) + (\alpha - \beta_i^{\downarrow h}(A)) & \text{if } \mathbf{w} \notin A^{\uparrow g} \end{cases}$$

Note that $\mathbf{w} \in A^{\uparrow g}$ iff $\mathbf{w}^{\downarrow h} \in A$, and $\mathbf{w} \notin A^{\uparrow g}$ iff $\mathbf{w}^{\downarrow h} \notin A$. From the definition of combination, clearly, $\delta_i \oplus \delta_\Delta = \delta_f$. Since $\delta_i(A^{\uparrow g}) = 0$, the normalization constant $K = 0$.

Case (ii): $0 \leq \alpha \leq \beta_i^{\downarrow h}(A)$ In this case, whatever we have learned has resulted in a decrease in belief for A from degree $\beta_i^{\downarrow h}(A)$ to degree α. Define the disbelief function δ_Δ for h as follows:

$$\delta_\Delta(\{\mathbf{x}\}) = \begin{cases} \beta_i^{\downarrow h}(A) - \alpha & \text{if } \mathbf{x} \in A \\ 0 & \text{if } x \notin A \end{cases}$$

Since $\beta_i^{\downarrow h}(A) > 0$, we have

$$\delta_i^{\downarrow h}(A) = \delta_i(A^{\uparrow g}) = 0$$

and

$$\delta_i(\sim A^{\uparrow g}) = \beta_i(A^{\uparrow g}) = \beta_i^{\downarrow h}(A) > 0$$

Therefore, from Definition 5, we have

$$\delta_f(\{\mathbf{w}\}) = \begin{cases} \delta_i(\{\mathbf{w}\}) & \text{if } \mathbf{w} \in A^{\uparrow g} \\ \delta_i(\{\mathbf{w}\}) + (\alpha - \beta_i^{\downarrow h}(A)) & \text{if } \mathbf{w} \notin A^{\uparrow g} \end{cases}$$

We will now show that $\delta_i \oplus \delta_\Delta = \delta_f$. Note that since $\delta_i(A^{\uparrow g}) = 0$ and $\delta_i(\sim A^{\uparrow g}) > 0$,

$$\{\mathbf{w} \in \mathscr{W}_g \mid \delta_i(\{\mathbf{w}\}) = 0\} \subseteq A^{\uparrow g}$$

Therefore, after pointwise addition of δ_i and δ_Δ, the normalization constant $K = \beta_i^{\downarrow h}(A) - \alpha$. From the definition of combination, we have

$$(\delta_i \oplus \delta_\Delta)(\{\mathbf{w}\}) = \begin{cases} \delta_i(\{\mathbf{w}\}) + (\beta_i^{\downarrow h}(A) - \alpha) - K & \text{if } \mathbf{w} \in A^{\uparrow g} \\ \delta_i(\{\mathbf{w}\}) + 0 - K & \text{if } \mathbf{w} \notin A^{\uparrow g} \end{cases}$$

for all $\mathbf{w} \in \mathscr{W}_g$. Since $K = \beta_i^{\downarrow h}(A) - \alpha$, we get the result.

Case (iii): $\beta_i^{\downarrow h}(A) \le 0 \le \alpha$ In this case, whatever we have learned has resulted in an increase of belief for A. Define the disbelief function δ_Δ for h as follows:

$$\delta_\Delta(\{\mathbf{x}\}) = \begin{cases} 0 & \text{if } \mathbf{x} \in A \\ \alpha - \beta_i^{\downarrow h}(A) & \text{if } \mathbf{x} \notin A \end{cases}$$

Since $\beta_i^{\downarrow h}(A) \le 0$, $\delta_i^{\downarrow h}(A) = \delta_i(A^{\uparrow g}) = -\beta_i^{\downarrow h}(A) \ge 0$, and $\delta_i(\sim A^{\uparrow g}) = 0$. Therefore, from Definition 5, we have

$$\delta_f(\{\mathbf{w}\}) = \begin{cases} \delta_i(\{\mathbf{w}\}) + \beta_i^{\downarrow h}(A) & \text{if } \mathbf{w} \in A^{\uparrow g} \\ \delta_i(\{\mathbf{w}\}) + \alpha & \text{if } \mathbf{w} \notin A^{\uparrow g} \end{cases}$$

We will now show that $\delta_i \oplus \delta_\Delta = \delta_f$. First notice that after pointwise addition, the normalization constant $K = -\beta_i^{\downarrow h}(A)$. This is because

$$\text{MIN}\,\{\delta_i(\{\mathbf{w}\}) \mid \mathbf{w} \in A^{\uparrow g}\} = \delta_i(A^{\uparrow g}) = -\beta_i^{\downarrow h}(A)$$

and

$$\text{MIN}\,\{\delta_i(\{\mathbf{w}\}) + \alpha - \beta_i^{\downarrow h}(A) \mid \mathbf{w} \in \sim A^{\uparrow g}\} = \alpha - \beta_i^{\downarrow h}(A).$$

Therefore, from the definition of combination we have

$$(\delta_i \oplus \delta_\Delta)(\{\mathbf{w}\}) = \begin{cases} \delta_i(\{\mathbf{w}\}) + \beta_i^{\downarrow h}(A) & \text{if } \mathbf{w} \in A^{\uparrow g} \\ \delta_i(\{\mathbf{w}\}) + \alpha & \text{if } \mathbf{w} \notin A^{\uparrow g} \end{cases}$$

$$= \delta_f(\{\mathbf{w}\})$$

for all $\mathbf{w} \in \mathscr{W}_g$.

Thus in all three cases, which is an exhaustive list, belief revision reduces to the rule of combination.

In the next section we will show that our rule of combination and Spohn's λ-conditionalization are mathematically equivalent. But first we pause for a numerical example to illustrate Theorem 6.

EXAMPLE 7 (A, α-conditionalization as a rule of combination)   Consider the example of belief revision given in Example 4. The initial epistemic state is given by the disbelief function δ for $\{R, Q, P\}$. Since the initial belief for proposition $\{r\}$ is 0 and the belief after the first evidence is 3, we can describe what we have learned from the first body of evidence by disbelief function δ_Δ for $\{R\}$ as follows:

$$\delta_\Delta'(\{r\}) = 0 \qquad \text{and} \qquad \delta_\Delta'(\{\sim r\}) = 3$$

Note that the resulting disbelief function $\delta' = \delta \oplus \delta_\Delta'$ (the normalization constant in this combination is $K = 0$).

As per δ', we disbelieve $\{q\}$ to degree 1. After the second body of evidence, we believe $\{q\}$ to degree 3. Therefore, what we have learned from the second body of evidence can be represented by disbelief function δ_Δ'' for $\{Q\}$ as follows:

$$\delta_\Delta''(\{q\}) = 0 \qquad \text{and} \qquad \delta_\Delta''(\{\sim q\}) = 4$$

Note that $\delta'' = \delta' \oplus \delta_\Delta''$ (the normalization constant in this combination is $K = 1$).

8. A RULE OF SUBTRACTION

In this section, first, we define a rule of subtraction for disbelief functions. Second, we show that we can always recover the incremental disbelief function from the final and initial disbelief functions. Third, we show the mathematical equivalence between Spohn's λ-conditionalization and our rule of combination.

Spohn's belief revision rules were described in terms of the initial disbelief function δ_i and characteristics of the final disbelief function (proposition A

and its degree of belief α in A, α-conditionalization and disbelief function λ in λ-conditionalization). On the other hand, the rule of combination describes the final disbelief function in terms of the initial disbelief function and the incremental disbelief function representing the evidence. If we are given the initial and the final disbelief function, can we always recover the incremental disbelief function? The answer is yes and is stated below as Theorem 7. First we need a definition.

DEFINITION 8 (A RULE OF SUBTRACTION) *Suppose δ_1 is a disbelief function for g, suppose δ_2 is a disbelief function for h, and suppose $h \subseteq g$. Then the subtraction of δ_2 from δ_1, denoted by $\delta_1 - \delta_2$, is a disbelief function for g given by*

$$(\delta_1 - \delta_2)(\{\mathbf{w}\}) = \delta_1(\{\mathbf{w}\}) - \delta_2(\{\mathbf{w}^{\downarrow h}\}) - K$$

for all $\mathbf{w} \in \mathscr{W}_g$, where K is a normalization constant given by

$$K = \text{MIN} \{\delta_1(\{\mathbf{w}\}) - \delta_2(\{\mathbf{w}^{\downarrow h}\}) \mid \mathbf{w} \in \mathscr{W}_g\}$$

It is clear from the definition of the normalization constant K that $\delta_1 - \delta_2$ is a disbelief function. The next theorem states that we can recover the incremental disbelief function from the initial and final disbelief functions.

THEOREM 7 *Suppose δ_i and δ_Δ are disbelief functions for g and h, respectively. Then*

$$\left((\delta_i \oplus \delta_\Delta) - \delta_i\right)^{\downarrow h} = \delta_\Delta$$

Proof Suppose $\mathbf{w} \in \mathscr{W}_h$. Then

$$\left((\delta_i \oplus \delta_\Delta) - \delta_i\right)^{\downarrow h}(\{\mathbf{w}\})$$

$$= \left((\delta_i \oplus \delta_\Delta) - \delta_i\right)(\{\mathbf{w}\}^{\uparrow(g \cup h)})$$

$$= \text{MIN} \left\{\left((\delta_i \oplus \delta_\Delta) - \delta_i\right)(\{\mathbf{z}\}) \mid \mathbf{z} \in \{\mathbf{w}\}^{\uparrow(g \cup h)}\right\}$$

$$= \text{MIN} \left\{(\delta_i \oplus \delta_\Delta)(\{\mathbf{z}\}) - \delta_i(\{\mathbf{z}^{\downarrow g}\}) - K_1 \mid \mathbf{z} \in \{\mathbf{w}\}^{\uparrow(g \cup h)}\right\}$$

$$= \text{MIN} \left\{\delta_i(\{\mathbf{z}^{\downarrow g}\}) + \delta_\Delta(\{\mathbf{z}^{\downarrow h}\}) - K_2\right.$$

$$\left. - \delta_i(\{\mathbf{z}^{\downarrow g}\}) - K_1 \mid \mathbf{z} \in \{\mathbf{w}\}^{\uparrow(g \cup h)}\right\}$$

$$= \text{MIN} \left\{\delta_\Delta(\{\mathbf{z}^{\downarrow h}\}) - K_2 - K_1 \mid \mathbf{z} \in \{\mathbf{w}\}^{\uparrow(g \cup h)}\right\}$$

$$= \delta_\Delta(\{\mathbf{w}\}) - K_2 - K_1 \quad (\text{since } \mathbf{z}^{\downarrow h} = \mathbf{w} \text{ for all } \mathbf{z} \in \{\mathbf{w}\}^{\uparrow(g \cup h)})$$

Table 5.   The Rule of Subtraction

w	$\delta'(\{w\})$	$\delta''(\{w\})$	$(\delta'' - \delta')(\{w\})$
(r,q,p)	1	0	0
$(r,q,\sim p)$	2	1	0
$(r,\sim q,p)$	1	4	4
$(r,\sim q,\sim p)$	0	3	4
$(\sim r,q,p)$	3	2	0
$(\sim r,q,\sim p)$	5	4	0
$(\sim r,\sim q,p)$	3	6	4
$(\sim r,\sim q,\sim p)$	3	6	4

The normalization constant K_1 in the subtraction operation simplifies as follows:

$$K_1 = \text{MIN}\left\{(\delta_i \oplus \delta_\Delta)(\{z\}) - \delta_i(\{z^{\downarrow g}\}) \,|\, z \in \mathscr{W}_{g \cup h}\right\}$$

$$= \text{MIN}\left\{\delta_i(\{z^{\downarrow g}\}) + \delta_\Delta(\{z^{\downarrow h}\}) - K_2 - \delta_i(\{z^{\downarrow g}\}) \,|\, z \in \mathscr{W}_{g \cup h}\right\}$$

$$= \text{MIN}\left\{\delta_\Delta(\{z^{\downarrow h}\}) - K_2 \,|\, z \in \mathscr{W}_{g \cup h}\right\}$$

$$= \text{MIN}\left\{\delta_\Delta(\{z^{\downarrow h}\}) \,|\, z \in \mathscr{W}_{g \cup h}\right\} - K_2$$

$$= -K_2 \quad \left(\text{since } \delta_\Delta \text{ is a disbelief function}\right)$$

Therefore, $((\delta_i \oplus \delta_\Delta) - \delta_i)^{\downarrow h} = \delta_\Delta$.

COROLLARY TO THEOREM 7   *Suppose δ is a disbelief function for g, suppose δ_Δ is a disbelief function for h, and suppose $h \subseteq g$. Then $(\delta - \delta_\Delta) \oplus \delta_\Delta = \delta$.*

Proof   Suppose δ is a disbelief function for g, suppose δ_Δ is a disbelief function for h, and suppose $h \subseteq g$. Let $u \in \mathscr{W}_{g-h}$, $v \in \mathscr{W}_h$. Then

$$\left((\delta - \delta_\Delta) \oplus \delta_\Delta\right)(\{(u,v)\})$$

$$= (\delta - \delta_\Delta)(\{(u,v)\}) + \delta_\Delta(\{v\}) - K_1$$

$$= \delta(\{(u,v)\}) - \delta_\Delta(\{v\}) - K_2 + \delta_\Delta(\{v\}) - K_1$$

$$= \delta(\{(u,v)\}) - K_2 - K_1$$

As in the proof of Theorem 7, it is easy to show that $K_1 = -K_2$. Hence the proof follows.

The following example illustrates the rule of subtraction.

EXAMPLE 8 Consider the disbelief functions δ' and δ'' for $\{R,Q,P\}$ from Example 4 reproduced in Table 5. Table 5 also shows the disbelief function

$\delta'' - \delta'$. The normalization constant in this case is $K = -1$. Also, notice that $(\delta'' - \delta')^{\downarrow\{Q\}}$ is the same as δ_Δ'' described in Example 7.

The property of disbelief functions of being always able to recover the addendum from the sum is unique to this theory and is not shared either by probability theory or by the Dempster–Shafer theory of belief functions. This property is useful for two reasons. First, in cases where it is easier to describe evidence by reference to the final disbelief function, we can always recover the incremental belief function that represents just the evidence. Second, this property is useful in nonmonotonic reasoning because it allows us to retract the conclusion of an earlier inference without influencing conclusions drawn using other means (see e.g., Ginsberg [14]).

We now have the necessary tools to prove that Spohn's λ-conditionalization and our rule of combination are mathematically equivalent.

THEOREM 8   *Suppose δ is a disbelief function for g and λ is a disbelief function for h. Let δ_λ denote the λ-conditionalization of δ. Then there exists a disbelief function δ_Δ for h such that $\delta \oplus \delta_\Delta = \delta_\lambda$. Conversely, suppose δ is a disbelief function for g and δ_Δ is a disbelief function for h. Then there exists a disbelief function λ for h such that $\delta_\lambda = \delta \oplus \delta_\Delta$.*

Proof $(\Rightarrow)$ Suppose δ is a disbelief function for g and λ is a disbelief function for h. Note that δ_λ is a disbelief function for $g \cup h$. Define $\delta_\Delta = (\delta_\lambda - \delta)^{\downarrow h}$. Let $\mathbf{w} \in \mathscr{W}_{g-h}$, $\mathbf{u} \in \mathscr{W}_{g \cap h}$, and $\mathbf{v} \in \mathscr{W}_{h-g}$. Then

$$\left(\delta \oplus \delta_\Delta\right)\left(\{(\mathbf{w},\mathbf{u},\mathbf{v})\}\right) = \delta\left(\{(\mathbf{w},\mathbf{u})\}\right) + \left(\delta_\lambda - \delta\right)^{\downarrow h}\left(\{(\mathbf{u},\mathbf{v})\}\right) - K_1$$

Simplifying the second term on the right-hand side, we get

$$\left(\delta_\lambda - \delta\right)^{\downarrow h}\left(\{(\mathbf{u},\mathbf{v})\}\right)$$

$$= \left(\delta_\lambda - \delta\right)\left(\{(\mathbf{u},\mathbf{v})\}^{\uparrow(g\cup h)}\right)$$

$$= \mathrm{MIN}\left\{\left(\delta_\lambda - \delta\right)\left(\{(\mathbf{x},\mathbf{u},\mathbf{v})\}\right) \,\middle|\, \mathbf{x} \in \mathscr{W}_{g-h}\right\}$$

$$= \mathrm{MIN}\left\{\delta_\lambda\left(\{(\mathbf{x},\mathbf{u},\mathbf{v})\}\right) - \delta\left(\{(\mathbf{x},\mathbf{u})\}\right) - K_2 \,\middle|\, \mathbf{x} \in \mathscr{W}_{g-h}\right\}$$

$$= \mathrm{MIN}\left\{\delta\left(\{(\mathbf{x},\mathbf{u})\}\right) + \lambda\left(\{(\mathbf{u},\mathbf{v})\}\right) - \delta\left(\{\mathbf{u}\}^{\uparrow g}\right)\right.$$

$$\left. -\,\delta\left(\{(\mathbf{x},\mathbf{u})\}\right) - K_2 \,\middle|\, \mathbf{x} \in \mathscr{W}_{g-h}\right\}$$

$$= \mathrm{MIN}\left\{\lambda\left(\{(\mathbf{u},\mathbf{v})\}\right) - \delta\left(\{\mathbf{u}\}^{\uparrow g}\right) - K_2 \,\middle|\, \mathbf{x} \in \mathscr{W}_{g-h}\right\}$$

$$= \lambda\left(\{(\mathbf{u},\mathbf{v})\}\right) - \delta\left(\{\mathbf{u}\}^{\uparrow g}\right) - K_2$$

Therefore we have

$$(\delta \oplus \delta_\Delta)(\{(w,u,v)\})$$

$$= \delta(\{(w,u)\}) + \lambda(\{(u,v)\}) - \delta(\{u\}^{\uparrow g}) - K_2 - K_1$$

$$= \delta_\lambda(\{(w,u,v)\}) - K_2 - K_1$$

The normalization constant K_1 in the combination operation simplifies as follows:

$$K_1 = \text{MIN} \left\{ \delta(\{(w,u)\}) + (\delta_\lambda - \delta)^{\downarrow h}(\{(u,v)\}) \mid (w,u,v) \in \mathscr{W}_{g \cup h} \right\}$$

$$= \text{MIN} \left\{ \delta(\{(w,u)\}) + \lambda(\{(u,v)\}) \right.$$

$$\left. - \delta(\{u\}^{\uparrow g}) - K_2 \mid (w,u,v) \in \mathscr{W}_{g \cup h} \right\}$$

$$= \text{MIN} \left\{ \delta_\lambda(\{(w,u,v)\}) \mid (w,u,v) \in \mathscr{W}_{g \cup h} \right\} - K_2$$

$$= -K_2 \quad (\text{since } \delta_\lambda \text{ is a disbelief function})$$

Therefore $\delta \oplus \delta_\Delta = \delta_\lambda$.

$(\Leftarrow)$ Suppose δ is a disbelief function for g and δ_Δ is a disbelief function for h. Define λ to be $(\delta \oplus \delta_\Delta)^{\downarrow h}$. We need to show that the λ-conditionalization of δ is indeed $\delta \oplus \delta_\Delta$. Let $w \in \mathscr{W}_{g-h}$, $u \in \mathscr{W}_{g \cap h}$, and $v \in \mathscr{W}_{h-g}$. Then

$$\delta_\lambda(\{(w,u,v)\})$$

$$= \delta(\{(w,u)\}) + (\delta \oplus \delta_\Delta)^{\downarrow h}(\{(u,v)\}) - \delta(\{u\}^{\uparrow g})$$

$$= \delta(\{(w,u)\}) + (\delta \oplus \delta_\Delta)(\{(u,v)\}^{\uparrow(g \cup h)}) - \delta(\{u\}^{\uparrow g})$$

$$= \delta(\{(w,u)\}) + \text{MIN} \left\{ (\delta \oplus \delta_\Delta)(\{(x,u,v)\}) \mid x \in \mathscr{W}_{g-h} \right\} - \delta(\{u\}^{\uparrow g})$$

$$= \delta(\{(w,u)\}) + \text{MIN} \left\{ \delta(\{(x,u)\}) + \delta_\Delta(\{(u,v)\}) \right.$$

$$\left. - K \mid x \in \mathscr{W}_{g-h} \right\} - \delta(\{u\}^{\uparrow g})$$

$$= \delta(\{(w,u)\}) + \text{MIN} \left\{ \delta(\{(x,u)\}) \mid x \in \mathscr{W}_{g-h} \right\}$$

$$+ \delta_\Delta(\{(u,v)\}) - K - \delta(\{u\}^{\uparrow g})$$

$$= \delta(\{(w,u)\}) + \delta_\Delta(\{(u,v)\}) - K$$

$$= (\delta \oplus \delta_\Delta)(\{(w,u,v)\})$$

9. PROPERTIES OF THE COMBINATION RULE

In this section, we discuss several important properties of the combination rule described in Section 7. First, we state some elementary properties of the

Table 6. Spohn's Rule for Belief Revision as a Rule of Combination

w	δ_1	δ_2	$\delta_1 \oplus \delta_2$	δ_3	$\delta_1 \oplus \delta_2 \oplus \delta_3$
s	0	0	0	2	1
$\sim s$	0	1	1	0	0

rule of combination. Second, we describe how the rule of combination can be used to construct a disbelief function for many variables from independent disbelief functions each of which involves only a few variables. We sketch what we mean by independent disbelief functions. Third, we show that Spohn's theory of disbelief functions fits into the same abstract framework as that for the theory of probability and the Dempster–Shafer theory of belief functions.

First, we state some elementary properties of the rule of combination.

PROPOSITION 3   *The rule of combination described in* (2) *has the following properties*
   C1. (Commutativity) $\delta_1 \oplus \delta_2 = \delta_2 \oplus \delta_1$.
   C2. (Associativity) $(\delta_1 \oplus \delta_2) \oplus \delta_3 = \delta_1 \oplus (\delta_2 \oplus \delta_3)$.
   C3. If δ_1 is vacuous, then $\delta_1 \oplus \delta_2 = \delta_2$.
   C4. In general, $\delta_1 \oplus \delta_1 \neq \delta_1$. The disbelief function $\delta_1 \oplus \delta_1$ disbelieves the same propositions as δ_1, but it will do so with twice the degree, as it were.

   Proof   All four properties follow trivially from the definition of combination.

Note that although Spohn's belief revision rule is partially commutative, the rule of combination is always commutative. There is no conflict here. Spohn's belief revision rule is described in terms of the initial epistemic state and some features of the final epistemic state. On the other hand, the rule of combination describes belief revision in terms of the initial epistemic state and the epistemic state representing the evidence. Thus commutativity for Spohn's rule and commutativity for the rule of combination are two different relations. The following example illustrates this for the noncommutative belief revision case of Example 5.

EXAMPLE 9 (Noncommutative belief revision as a commutative combination) Consider the story of Example 5. The initial epistemic state for S is the vacuous disbelief function represented by δ_1 in Table 6. Following the weather announcement, I now believe that it is snowing to degree 1. Using the rule of subtraction, the incremental epistemic state representing the weather announcement, denoted by δ_2, is shown in Table 6. The current epistemic state is given by $\delta_1 \oplus \delta_2$. After I look out the window, I now believe that it is not snowing to degree 1. Again, using the rule of subtraction, the incremental epistemic state representing the evidence obtained from looking out the window, denoted

by δ_3, is shown in Table 6. The final epistemic state is given by $\delta_1 \oplus \delta_2 \oplus \delta_3$. Note that $\delta_1 \oplus \delta_2 \oplus \delta_3 = \delta_1 \oplus \delta_3 \oplus \delta_2$.

Second, the rule of combination is valid not only for belief revision but also for the construction of an initial disbelief function for many variables when this information is given in terms of independent disbelief functions each of which is for small number of variables. What do we mean by independent disbelief functions. Here, we will just sketch an answer by analogy with the theory of probability and the Dempster–Shafer theory of belief functions. A complete answer merits a separate paper.

By independent beliefs, we mean the same as in the theory of probability and the Dempster–Shafer theory of belief functions. In probability theory, beliefs are represented by functions called *potentials*, and the rule of combination is pointwise multiplication (see, e.g., Refs. 10 and 11). However, combining two potentials gives us meaningful results only when the potentials being combined are independent. For example, suppose we have two variables X and Y with frames $\{x, \sim x\}$ and $\{y, \sim y\}$, respectively. Consider a potential $p_1 = (p(x), p(\sim x))$ for X, representing a (prior) probability distribution for X. Consider another potential $p_2 = (p(y \mid x), p(\sim y \mid x), p(y \mid \sim x), p(\sim y \mid \sim x))$ for $\{X, Y\}$, representing the conditional distributions for Y given X. In this case, the potentials p_1 and p_2 are independent. Combining these pointwise multiplication, denoted by $\otimes$, gives us the joint potential $p_1 \otimes p_2$ for $\{X, Y\}$ as follows

$$\left(p(y|x)p(x), \ p(\sim y|x)p(x), \ p(y| \sim x)p(\sim x), \ p(\sim y| \sim x)p(\sim x) \right)$$

We recognize this potential as the joint probability distribution for $\{X, Y\}$. Consider another potential $p_3 = (p(y), p(\sim y))$ for Y representing a probability distribution of Y. In general, p_1 and p_3 are not independent. If we combine these two potentials, the result

$$p_1 \otimes p_3 = \left(p(x)p(y), \ p(x)p(\sim y), \ p(\sim x)p(y), \ p(\sim x)p(\sim y) \right)$$

does not necessarily represent pooling of evidence. We know from probability theory that $p_1 \otimes p_3$ represents combination of evidence if and only if X and Y are probabilistically independent, that is, if and only if p_1 and p_3 are independent potentials.

In the Dempster–Shafer theory, Dempster's rule for combining belief functions represents pooling of evidence only when the belief functions being combined are independent. Shafer [15, 16] has described in detail precisely what is meant by independent belief functions in terms of canonical examples for belief functions. In fact, most of the examples that Pearl [17, pp. 447–450] describes to show that application of Dempster's rule given nonintuitive results do so precisely because the belief functions being combined are not independent.

Spohn's theory of disbelief functions is closely analogous to the theory of probability (see Spohn [2, pp. 318–320] for a comparison of his theory with probability theory). Analogous to the concept of probabilistic conditional independence, Spohn [1, pp. 120–125; 2, p. 318] has described conditional independence for disbelief functions. As in the probabilistic case, Hunter [18] has shown that the conditional independence relation for disbelief functions forms a "graphoid" (see, e.g., Geiger and Pearl [19] or Verma and Pearl [20] for definition of a graphoid). Using the notion of conditional independence, we can define when two disbelief functions are independent by direct analogy with the theory of probability.

EXAMPLE 10 (Construction of an initial disbelief function)   Consider two sets of beliefs as follows:
  1. Most Republicans are not pacifists.
  2. Most Quakers are pacifists.

Suppose further that these two sets of beliefs are independent. Then, if δ_1 is a disbelief representation of the first set of beliefs and δ_2 is a disbelief representation of the second set of beliefs, then $\delta_1 \oplus \delta_2$ will represent the aggregation of these two sets of beliefs. In particular, suppose δ_1 is a disbelief function for $\{R, P\}$ as follows:

$$\delta_1(\{(r,p)\}) = 1$$

$$\delta_1(\{(r,\sim p)\}) = \delta_1(\{(\sim r,p)\}) = \delta_1(\{(\sim r,\sim p)\}) = 0$$

(i.e., we disbelieve pacifist Republicans), and suppose δ_2 is a disbelief function for $\{Q, P\}$ as follows:

$$\delta_2(\{(q,\sim p)\}) = 2$$

$$\delta_2(\{(q,p)\}) = \delta_2(\{(\sim q,p)\}) = \delta_2(\{(\sim q,\sim p)\}) = 0$$

(i.e., we disbelieve nonpacifist Quakers). Then the disbelief function $\delta_1 \oplus \delta_2$ = δ, say, shown in Table 2, represents the aggregate belief. Note that there is no belief revision going on here. Of course, Theorem 8 tells us that we can mathematically describe the aggregation of δ_1 and δ_2 using λ-conditionalization (λ is the same as δ_2 for this example). But in general, it is neither practical nor intuitive.

Third, as per our reformulation, Spohn's theory of epistemic beliefs shares the essential abstract features of probability theory and the Dempster–Shafer theory of belief functions as described by Shenoy and Shafer [10, 11]. These features are (1) a functional representation of knowledge (or beliefs), (2) a rule of marginalization, and (3) a rule of combination. In Refs. 10 and 11, we also state three axioms for the marginalization and combination rules that enable

one to use local computations in the calculation of the marginals of a joint disbelief function with explicitly having to compute the joint disbelief function. These three axioms are as follows (stated in the notation of disbelief functions).

L1. (Commutativity and Associativity of Combination) Suppose δ_1, δ_2, and δ_3 are disbelief functions for g, h, and k respectively. Then,

$$\delta_1 \oplus \delta_2 = \delta_2 \oplus \delta_1,$$

and

$$(\delta_1 \oplus \delta_2) \oplus \delta_3 = \delta_1 \oplus (\delta_2 \oplus \delta_3)$$

L2. (Consonance of Marginalization) Suppose δ is a disbelief function for g, and suppose $k \subseteq h \subseteq g$. Then

$$\left(\delta^{\downarrow h}\right)^{\downarrow k} = \delta^{\downarrow k}$$

L3. (Distributivity of Marginalization over Combination) Suppose δ_1 and δ_2 are disbelief functions for g and h, respectively, Then

$$(\delta_1 \oplus \delta_2)^{\downarrow g} = \delta_1 \oplus \left(\delta_2^{\downarrow(g \cap h)}\right)$$

We have already shown that axioms L1 and L2 are valid for disbelief functions. Theorem 9 states that axiom L3 is also satisfied.

THEOREM 9 *Suppose δ_1 and δ_2 are disbelief functions for g and h, respectively. Then*

$$(\delta_1 \oplus \delta_2)^{\downarrow g} = \delta_1 \oplus \left(\delta_2^{\downarrow(g \cap h)}\right)$$

Proof Note that $g \cup h = (g - h) \cup (g \cap h) \cup (h - g)$, $g = (g - h) \cup (g \cap h)$, and $h = (g \cap h) \cup (h - g)$. Suppose $\mathbf{w} \in \mathcal{W}_{g-h}$, $\mathbf{u} \in \mathcal{W}_{g \cap h}$, and $\mathbf{v} \in \mathcal{W}_{h-g}$. Then $(\mathbf{w},\mathbf{u}) \in \mathcal{W}_g$ and $(\mathbf{u},\mathbf{v}) \in \mathcal{W}_h$. First, note that the normalization factor in the combination of the left-hand side, say K_1, is the same as the normalization factor in the combination on the right-hand side, say K_2, that is, $K_1 = K_2$, as shown below.

$$K_1 = \text{MIN}\left\{\delta_1(\{(\mathbf{w},\mathbf{u})\}) + \delta_2(\{(\mathbf{u},\mathbf{v})\}) \mid (\mathbf{w},\mathbf{u},\mathbf{v}) \in \mathcal{W}_{g \cup h}\right\}$$

$$= \text{MIN}\left\{\text{MIN}\left\{\delta_1(\{(\mathbf{w},\mathbf{u})\}) + \delta_2(\{(\mathbf{u},\mathbf{v})\}) \mid \mathbf{v} \in \mathcal{W}_{h-g}\right\} \mid (\mathbf{w},\mathbf{u}) \in \mathcal{W}_g\right\}$$

$$= \text{MIN}\left\{\delta_1(\{(\mathbf{w},\mathbf{u})\}) + \text{MIN}\left\{\delta_2(\{(\mathbf{u},\mathbf{v})\}) \mid \mathbf{v} \in \mathcal{W}_{h-g}\right\} \mid (\mathbf{w},\mathbf{u}) \in \mathcal{W}_g\right\}$$

$$= \text{MIN}\left\{\delta_1(\{(\mathbf{w},\mathbf{u})\}) + \left(\delta_2^{\downarrow(g \cap h)}\right)(\{(\mathbf{u})\}) \mid (\mathbf{w},\mathbf{u}) \in \mathcal{W}_g\right\}$$

$$= K_2 = K, \text{ say}.$$

Next observe that

$$
\begin{aligned}
(\delta_1 \oplus \delta_2)^{\downarrow g}(\{(\mathbf{w},\mathbf{u})\}) \\
&= \mathrm{MIN}\left\{(\delta_1 \oplus \delta_2)(\{(\mathbf{w},\mathbf{u},\mathbf{v})\}) \mid \mathbf{v} \in \mathscr{W}_{h-g}\right\} \\
&= \mathrm{MIN}\left\{\delta_1(\{(\mathbf{w},\mathbf{u})\}) + \delta_2(\{(\mathbf{u},\mathbf{v})\}) - K \mid \mathbf{v} \in \mathscr{W}_{h-g}\right\} \\
&= \delta_1(\{(\mathbf{w},\mathbf{u})\}) + \mathrm{MIN}\left\{\delta_2(\{(\mathbf{u},\mathbf{v})\}) \mid \mathbf{v} \in \mathscr{W}_{h-g}\right\} - K \\
&= \delta_1(\{(\mathbf{w},\mathbf{u})\}) + \left(\delta_2^{\downarrow(g \cap h)}\right)(\{\mathbf{u}\}) - K \\
&= \left(\delta_1 \oplus \left(\delta_2^{\downarrow(g \cap h)}\right)\right)(\{(\mathbf{w},\mathbf{u})\})
\end{aligned}
$$

Since all three axioms required for local computation of marginals are satisfied, the scheme described in Ref. 11 can be used for belief updating. Hunter [12] describes an analogous scheme for belief revision.

10. DISCUSSION

In Ref. 21 we describe a valuation-based language for representing and reasoning with knowledge. In such a language, knowledge is represented by functions called valuations, and inferences are made from the knowledge base using two operators called combination and marginalization. Combination corresponds to aggregation of knowledge, and marginalization corresponds to crystallization of knowledge. Conceptually, all the valuations are combined to obtain what is called the *joint valuation*. The marginals of the joint valuation are then found for each variable. If combination and marginalization operators satisfy three axioms, then the marginals of the joint valuation can be found using local computation without actually computing the joint valuation. In Refs. 10 and 11 we show that Bayesian probability theory and the Dempster–Shafer theory of belief functions fit in the abstract framework of valuation-based languages. In this paper, we have shown that Spohn's theory of epistemic beliefs also fits in this abstract framework. One implication of this is that we have a ready-made algorithm for propagating disbelief functions that used only local computation. Another implication is that we now have a better understanding of the sense in which Spohn's theory differs from probability theory and the Dempster–Shafer theory of belief functions (in the functional representation of knowledge, and rules of combination and marginalization), and the sense in which it is similar to these alternative theories of uncertain reasoning (the abstract features of the axiomatic framework).

ACKNOWLEDGMENTS

References

1. Spohn, W., Ordinal conditional functions: a dynamic theory of epistemic states, in, *Causation in Decision, Belief Change, and Statistics,* vol. 2, (W. L. Harper and B. Skyrms, Eds.), Reidel, Dordrecht, Holland, 1988, pp. 105–134.

2. Spohn, W., A general non-probabilistic theory of inductive reasoning, *Proceedings of the 4th Workshop on Uncertainty in Artificial Intelligence,* Minneapolis, Minn., 315–322, 1988. Also in *Uncertainty in Artificial Intelligence,* vol. 4, (R. Shachter, T. Levitt, J. Lemmer, and L. Kanal, Eds.), North-Holland, Amsterdam, 1990.

3. Dempster, A. P., A generalization of Bayesian inference *J. Roy. Stat. Soc., Ser. B* **30,** 205–247, 1968.

4. Shafer, G. *A Mathematical Theory of Evidence,* Princeton Univ. Press, Princeton, N.J., 1976.

5. Pearl, J., Probabilistic semantics for nonmonotonic reasoning: a survey, *Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning,* Toronto, Canada, 505–516, 1989.

6. Dubois, D., and Prade, H., Epistemic entrenchment and possibilistic logic, unpublished manuscript, 1990.

7. Zadeh, L. A., Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets Sys.* **1,** 3–28, 1978.

8. Dubois, D., and Prade, H., *Possibility Theory: An Approach to Computerized Processing of Uncertainty,* Plenum, New York, 1988.

9. Gardenfors, P., *Knowledge in Flux: Modeling the Dynamics of Epistemic States,* MIT Press, Cambridge, Mass., 1988.

10. Shenoy, P. P., and Shafer, G., An axiomatic framework for Bayesian and belief-function propagation, *Proceedings of the 4th Workshop on Uncertainty in Artificial Intelligence,* Minneapolis, Minn., 307–314, 1988.

11. Shenoy, P. P., and Shafer, G., Axioms for probability and belief-function propagation, Working Paper No. 209, School of Business, University of Kansas, Lawrence, Kan. Also in *Uncertainty in Artificial Intelligence,* vol. 4, (R. Shachter, T. Levitt, J. Lemmer, and L. Kanal, Eds.), North-Holland, Amsterdam, 169–198, 1990.

12. Hunter, D., Parallel belief revision, *Proceedings of the 4th Workshop on Uncertainty in Artificial Intelligence,* Minneapolis, Minn., 170–176, 1988. Also

in *Uncertainty in Artificial Intelligence*, vol. 4, (R. Shachter, T. Levitt, J. Lemmer, and L. Kanal, Eds.), North-Holland, Amsterdam, 241–251, 1990.

13. Jeffrey, R. C., *The Logic of Decision*, 2nd ed., Chicago Univ. Press, Chicago, 1983.

14. Ginsberg, M. L., Non-monotonic reasoning using Dempster's rule, *Proceedings of the 3rd National Conference on Artificial Intelligence* (AAAI-84), Austin, Tex., 126–129, 1984.

15. Shafer, G., The problem of dependent evidence, School of Business Working Paper No. 164, University of Kansas, Lawrence, Kansas, 1984.

16. Shafer, G., Belief functions and possibility measures in *Analysis of Fuzzy Information*, Vol. I, *Mathematics and Logic* (J. C. Bezdek, Ed.), CRC Press, Boca Raton, Fla., 51–84, 1987.

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

18. Hunter, D., Graphoids, semi-graphoids, and ordinal conditional functions, unpublished manuscript, 1988.

19. Geiger, D., and Pearl, J., On the logic of causal models, *Proceedings of 4th Workshop on Uncertainty in Artificial Intelligence,* 1988, 136–147, Minneapolis, Minn. Also in *Uncertainty in Artificial Intelligence*, vol. 4, (R. Shachter, T. Levitt, J. Lemmer, and L. Kanal, Eds.), North-Holland, Amsterdam, 3–14, 1990.

20. Verma, T., and Pearl, J., Causal networks: semantics and expressiveness, *Proceedings of the 4th Workshop on Uncertainty in Artificial Intelligence,* Minneapolis, Minn., 352–359. Also in *Uncertainty in Artificial Intelligence*, vol. 4, (R. Shachter, T. Levitt, J. Lemmer, and L. Kanal, Eds), North-Holland, Amsterdam, 69–76, 1990.

21. Shenoy, P. P., A valuation-based language for expert systems, *Int. J. Approximate Reasoning* 3(5), 383–411, 1989.