

USING POSSIBILITY THEORY IN EXPERT SYSTEMS

by

Prakash P. Shenoy

*School of Business, University of Kansas, Lawrence, KS 66045-7585, USA
Tel: (785)-864-7551, Fax: (785)-864-5328, Email: PSHENOY@KU.EDU*

ABSTRACT

This paper has two main objectives. The first objective is to give a characterization of a qualitative description of a possibility function. A qualitative description of a possibility function is called a consistent possibilistic state. The qualitative description and its characterization serve as qualitative semantics for possibility functions. These semantics are useful in representing knowledge as possibility functions. The second objective is to describe how Zadeh's theory of possibility fits in the framework of valuation-based systems (VBS). Since VBS serve as a framework for managing uncertainty and imprecision in expert systems, this facilitates the use of possibility theory in expert systems.

Key Words: Possibility theory, consistent possibilistic state, valuation-based systems, expert systems

1. INTRODUCTION

The main goal of this paper is to describe how Zadeh's theory of possibility can be used for managing uncertainty and imprecision in expert systems. We give a characterization of a qualitative description of a possibility function. This is a useful first step in representing knowledge as possibility functions. Also, we describe how possibility theory fits in the framework of valuation-based systems (VBS). Since VBS serve as a framework for managing uncertainty and imprecision in expert systems, this facilitates the use of possibility theory in expert systems.

The theory of possibility was first described by Zadeh [1978, 1979] and further developed by Dubois and Prade [1988]. In possibility theory, the basic representational unit is called a possibility function. The two main operations for manipulating possibility functions are called projection and particularization [Zadeh 1979].

In this paper, we are concerned with two issues related to using possibility theory in expert systems. The first issue is in representing knowledge as possibility functions. To do this effectively, we need semantics for possibility functions. What does it mean, for example, to say that

possibility of a proposition is 0.8? In probability theory, for example, the probability of a proposition can be interpreted as a relative frequency or as a betting rate. With this goal in mind, we first define a qualitative description of a possibility function called a consistent possibilistic state. This is analogous to the definition of a consistent epistemic state [Shenoy 1991a] in Spohn's theory of epistemic beliefs [Spohn 1988]. Next, we give a characterization of a consistent possibilistic state in terms of what we call the content of a consistent possibilistic state. Again, this characterization is analogous to Spohn's characterization of a consistent epistemic state in terms of its content [Spohn 1988]. This is only one step toward developing semantics for possibility functions. Semantics for the quantitative aspects of possibility functions have been studied by, for example, Dubois et al. [1989] and Ruspini [1991].

The second issue related to using possibility theory in expert systems is making the correspondence between concepts in possibility theory and concepts in expert systems. If we interpret possibility functions as knowledge, what do the projection and particularization operations mean in the context of expert systems? How do we use these operations to make inferences in a knowledge-based system? Also, since expert systems typically deal with many propositions, there is a computational issue of the tractability of using possibility theory to solve large problems. To help answer these questions, we describe how possibility theory fits in the framework of VBS.

The framework of VBS was first defined by Shenoy [1989] as a general language for incorporating uncertainty in expert systems. It was further elaborated in [Shenoy 1991c] to include axioms that permit local computation in solving a VBS, and a fusion algorithm for solving a VBS using local computation. VBS encode knowledge using functions defined on frames of variables. The functions are called valuations. VBS include two operators called combination and marginalization that operate on valuations. Combination corresponds to aggregation of knowledge. Marginalization corresponds to coarsening of knowledge. The process of reasoning in VBS can be described simply as finding the marginal of the joint valuation for each variable in the system. The joint valuation is the valuation obtained by combining all valuations. In systems with many variables, it is computationally intractable to explicitly compute the joint valuation. However, if combination and marginalization satisfy certain axioms, it is possible to compute the marginals of the joint valuation without explicitly computing the joint valuation.

The framework of VBS is general enough to represent many domains such as Bayesian probability theory [Shenoy 1991c], Dempster-Shafer's theory of belief functions [Dempster 1967, Shafer 1976, Shenoy 1991d], Spohn's theory of epistemic beliefs [Spohn 1988, Shenoy 1991c], discrete optimization [Shenoy 1991b], propositional logic [Shenoy 1990a, 1990b], constraint satisfaction [Shenoy and Shafer 1988], and Bayesian decision theory [Shenoy 1990c, 1990d]. Saffiotti and Umkehrer [1991] describe an efficient implementation of VBS called Pulcinella.

The correspondence between possibility theory and VBS is as follows. The particularization operation in possibility theory corresponds to the combination operation in VBS. And the projection operation in possibility theory corresponds to the marginalization operation in VBS. This cor-

responsiveness was first pointed out explicitly by Dubois and Prade [1991]. They also state that the combination and marginalization operations in possibility theory satisfy the three axioms that enable the use of local computation in computing marginals of the joint possibility function. In this paper, we reiterate this correspondence in greater detail.

An outline of this paper is as follows. Section 2 describes the framework of VBS. Section 3 contains an axiomatic definition and a characterization of a consistent possibilistic state. Section 4 describes the main features of possibility theory in terms of the framework of VBS. Section 5 describes a small example illustrating the use of possibility theory for managing uncertainty and imprecision in expert systems. Section 6 contains some concluding remarks. Finally, section 7 contains proofs of all results in the paper.

2. VALUATION-BASED SYSTEMS

In this section, we will sketch the basic features of VBS. Also, we describe three axioms that permit the use of local computation, and describe a fusion algorithm for solving a VBS using local computation.

2.1. The Framework

This subsection describes the framework of valuation-based systems. In a VBS, we represent knowledge by functions called valuations. Valuations are functions that assign values to the elements of frames for sets of variables. We make inferences in a VBS using two operators called combination and marginalization. We use these operators on valuations.

Variables and Configurations. We use the symbol \mathcal{W}_X for the set of possible values of a variable X , and we call \mathcal{W}_X the *frame for X* . We assume that one and only one of the elements of \mathcal{W}_X is the true value of X . We are concerned with a finite set \mathcal{X} of variables, and we assume that all the variables in \mathcal{X} have finite frames.

Given a nonempty set s of variables, let \mathcal{W}_s denote the Cartesian product of \mathcal{W}_X for X in s ; $\mathcal{W}_s = \times\{\mathcal{W}_X \mid X \in s\}$. We call \mathcal{W}_s the *frame for s* . We call the elements of \mathcal{W}_s *configurations of s* .

Valuations. Given a subset s of variables, there is a set \mathcal{V}_s . We call the elements of \mathcal{V}_s *valuations for s* . Let \mathcal{V} denote the set of all valuations, i.e., $\mathcal{V} = \cup\{\mathcal{V}_s \mid s \subseteq \mathcal{X}\}$. If σ is a valuation for s , then we say s is the *domain* of σ .

Valuations are primitives in our abstract framework and as such require no definition. But as we shall see shortly, they are objects which can be combined and marginalized. Intuitively, a valuation for s represents some knowledge about the variables in s .

Nonzero valuations. For each $s \subseteq \mathcal{X}$, there is a nonempty subset \mathcal{P}_s of \mathcal{V}_s whose elements are called *nonzero valuations for s* . Let \mathcal{P} denote $\cup\{\mathcal{P}_s \mid s \subseteq \mathcal{X}\}$, the set of all nonzero valuations.

Intuitively, a nonzero valuation represents knowledge that is internally consistent. The notion of nonzero valuations is important as it enables us to define combinability of valuations, and it

allows us to constrain the definitions of combination and marginalization to meaningful operators. An example of a nonzero valuation is a possibility function.

Combination. We assume there is a mapping $\otimes: \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}$, called *combination*, such that

- (i) if ρ and σ are valuations for r and s , respectively, then $\rho \otimes \sigma$ is a valuation for $r \cup s$;
- (ii) if either ρ or σ is not a nonzero valuation, then $\rho \otimes \sigma$ is not a nonzero valuation;
- and
- (iii) if ρ and σ are both nonzero valuations, then $\rho \otimes \sigma$ may or may not be a nonzero valuation.

We call $\rho \otimes \sigma$ the *combination of ρ and σ* .

Intuitively, combination corresponds to aggregation of knowledge. If ρ and σ are valuations for r and s representing knowledge about variables in r and s , respectively, then $\rho \otimes \sigma$ represents the aggregated knowledge about variables in $r \cup s$. For possibility functions, combination is pointwise multiplication [Zadeh 1965].

Marginalization. We assume that for each $s \subseteq \mathcal{X}$, and for each $X \in s$, there is a mapping $\downarrow(s-\{X\}): \mathcal{V}_s \rightarrow \mathcal{V}_{s-\{X\}}$, called *marginalization to $s-\{X\}$* such that

- (i) If σ is a valuation for s , then $\sigma^{\downarrow(s-\{X\})}$ is a valuation for $s-\{X\}$; and
- (ii) $\sigma^{\downarrow(s-\{X\})}$ is a nonzero valuation if and only if σ is a nonzero valuation.

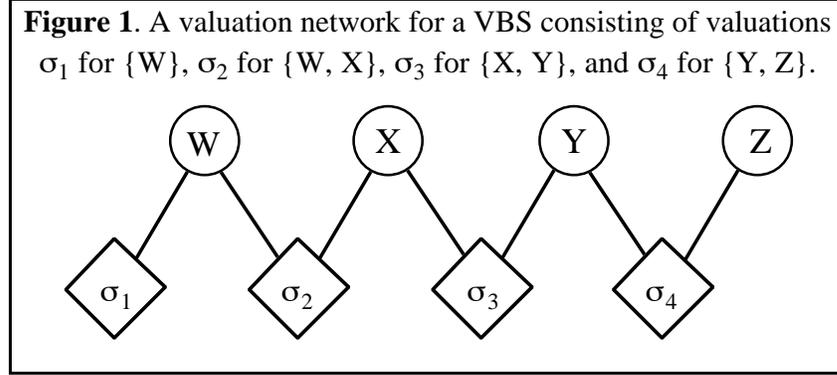
We call $\sigma^{\downarrow(s-\{X\})}$ the *marginal of σ for $s-\{X\}$* .

Intuitively, marginalization corresponds to coarsening of knowledge. If σ is a valuation for s representing some knowledge about variables in s , and $X \in s$, then $\sigma^{\downarrow(s-\{X\})}$ represents the knowledge about variables in $s-\{X\}$ implied by σ if we disregard variable X . In the case of possibility functions, marginalization from s to $s-\{X\}$ is maximization over the frame for X [Zadeh 1979].

In summary, a *valuation-based system* consists of a 3-tuple $\{\{\sigma_1, \dots, \sigma_m\}, \otimes, \downarrow\}$ where $\{\sigma_1, \dots, \sigma_m\}$ is a collection of valuations, \otimes is the combination operator, and \downarrow is the marginalization operator.

Valuation Networks. A graphical depiction of a valuation-based system is called a *valuation network*. In a valuation network, variables are represented by circular nodes, and valuations are represented by diamond-shaped nodes. Also, each valuation node is connected by an undirected edge to each variable node in its domain. Figure 1 shows a valuation network for a VBS that consists of valuations σ_1 for $\{W\}$, σ_2 for $\{W, X\}$, σ_3 for $\{X, Y\}$, and σ_4 for $\{Y, Z\}$.

Making Inference in VBS. In a VBS, the combination of all valuations is called the *joint valuation*. Given a VBS, we make inferences by computing the marginal of the joint valuation for each variable in the system.



If there are n variables in the system, and each variable has two configurations in its frame, then there are 2^n configurations of all variables. Hence, it is not computationally feasible to compute the joint valuation when there are a large number of variables. In Section 2.3, we describe an algorithm for computing marginals of the joint valuation without explicitly computing the joint valuation. So that this algorithm gives us the correct answers, we require combination and marginalization to satisfy three axioms. The axioms and the algorithm are described in the next two subsections, respectively.

2.2. Axioms

In this section, we state three simple axioms that enable local computation of marginals of the joint valuation. These axioms were first formulated by Shenoy and Shafer [1990]. The axioms are stated slightly differently here.

Axiom A1 (*Commutativity and associativity of combination*): Suppose ρ , σ , and τ are valuations for r , s , and t respectively. Then

$$\rho \otimes \sigma = \sigma \otimes \rho, \text{ and } \rho \otimes (\sigma \otimes \tau) = (\rho \otimes \sigma) \otimes \tau.$$

Axiom A2 (*Order of deletion does not matter*): Suppose σ is a valuation for s , and suppose $X_1, X_2 \in s$. Then

$$(\sigma \downarrow^{(s-\{X_1\})}) \downarrow^{(s-\{X_1, X_2\})} = (\sigma \downarrow^{(s-\{X_2\})}) \downarrow^{(s-\{X_1, X_2\})}.$$

Axiom A3 (*Distributivity of marginalization over combination*): Suppose ρ and σ are valuations for r and s , respectively. Suppose $X \notin r$, and suppose $X \in s$. Then

$$(\rho \otimes \sigma) \downarrow^{((r \cup s) - \{X\})} = \rho \otimes (\sigma \downarrow^{(s - \{X\})}).$$

One implication of Axiom A1 is that when we have multiple combinations of valuations, we can write it without using parenthesis. For example, $(\dots((\sigma_1 \otimes \sigma_2) \otimes \sigma_3) \otimes \dots \otimes \sigma_m)$ can be written simply as $\otimes\{\sigma_i \mid i = 1, \dots, m\}$ or as $\sigma_1 \otimes \dots \otimes \sigma_m$, i.e., we need not indicate the order in which the combinations are carried out.

If we regard marginalization as a coarsening of a valuation by deleting variables, then axiom A2 says that the order in which the variables are deleted does not matter. One implication of this axiom is that $(\sigma^{\downarrow(s-\{X_1\})})^{\downarrow(s-\{X_1, X_2\})}$ can be written simply as $\sigma^{\downarrow(s-\{X_1, X_2\})}$, i.e., we need not indicate the order in which the variables are deleted.

Axiom A3 is the crucial axiom that makes local computation possible. Axiom A3 states that the computation of $(\rho \otimes \sigma)^{\downarrow((r \cup s) - \{X\})}$ can be accomplished without having to compute $\rho \otimes \sigma$. The combination operation in $\rho \otimes \sigma$ is on the frame for $r \cup s$ whereas the combination operation in $\rho \otimes (\sigma^{\downarrow(s-\{X\})})$ is on the frame for $(r \cup s) - \{X\}$. In the next subsection, we describe a fusion algorithm that applies this axiom repeatedly resulting in an efficient method for computing marginals.

2.3. A Fusion Algorithm

In this subsection, we describe a fusion algorithm for making inferences in a VBS using local computation. Suppose $\{\{\sigma_1, \dots, \sigma_m\}, \otimes, \downarrow\}$ is a VBS with n variables and m valuations. Suppose that combination and marginalization satisfy the three axioms stated in section 2.2. Suppose we have to compute the marginal of the joint valuation for variable X , $(\sigma_1 \otimes \dots \otimes \sigma_m)^{\downarrow\{X\}}$. The basic idea of the fusion algorithm is to successively delete all variables but X from the VBS. The variables may be deleted in any sequence. Axiom A 2 tells us that all deletion sequences lead to the same answers. But, different deletion sequences may involve different computational costs. We will comment on good deletion sequences at the end of this section.

When we delete a variable, we have to do a “fusion” operation on the valuations. Consider a set of k valuations ρ_1, \dots, ρ_k . Suppose ρ_i is a valuation for r_i . Let $\text{Fus}_X\{\rho_1, \dots, \rho_k\}$ denote the collection of valuations after fusing the valuations in the set $\{\rho_1, \dots, \rho_k\}$ with respect to variable X . Then

$$\text{Fus}_X\{\rho_1, \dots, \rho_k\} = \{\rho^{\downarrow(r-\{X\})}\} \cup \{\rho_i \mid X \notin r_i\}$$

where $\rho = \otimes\{\rho_i \mid X \in r_i\}$, and $r = \cup\{r_i \mid X \in r_i\}$. After fusion, the set of valuations is changed as follows. All valuations that bear on X are combined, and the resulting valuation is marginalized such that X is eliminated from its domain. The valuations that do not bear on X remain unchanged.

We are ready to state the main theorem that describes the fusion algorithm

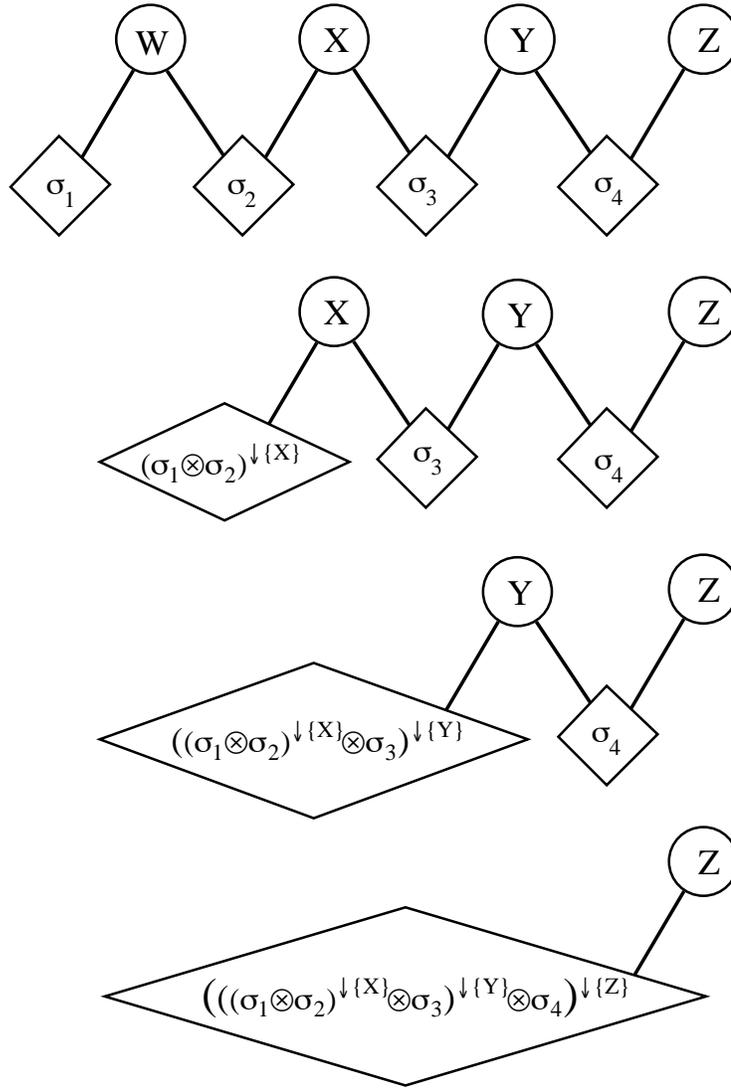
Theorem 1 [Shenoy 1991c]. Suppose $\{\{\sigma_1, \dots, \sigma_m\}, \otimes, \downarrow\}$ is a VBS where σ_i is a valuation for s_i , and suppose \otimes and \downarrow satisfy axioms A1–A3. Let \mathcal{X} denote $s_1 \cup \dots \cup s_m$. Suppose $X \in \mathcal{X}$, and suppose $X_1 X_2 \dots X_{n-1}$ is a sequence of variables in $\mathcal{X} - \{X\}$. Then

$$(\sigma_1 \otimes \dots \otimes \sigma_m)^{\downarrow\{X\}} = \otimes \left\{ \text{Fus}_{X_{n-1}} \left\{ \dots \text{Fus}_{X_2} \left\{ \text{Fus}_{X_1} \{ \sigma_1, \dots, \sigma_m \} \right\} \right\} \right\}.$$

To illustrate Theorem 1, consider the VBS shown in Figure 1. Suppose we need to compute the marginal of the joint for Z , $(\sigma_1 \otimes \dots \otimes \sigma_4)^{\downarrow\{Z\}}$. Consider the deletion sequence WXY . After

fusion with respect to W , we have $(\sigma_1 \otimes \sigma_2)^{\downarrow\{X\}}$ for $\{X\}$, σ_3 for $\{X, Y\}$, and σ_4 for $\{Y, Z\}$. After fusion with respect to X , we have $((\sigma_1 \otimes \sigma_2)^{\downarrow\{X\}} \otimes \sigma_3)^{\downarrow\{Y\}}$ for $\{Y\}$, and σ_4 for $\{Y, Z\}$. Finally, after fusion with respect to Y , we have $((\sigma_1 \otimes \sigma_2)^{\downarrow\{X\}} \otimes \sigma_3)^{\downarrow\{Y\}} \otimes \sigma_4)^{\downarrow\{Z\}}$ for $\{Z\}$. Theorem 1 tells us that $((\sigma_1 \otimes \sigma_2)^{\downarrow\{X\}} \otimes \sigma_3)^{\downarrow\{Y\}} \otimes \sigma_4)^{\downarrow\{Z\}} = (\sigma_1 \otimes \dots \otimes \sigma_4)^{\downarrow\{Z\}}$. Thus, instead of doing combinations on the frame for $\{W, X, Y, Z\}$, we do combinations on the frame for $\{W, X\}$, $\{X, Y\}$, and $\{Y, Z\}$. The fusion algorithm is shown graphically in Figure 2.

Figure 2. The first valuation network shows the initial VBS. The second network is the result after fusion with respect to W . The third network is the result after fusion with respect to X . The fourth network is the result after fusion with respect to Y .



If we can compute the marginal of the joint valuation for one variable, then we can compute the marginals for all variables. We simply compute them one after the other. It is obvious, however,

that this will involve much duplication of effort. Shenoy and Shafer [1990] describe an efficient algorithm for simultaneous computation of all marginals without duplication of effort. Regardless of the number of variables in a VBS, we can compute marginals of the joint valuation for all variables for roughly three times the computational effort required to compute one marginal [Shenoy and Shafer 1990].

Deletion Sequences. Different deletion sequences may involve different computational efforts. For example, consider the VBS in the above example. In this example, deletion sequence WXY involves less computational effort than, for example, XYW, as the former involves combinations on the frame for two variables only whereas the latter involves combination on the frame for three variables. Finding an optimal deletion sequence is a secondary optimization problem that has shown to be NP-complete [Arnborg et al. 1987]. But, there are several heuristics for finding good deletion sequences [Kong 1986, Mellouli 1987, Zhang 1988, Kjærulff 1990].

One such heuristic is called one-step-look-ahead [Kong 1986]. This heuristic tells us which variable to delete next. As per this heuristic, the variable that should be deleted next is one that leads to combination over the smallest frame. For example, in the VBS described above, if we assume that each variable has a frame consisting of two configurations, then this heuristic would pick W over X and Y for first deletion since deletion of W involves combination on the frame for {W, X} whereas deletion of X involves combination on the frame for {W, X, Y}, and deletion of Y involves combination on the frame for {X, Y, Z}. After W is deleted, for second deletion, this heuristic would pick X over Y. Thus, this heuristic would choose deletion sequence WXY.

3. CONSISTENT POSSIBILISTIC STATES

In this section, first we define axiomatically a consistent possibilistic state. A consistent possibilistic state serves as a qualitative description of a possibility function. The definition of a consistent possibilistic state is analogous to the definition of a consistent epistemic state in Spohn's theory of epistemic beliefs [Shenoy 1991a, Spohn 1988]. Next, we give a characterization of a consistent possibilistic state. Again, this characterization is analogous to the characterization of a consistent epistemic state given by Spohn [1988]. Before we define a consistent possibilistic state, we need to establish our notation.

Consider a variable X with frame \mathcal{W}_X . Suppose \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 s,' where s 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 [Shafer 1976, p. 36].

The correspondence between subsets and propositions is useful since it translates the logical notions of conjunction, disjunction, implication, and negation into the set-theoretic notions of intersection, union, inclusion, and complementation [Shafer 1976, pp. 36-37]. Thus, if r and s are two subsets of \mathcal{W}_X , and r' and s' are the corresponding propositions, then $r \cap s$ corresponds to the

conjunction of r' and s' , $r \cup s$ corresponds to the disjunction of r' and s' , $r \subseteq s$ if and only if r' implies s' , and r is the set-theoretic complement of s with respect to \mathcal{W}_X (written as $r = \sim s$) if and only if r' is the negation of s' . Notice also that the proposition that corresponds to \emptyset is false, and the proposition that corresponds to \mathcal{W}_X is true. 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 a possibilistic state for X , propositions are said to be either possibly true (or simply, *possible*) or possibly false (or simply, *not possible*). Thus a possibilistic state for X is an assignment of labels ‘possible’ and ‘not possible’ to each proposition in \mathcal{W}_X . Logical consistency requires that a possibilistic state satisfy certain conditions (axioms). A definition of a *consistent* possibilistic state is as follows.

Definition 1. A possibilistic state for X is said to be *consistent* if the following four axioms are satisfied:

B1. For any proposition s , one and only one of the following conditions holds:

- (i) s is possible.
- (ii) s is not possible.

B2. \mathcal{W}_X is possible, and \emptyset is not possible.

B3. If s is not possible and $r \subseteq s$, then r is not possible.

B4. If r and s are not possible, then $r \cup s$ is not possible.

Some simple consequences of Definition 1 are as follows.

Proposition 1. The following conditions always hold in any consistent possibilistic state:

B5. If s is possible and $r \supseteq s$, then r is possible.

B6. If s is not possible, then $\sim s$ is possible.

B7. If s is possible, then this does not necessarily imply that $\sim s$ is not possible; $\sim s$ could also be possible.

It is clear from axiom B1 that we can specify a consistent possibilistic state simply by listing all propositions that are not possible. Then axiom B1 tells us that the remaining propositions are possible. Let \acute{u} denote the set of all propositions (subsets) that are not possible. Theorem 2 gives a characterization of \acute{u} .

Theorem 2. Suppose \acute{u} denotes the set of all propositions that are not possible in some possibilistic state for X . Then the possibilistic state is consistent if and only if there exists a unique proper subset c of \mathcal{W}_X such that $\acute{u} = \{s \in 2^{\mathcal{W}_X} \mid s \subseteq c\}$.

Subset c in Theorem 2 is called the *content* of the consistent possibilistic state. Note that the content constitutes a complete specification of a possibilistic state. Thus a simple corollary of

Theorem 2 is that in a frame \mathcal{W}_X consisting of n elements, there are exactly $2^n - 1$ distinct possible consistent possibilistic states (corresponding to each proper subset of \mathcal{W} as the content). The content c represents the largest subset (proposition) that is not possible.

Example 1. (*Consistent possibilistic states*) Consider a frame $\mathcal{W} = \{x, y, z\}$. Table 1 lists all seven possible consistent possibilistic states. Possibilistic state 1 represents a state of complete ignorance in which every proposition (except the null proposition) is possible. Possibilistic states 5, 6, and 7 represent a state of complete information where one and only one element of the frame is possibly true. Possibilistic states 2, 3, and 4 are states of partial information where one element of the frame is not possible

Table 1. The seven consistent possibilistic states for a variable with frame $\mathcal{W} = \{x, y, z\}$.

State	Content c	Possible	Not Possible \bar{c}
1	\emptyset	$\{x\}, \{y\}, \{z\}, \{x, y\}, \{y, z\}, \{x, z\}, \{x, y, z\}$	\emptyset
2	$\{x\}$	$\{y\}, \{z\}, \{x, y\}, \{x, z\}, \{y, z\}, \{x, y, z\}$	$\emptyset, \{x\}$
3	$\{y\}$	$\{x\}, \{z\}, \{x, y\}, \{x, z\}, \{y, z\}, \{x, y, z\}$	$\emptyset, \{y\}$
4	$\{z\}$	$\{x\}, \{y\}, \{x, y\}, \{x, z\}, \{y, z\}, \{x, y, z\}$	$\emptyset, \{z\}$
5	$\{x, y\}$	$\{z\}, \{x, z\}, \{y, z\}, \{x, y, z\}$	$\emptyset, \{x\}, \{y\}, \{x, y\}$
6	$\{x, z\}$	$\{y\}, \{x, y\}, \{y, z\}, \{x, y, z\}$	$\emptyset, \{x\}, \{z\}, \{x, z\}$
7	$\{y, z\}$	$\{x\}, \{x, y\}, \{x, z\}, \{x, y, z\}$	$\emptyset, \{y\}, \{z\}, \{y, z\}$

4. POSSIBILITY THEORY

In this section, we describe the main features of Zadeh's theory of possibility in terms of the framework of VBS described in the previous section, i.e., we will use the terminology of VBS instead of the terminology used in [Zadeh 1979]. Thus the operation Zadeh calls particularization we call combination, and the operation Zadeh calls projection we call marginalization. The basic unit of knowledge representation is called a possibility function.

Definition 2. A possibility function π for p is a function $\pi: 2^{\mathcal{W}_p} \rightarrow [0, 1]$ such that

- P1. There exists a configuration $\mathbf{w} \in \mathcal{W}_p$ such that $\pi(\{\mathbf{w}\}) = 1$;
- P2. For any $r \in (2^{\mathcal{W}_g} - \{\emptyset\})$, $\pi(r) = \text{MAX}\{\pi(\{\mathbf{w}\}) \mid \mathbf{w} \in r\}$; and
- P3. $\pi(\emptyset) = 0$.

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

A possibility function is a complete representation of a consistent possibilistic state. To see what propositions are possible and what propositions are not possible in state π , consider the subset $c = \{\mathbf{w} \in \mathcal{W}_p \mid \pi(\{\mathbf{w}\}) < 1\}$. By condition P1 in Definition 2, c is always a proper subset of \mathcal{W}_p . c represents the content of the possibilistic state π , i.e., r is not possible in state π iff $r \subseteq c$. Thus r is possible in state π iff $\pi(r) = 1$, and r is not possible in state π iff $\pi(r) < 1$.

A possibility function consists of more than a representation of a consistent possibilistic state. It also includes degrees to which proposition are possible and degrees to which propositions are not possible. $\pi(r)$ can be interpreted as the degree to which proposition r is possible, and $1 - \pi(r)$ can be interpreted as the degree to which proposition r is not possible, i.e., r is more possible than t if $\pi(r) > \pi(t)$ and conversely, r is more impossible than t if $\pi(r) < \pi(t) < 1$.

Consider the following possibility function for s : $\pi(\{\mathbf{w}\}) = 1$ for all $\mathbf{w} \in \mathcal{W}_s$. This means that the only proposition that is not possible is \emptyset (and all other propositions are possible). We shall call such a possibility function *vacuous*. It represents a state of complete ignorance.

Proposition 2. Suppose π is a possibility function for p . Then

P4. For each $r \in 2^{\mathcal{W}_p}$, either $\pi(r) = 1$ or $\pi(\sim r) = 1$ or both.

P5. For each $r, t \in 2^{\mathcal{W}_p}$, $\pi(r \cup t) = \text{MAX}\{\pi(r), \pi(t)\}$

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

Marginalization. Suppose π is a possibility function for p . Suppose $q \subseteq p$. We may be interested only in propositions about variables in q . In this case, we would like to marginalize π to q . The following definition of marginalization is motivated by the fact that each proposition $q \in 2^{\mathcal{W}_q}$ about variables in q can be regarded as a proposition $q^{\uparrow p} \in 2^{\mathcal{W}_p}$ about variables in p .

Definition 3 (Marginalization). Suppose π is a possibility function for p , and suppose $q \subseteq p$. The *marginal of π for q* , denoted by $\pi^{\downarrow q}$, is a possibility function for q given as follows:

$$\pi^{\downarrow q}(q) = \pi(q^{\uparrow p}) = \text{MAX}\{\pi(\{\mathbf{x}, \mathbf{y}\}) \mid \mathbf{x} \in q, \mathbf{y} \in \mathcal{W}_{p-q}\} \quad (1)$$

for all $q \in 2^{\mathcal{W}_q}$.

In particular, if q is a singleton subset, i.e., $q = \{\mathbf{x}\}$ for some $\mathbf{x} \in \mathcal{W}_q$, then (1) simplifies to:

$$\pi^{\downarrow q}(\{\mathbf{x}\}) = \text{MAX} \left\{ \pi(\{(\mathbf{x}, \mathbf{y})\}) \mid \mathbf{y} \in \mathcal{W}_{p-q} \right\}.$$

Note that if π is a possibility function for p , and $X_1, X_2 \in p$, then $(\pi^{\downarrow(p-\{X_1\})})^{\downarrow(p-\{X_1, X_2\})} = (\pi^{\downarrow(p-\{X_2\})})^{\downarrow(p-\{X_1, X_2\})}$. In words, if we regard marginalization as reduction of π by deletion of variables, then the order in which the variables are deleted makes no difference in the final answer.

Example 2. (Possibility function and marginalization) We would like to determine whether a stranger (about who we know nothing about) is a pacifist or not depending on whether he is a Republican or not and whether he is a Quaker or not. Consider three variables R, Q and P . R has two configurations: r (for Republican), $\sim r$ (not Republican); Q has two configurations: q (Quaker) and $\sim q$ (not Quaker); and P has two configurations: p (pacifist) and $\sim p$ (not pacifist). Our knowledge that most (at least 80 percent) Republicans are not pacifists and that most (at least 90 percent) Quakers are pacifists is represented by the possibility function π for $\{R, Q, P\}$ shown in Table 2 (the construction of this possibility function will be explained later in this section—see Example 3).

Note that the marginal of π for R is the vacuous possibility function for R , i.e., $\pi^{\downarrow\{R\}}(\{r\}) = \pi^{\downarrow\{R\}}(\{\sim r\}) = 1$. Thus in possibilistic state π , we have no knowledge whether the stranger is a republican or not. Similarly, notice that the marginals of π for $\{Q\}$ and $\{P\}$ are also vacuous. The marginal of π for $\{R, P\}$ is as follows:

$$\pi^{\downarrow\{R, P\}}(\{r, p\}) = 0.2, \pi^{\downarrow\{R, P\}}(\{r, \sim p\}) = \pi^{\downarrow\{R, P\}}(\{\sim r, p\}) = \pi^{\downarrow\{R, P\}}(\{\sim r, \sim p\}) = 1.$$

Thus pacifist Republicans are not possible (to degree 0.8). Similarly, notice that the marginal of π for $\{Q, P\}$ is as follows:

$$\pi^{\downarrow\{Q, P\}}(\{q, \sim p\}) = 0.1, \pi^{\downarrow\{Q, P\}}(\{q, p\}) = \pi^{\downarrow\{Q, P\}}(\{\sim q, p\}) = \pi^{\downarrow\{Q, P\}}(\{\sim q, \sim p\}) = 1.$$

Thus non-pacifist Quakers are not possible (to degree 0.9). Finally note that the marginal of π for $\{R, Q\}$ is as follows:

$$\pi^{\downarrow\{R, Q\}}(\{r, q\}) = 0.2, \pi^{\downarrow\{R, Q\}}(\{r, \sim q\}) = \pi^{\downarrow\{R, Q\}}(\{\sim r, q\}) = \pi^{\downarrow\{R, Q\}}(\{\sim r, \sim q\}) = 1.$$

Thus Republican Quakers are not possible (to degree 0.8).

Table 2. A possibility function π for $\{R,Q,P\}$.

$\mathcal{W}_{\{R,P,Q\}}$			π
r	p	q	0.2
r	p	\sim q	0.2
r	\sim p	q	0.1
r	\sim p	\sim q	1.0
\sim r	p	q	1.0
\sim r	p	\sim q	1.0
\sim r	\sim p	q	0.1
\sim r	\sim p	\sim q	1.0

Next, we state a rule for combining possibility function. This rule is called particularization by Zadeh [1979]. Zadeh [1965] has given two rules for combining possibility functions—minimization and multiplication. Axiom B2 (\mathcal{W}_X is possible) requires that we include normalization in the definition of combination. Since minimization followed by normalization is not associative, multiplication followed by normalization is associative, and associativity of combination is a requirement in VBS (axiom A1), we use multiplication followed by normalization as our definition of combination.

Definition 4. Suppose π_1 and π_2 are possibility functions for p_1 and p_2 , respectively.

Suppose $K = \text{MAX} \{ \pi_1(\{\mathbf{w}^{\downarrow p_1}\}) \pi_2(\{\mathbf{w}^{\downarrow p_2}\}) \mid \mathbf{w} \in \mathcal{W}_{p_1 \cup p_2} \}$. The *combination* of π_1 and π_2 , denoted by $\pi_1 \otimes \pi_2$, is the function for $p_1 \cup p_2$ given by

$$(\pi_1 \otimes \pi_2)(\{\mathbf{w}\}) = \begin{cases} K^{-1} \pi_1(\{\mathbf{w}^{\downarrow p_1}\}) \pi_2(\{\mathbf{w}^{\downarrow p_2}\}) & \text{if } K \neq 0 \\ 0 & \text{if } K = 0 \end{cases} \quad (2)$$

for all $\mathbf{w} \in \mathcal{W}_{p_1 \cup p_2}$.

If $K = 0$, then $\pi_1 \otimes \pi_2$ is not a possibility function. In this case, $\pi_1 \otimes \pi_2$ is not a nonzero valuation. This means that the knowledge in π_1 and π_2 are inconsistent. If $K \neq 0$, then K is a normalization constant that ensures that $\pi_1 \otimes \pi_2$ is a possibility function. Thus, combination consists of pointwise multiplication followed by normalization (if normalization is possible).

Example 3. (The rule of combination) Consider two pieces of knowledge as follows:

1. Most Republicans (at least 80 percent) are not pacifists.
2. Most Quakers (at least 90 percent) are pacifists.

If π_1 is a possibility function representation of the first piece of knowledge, and π_2 is a possibility function representation of the second piece of knowledge, then $\pi_1 \otimes \pi_2$ will represent the aggregation of these two pieces of knowledge. In particular, suppose π_1 is a possibility function for $\{R,P\}$ as follows:

$$\pi_1(\{(r,p)\}) = 0.2, \pi_1(\{(r,\sim p)\}) = \pi_1(\{(\sim r,p)\}) = \pi_1(\{(\sim r,\sim p)\}) = 1$$

(i.e., pacifist Republicans are not possible to degree 0.8), and suppose π_2 is a possibility function for $\{Q,P\}$ as follows:

$$\pi_2(\{(q,\sim p)\}) = 0.1, \pi_2(\{(q,p)\}) = \pi_2(\{(\sim q,p)\}) = \pi_2(\{(\sim q,\sim p)\}) = 1$$

(i.e., non-pacifist Quakers are not possible to degree 0.9). Then the possibility function $\pi_1 \otimes \pi_2 = \pi$, say, shown in Table 3, represents the aggregate knowledge.

Table 3. The combination of π_1 and π_2 .

$W_{\{R,P,Q\}}$	π_1	π_2	$\pi_1 \otimes \pi_2 = \pi$
r p q	0.2	1.0	0.2
r p \sim q	0.2	1.0	0.2
r \sim p q	1.0	0.1	0.1
r \sim p \sim q	1.0	1.0	1.0
\sim r p q	1.0	1.0	1.0
\sim r p \sim q	1.0	1.0	1.0
\sim r \sim p q	1.0	0.1	0.1
\sim r \sim p \sim q	1.0	1.0	1.0

Proposition 3. The rule of combination described in (2) has the following properties:

- C1. (Commutativity) $\pi_1 \otimes \pi_2 = \pi_2 \otimes \pi_1$.
- C2. (Associativity) $(\pi_1 \otimes \pi_2) \otimes \pi_3 = \pi_1 \otimes (\pi_2 \otimes \pi_3)$.
- C3. If π_1 is vacuous, then $\pi_1 \otimes \pi_2 = \pi_2$.
- C4. In general, $\pi_1 \otimes \pi_1 \neq \pi_1$.

The possibility function $\pi_1 \otimes \pi_2$ represents the aggregation of knowledge contained in possibility functions π_1 and π_2 if π_1 and π_2 are independent. Like Bayesian probability theory, Dempster-Shafer's theory of belief function, and Spohn's theory of epistemic beliefs, possibility theory requires that the possibility functions π_1 and π_2 be independent. This is because of property C4—double-counting of knowledge may lead to erroneous results.

We have already shown that axioms A1 and A2 are valid for possibility functions. Theorem 3 below states that axiom A3 is also satisfied. This result is stated in [Dubois and Prade 1991].

Theorem 3. Suppose π_1 and π_2 are possibility functions for p_1 and p_2 , respectively. Suppose $X \notin p_1$, and suppose $X \in p_2$. Then $(\pi_1 \otimes \pi_2)^{\downarrow((p_1 \cup p_2) - \{X\})} = \pi_1 \otimes (\pi_2^{\downarrow(p_2 - \{X\})})$.

Since all three axioms required for local computation of marginals are satisfied, the fusion algorithm described in section 2.3 can be used for reasoning from knowledge expressed as possibility functions. In the context of possibility theory, [Chatalic et al. 1987] describes a local computational method for reasoning with possibility functions. The next section describes a small example to illustrate the use of the fusion algorithm to find marginals.

5. AN EXAMPLE

In this section, we describe an example in complete detail to illustrate the use of possibility theory in managing imprecision in expert systems.

Is Dick a Pacifist? Consider the following items of evidence. Most Republicans (at least 80 percent) are not pacifists. Most Quakers (at least 90 percent) are pacifists. Dick is a Republican (and we are more than 99 percent certain of this). Dick is a Quaker (and we are more than 99 percent certain of this). Is Dick a pacifist or not?

We will model the four items of evidence as possibility functions π_1 for $\{R, P\}$, π_2 for $\{Q, P\}$, π_3 for $\{R\}$, and π_4 for $\{Q\}$, respectively, as displayed in Table 4. Figure 3 shows the valuation network for this example. If we apply the fusion algorithm using deletion sequence RQ, we get $(\pi_1 \otimes \pi_2 \otimes \pi_3 \otimes \pi_4)^{\downarrow\{P\}} = (\pi_1 \otimes \pi_3)^{\downarrow\{P\}} \otimes (\pi_2 \otimes \pi_4)^{\downarrow\{P\}}$. The details of the computations are shown in Table 5. The conclusion is that the proposition *Dick is a pacifist* is possible (to degree 1), and the proposition *Dick is not a pacifist* is not possible (to degree 0.5).

Table 4. The possibility functions $\pi_1, \pi_2, \pi_3,$ and π_4 in the *Is Dick a Pacifist?* example.

$\mathcal{W}_{\{R,P\}}$		π_1
r	p	0.2
r	$\sim p$	1
$\sim r$	p	1
$\sim r$	$\sim p$	1

$\mathcal{W}_{\{Q,P\}}$		π_2
q	p	1
q	$\sim p$	0.1
$\sim q$	p	1
$\sim q$	$\sim p$	1

$\mathcal{W}_{\{R\}}$		π_3
r		1
$\sim r$.01

$\mathcal{W}_{\{Q\}}$		π_4
q		1
$\sim q$.01

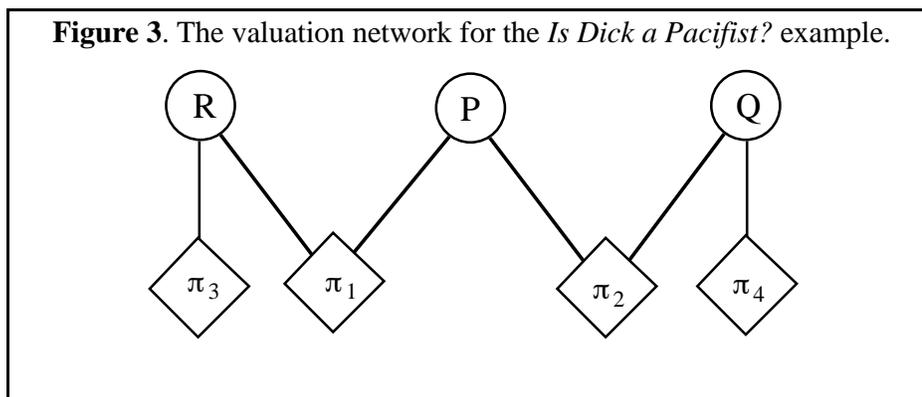


Table 5. The computation of $(\pi_1 \otimes \pi_3)^{\downarrow\{P\}} \otimes (\pi_2 \otimes \pi_4)^{\downarrow\{P\}}$.

$\mathcal{W}_{\{R,P\}}$		$\pi_1 \otimes \pi_3$
r	p	0.20
r	$\sim p$	1.00
$\sim r$	p	0.01
$\sim r$	$\sim p$	0.01

$\mathcal{W}_{\{Q,P\}}$		$\pi_2 \otimes \pi_4$
q	p	1.00
q	$\sim p$	0.10
$\sim q$	p	0.01
$\sim q$	$\sim p$	0.01

\mathcal{W}_P	$(\pi_1 \otimes \pi_3)^{\downarrow\{P\}}$	$(\pi_2 \otimes \pi_4)^{\downarrow\{P\}}$	$(\pi_1 \otimes \pi_3)^{\downarrow\{P\}} \otimes (\pi_2 \otimes \pi_4)^{\downarrow\{P\}}$
p	0.20	1.00	1.00
$\sim p$	1.00	0.10	0.50

6. CONCLUSIONS

We have given an axiomatic definition of a qualitative description of a possibility function called a consistent possibilistic state. We have also given a characterization of a consistent possibilistic state in terms of its content. Our objective here is to help develop semantics for possibility theory. Our contribution in this paper is only one step in this direction.

We have also described how possibility theory fits in the framework of valuation-based systems. This was first described by Dubois and Prade [1991]. We have expanded on this topic by providing a brief yet complete description of the VBS framework, and by describing the main features of possibility theory in terms of the VBS framework. We have also described a small example in complete detail to illustrate the use of possibility theory in managing uncertainty and imprecision in expert systems.

7. PROOFS

In this section, we give proofs for all results in the paper. First we state and prove a lemma needed to prove Theorem 1.

Lemma 1. Suppose $\{\{\sigma_1, \dots, \sigma_m\}, \otimes, \downarrow\}$ is a VBS where σ_i is a valuation for s_i , and suppose \otimes and \downarrow satisfy axioms A1–A3. Let \mathcal{X} denote $s_1 \cup \dots \cup s_m$. Suppose $X \in \mathcal{X}$. Then

$$(\sigma_1 \otimes \dots \otimes \sigma_m) \downarrow^{(\mathcal{X} - \{X\})} = \otimes \text{Fus}_{\mathcal{X}} \{\sigma_1, \dots, \sigma_m\}.$$

Proof of Lemma 1. Suppose σ_i is a valuation for s_i , $i = 1, \dots, m$. Let $s = \cup \{s_i \mid X \in s_i\}$, and let $r = \cup \{s_i \mid X \notin s_i\}$. Let $\rho = \otimes \{\sigma_i \mid X \notin s_i\}$, and $\sigma = \otimes \{\sigma_i \mid X \in s_i\}$. Note that $X \in s$, and $X \notin r$. Then

$$\begin{aligned} (\sigma_1 \otimes \dots \otimes \sigma_m) \downarrow^{(\mathcal{X} - \{X\})} &= (\rho \otimes \sigma) \downarrow^{((r \cup s) - \{X\})} \\ &= \rho \otimes (\sigma \downarrow^{(s - \{X\})}) && \text{(using axiom A3)} \\ &= (\otimes \{\sigma_i \mid X \notin s_i\}) \otimes (\sigma \downarrow^{(s - \{X\})}) \\ &= \otimes \text{Fus}_{\mathcal{X}} \{\sigma_1, \dots, \sigma_m\}. \quad \blacksquare \end{aligned}$$

Proof of Theorem 1. By axiom A2, $(\sigma_1 \otimes \dots \otimes \sigma_m) \downarrow^{\{X\}}$ is obtained by sequentially marginalizing all variables but X from the joint valuation. A proof of this theorem is obtained by repeatedly applying the result of Lemma 1. At each step, we delete a variable and fuse the set of all valuations with respect to this variable. Using Lemma 1, after fusion with respect to X_1 , the combination of all valuations in the resulting VBS is equal to $(\sigma_1 \otimes \dots \otimes \sigma_m) \downarrow^{(\mathcal{X} - \{X_1\})}$. Again, using Lemma 1, after fusion with respect to X_2 , the combination of all valuations in the resulting VBS is equal to $(\sigma_1 \otimes \dots \otimes \sigma_m) \downarrow^{(\mathcal{X} - \{X_1, X_2\})}$. And so on. When all variables but X have been deleted, we have the result. \blacksquare

Proof of Proposition 1. To show B5, suppose s is possible, and suppose $r \supseteq s$. Assume r is not possible. Then from B3, s is not possible contradicting our earlier supposition. Therefore r must be possible.

To show B6, suppose s is not possible. If $\sim s$ is also not possible, then from B4, $s \cup \sim s = \mathcal{W}_X$ is not possible contradicting B2. Therefore $\sim s$ must be possible.

To show that B7 is true, consider a possibilistic state in which the only proposition that is not possible is \emptyset (and all other propositions are possible). Clearly, this possibilistic state satisfies axioms B1 to B4. ■

Proof of Theorem 2. (Sufficiency). Assume that the possibilistic state is consistent. Consider the proposition $\cup \acute{u}$. It follows from B4 that $\cup \acute{u}$ is not possible, and follows from B2 that $\cup \acute{u}$ is a proper subset of \mathcal{W}_X . Define $c = \cup \acute{u}$. We need to show that $\acute{u} = \{s \in 2^{\mathcal{W}_X} \mid s \subseteq c\}$. Suppose $r \in \acute{u}$. Then $r \subseteq \cup \acute{u} = c$, i.e., $r \in \{s \in 2^{\mathcal{W}_X} \mid s \subseteq c\}$. Now suppose $r \in \{s \in 2^{\mathcal{W}_X} \mid s \subseteq c\}$, i.e., $r \subseteq c$. Then by axiom B3, $r \in \acute{u}$. Thus we have shown that $\acute{u} = \{s \in 2^{\mathcal{W}_X} \mid s \subseteq c\}$.

(Necessity). Suppose $\acute{u} = \{s \in 2^{\mathcal{W}_X} \mid s \subseteq c\}$ for some proper subset c of \mathcal{W}_X . We will show that this possibilistic state satisfies axioms B1 to B4. First note that B1 is satisfied by definition. B2 is satisfied since c is a proper subset of \mathcal{W}_X . To show B3, suppose that s is not possible and suppose $r \subseteq s$. Since s is not possible, by hypothesis, $s \subseteq c$. Therefore $r \subseteq c$. Therefore r is not possible. To show B4, suppose that r and s are not possible, i.e., r and s belong to \acute{u} . By hypothesis, $r \subseteq c$, and $s \subseteq c$. Hence, $r \cup s \subseteq c$. Therefore $r \cup s \in \acute{u}$, i.e., $r \cup s$ is not possible. ■

Proof of Proposition 2. P4 follows from condition P1 in Definition 2, and P5 follows from condition P2 in Definition 2. ■

Proof of Proposition 3. All four properties follow trivially from the definition of combination. ■

Proof of Theorem 3. Note that $p_1 \cup p_2 = (p_1 - p_2) \cup (p_1 \cap p_2) \cup (p_2 - p_1 - \{X\}) \cup \{X\}$, $p_1 = (p_1 - p_2) \cup (p_1 \cap p_2)$, and $p_2 = (p_1 \cap p_2) \cup (p_2 - p_1 - \{X\}) \cup \{X\}$. Suppose $\mathbf{w} \in \mathcal{W}_{p_1 - p_2}$, $\mathbf{u} \in \mathcal{W}_{p_1 \cap p_2}$, and $\mathbf{v} \in \mathcal{W}_{p_2 - p_1 - \{X\}}$, $\mathbf{x} \in \mathcal{W}_X$. Then $(\mathbf{w}, \mathbf{u}) \in \mathcal{W}_{p_1}$, and $(\mathbf{u}, \mathbf{v}, \mathbf{x}) \in \mathcal{W}_{p_2}$. First, note that the normalization factor in the combination on 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 , i.e., $K_1 = K_2$, as shown below.

$$\begin{aligned} K_1 &= \text{MAX} \left\{ \pi_1(\{(\mathbf{w}, \mathbf{u})\}) \pi_2(\{(\mathbf{u}, \mathbf{v}, \mathbf{x})\}) \mid (\mathbf{w}, \mathbf{u}, \mathbf{v}, \mathbf{x}) \in \mathcal{W}_{p_1 \cup p_2} \right\} \\ &= \text{MAX} \left\{ \pi_1(\{(\mathbf{w}, \mathbf{u})\}) \text{MAX} \left\{ \pi_2(\{(\mathbf{u}, \mathbf{v}, \mathbf{x})\}) \mid \mathbf{x} \in \mathcal{W}_X \right\} \mid (\mathbf{w}, \mathbf{u}, \mathbf{v}) \in \mathcal{W}_{((p_1 \cup p_2) - \{X\})} \right\} \\ &= \text{MAX} \left\{ \pi_1(\{(\mathbf{w}, \mathbf{u})\}) (\pi_2^{\downarrow(p_2 - \{X\})})(\{(\mathbf{u}, \mathbf{v})\}) \mid (\mathbf{w}, \mathbf{u}, \mathbf{v}) \in \mathcal{W}_{((p_1 \cup p_2) - \{X\})} \right\} \\ &= K_2 = K, \text{ say.} \end{aligned}$$

Next,

$$\begin{aligned}
(\pi_1 \otimes \pi_2)^{\downarrow((p_1 \cup p_2) - \{X\})}(\{(w, u, v)\}) &= \\
&= \text{MAX}\{(\pi_1 \otimes \pi_2)(\{(w, u, v, x)\}) \mid x \in \mathcal{W}_X\} \\
&= \text{MAX}\{K^{-1} \pi_1(\{(w, u)\}) \pi_2(\{(u, v, x)\}) \mid x \in \mathcal{W}_X\} \\
&= K^{-1} \pi_1(\{(w, u)\}) \text{MAX}\{\pi_2(\{(u, v, x)\}) \mid x \in \mathcal{W}_X\} \\
&= K^{-1} \pi_1(\{(w, u)\}) (\pi_2^{\downarrow(p_2 - \{X\})}(\{(u, v)\})) \\
&= (\pi_1 \otimes (\pi_2^{\downarrow(p_2 - \{X\})}))(\{(w, u, v)\}).
\end{aligned}$$

■

ACKNOWLEDGEMENTS

This work was supported in part by the National Science Foundation under grant IRI-8902444. I am grateful to Didier Dubois, Serafin Moral, and Leen-Kiat Soh for comments.

REFERENCES

- Arnborg, S., D. G. Corneil, and A. Proskurowski (1987), “Complexity of finding embeddings in a k-tree,” *SIAM Journal of Algebraic and Discrete Methods*, **8**, 277–284.
- Dempster, A. P. (1967), “Upper and lower probabilities induced by a multivalued mapping,” *Annals of Mathematical Statistics*, **38**, 325–339.
- Dubois, D., J. Lang, and H. Prade (1989), “Automated reasoning using possibilistic logic: Semantics, belief revision and variable certainty weights,” *Proceedings of the Fifth Workshop on Uncertainty in Artificial Intelligence*, 81–87, Windsor, Ontario.
- Dubois, D. and H. Prade (1988), *Possibility Theory: An Approach to Computerized Processing of Uncertainty*, Plenum Press, New York, NY.
- Dubois, D. and H. Prade (1991), “Inference in possibilistic hypergraphs,” in Bouchon-Meunier, B., R. R. Yager and L. A. Zadeh (eds.), *Uncertainty in Knowledge Bases*, Lecture Notes in Computer Science No. 521, 250–259, Springer-Verlag, Berlin.
- Chatalic, P., D. Dubois, and H. Prade (1987), “A system for handling relational dependencies in approximate reasoning,” *Proceedings of the Third International Expert Systems Conference*, 495–502, London, UK.
- Kjærulff, U. (1990), “Triangulation of graphs—Algorithms giving small total state space,” Technical report R90-09, Institute for Electronic Systems, University of Aalborg, Denmark.
- Kong, A. (1986), “Multivariate belief functions and graphical models,” Ph.D. dissertation, Department of Statistics, Harvard University, Cambridge, MA.
- Mellouli, K. (1987), “On the propagation of beliefs in networks using the Dempster-Shafer theory of evidence,” Ph.D. thesis, School of Business, University of Kansas, Lawrence, KS.
- Ruspini, E. H. (1991), “On the semantics of fuzzy logic,” *International Journal of Approximate Reasoning*, **5**(1), 45–88.
- Saffiotti, A. and E. Umkehrer (1991), “Pulcinella: A general tool for propagating uncertainty in valuation networks,” in D’Ambrosio, B., P. Smets, and P. P. Bonissone (eds.), *Uncertainty in Artificial Intelligence: Proceedings of the Seventh Conference*, 323–331, Morgan Kaufmann, San Mateo, CA.
- Shafer, G. (1976), *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ.

- Shenoy, P. P. (1989), "A valuation-based language for expert systems," *International Journal for Approximate Reasoning*, **3**(5), 383–411, 1989.
- Shenoy, P. P. (1990a), "Valuation-based systems for propositional logic," in Ras, Z. W., M. Zemankova, and M. L. Emrich (eds.), *Methodologies for Intelligent Systems*, **5**, 305–312, North-Holland, Amsterdam.
- Shenoy, P. P. (1990b), "Consistency in valuation-based systems," Working Paper No. 216, School of Business, Lawrence, KS.
- Shenoy, P. P. (1990c), "Valuation-based systems for Bayesian decision analysis," Working Paper No. 220, School of Business, University of Kansas, Lawrence, KS. To appear in 1992 in *Operations Research*, **40**(3).
- Shenoy, P. P. (1990d), "A new method for representing and solving Bayesian decision problems," Working Paper No. 223, School of Business, University of Kansas, Lawrence, KS. To appear in 1991 in Hand, D. J. (ed.), *Artificial Intelligence and Statistics III*, Chapman & Hall, London, England.
- Shenoy, P. P. (1991a), "On Spohn's rule for revision of beliefs," *International Journal of Approximate Reasoning*, **5**(2), 149–181.
- Shenoy, P. P. (1991b), "Valuation-based systems for discrete optimization," in Bonissone, P. P., M. Henrion, L. N. Kanal, and J. Lemmer (eds.), *Uncertainty in Artificial Intelligence*, **6**, 385–400, North-Holland, Amsterdam.
- Shenoy, P. P. (1991c), "Valuation-based systems: A framework for managing uncertainty in expert systems," in Zadeh, L. A. and J. Kacprzyk (eds.), *Fuzzy Logic for the Management of Uncertainty*, John Wiley & Sons, New York, NY.
- Shenoy, P. P. (1991d), "Using Dempster-Shafer's belief-function theory in expert systems," Working Paper No. 234, School of Business, University of Kansas, Lawrence, KS. To appear in *Proceedings of the SPIE Conference on Applications of Artificial Intelligence X*, Orlando, FL, April 1992.
- Shenoy, P. P. and G. Shafer (1988), "Constraint propagation," Working Paper No. 208, School of Business, University of Kansas, Lawrence, KS.
- Shenoy, P. P. and G. Shafer (1990), "Axioms for probability and belief-function propagation," in Shachter, R. D., T. S. Levitt, J. F. Lemmer and L. N. Kanal (eds.), *Uncertainty in Artificial Intelligence*, **4**, 169–198, North-Holland, Amsterdam. Reprinted in Shafer, G. and J. Pearl, eds. (1990), *Readings in Uncertain Reasoning*, 575–610, Morgan Kaufmann, San Mateo, CA.
- Spohn, W. (1988), "Ordinal conditional functions: A dynamic theory of epistemic states," in Harper, W. L. and B. Skyrms (eds.), *Causation in Decision, Belief Change, and Statistics*, **2**, 105–134, D. Reidel Publishing Company, Dordrecht.
- Spohn, W. (1990), "A general non-probabilistic theory of inductive reasoning," in Shachter, R. D., T. S. Levitt, J. F. Lemmer and L. N. Kanal (eds.), *Uncertainty in Artificial Intelligence*, **4**, 149–158, North-Holland, Amsterdam.
- Zadeh, L. A. (1965), "Fuzzy sets," *Information and Control*, **8**, 338–353.
- Zadeh, L. A. (1978), "Fuzzy sets as a basis for a theory of possibility," *Fuzzy Sets and Systems*, **1**, 3–28.
- Zadeh, L. A. (1979), "A theory of approximate reasoning," in Ayes, J. E., D. Mitchie, and L. I. Mikulich (eds.), *Machine Intelligence*, **9**, 149–194, Ellis Horwood, Chichester, U.K.
- Zhang, L. (1988), "Studies on finding hypertree covers of hypergraphs," Working Paper No. 198, School of Business, University of Kansas, Lawrence, KS.

ABOUT THE AUTHOR

Prakash P. Shenoy is the Director of the Doctoral Program and Professor in the School of Business at the University of Kansas at Lawrence where he has been since 1978. Before that, he was with the Mathematics Research Center at the University of Wisconsin at Madison for one year.

He received a B.Tech. in Mechanical Engineering from the Indian Institute of Technology at Bombay, India in 1973, and a M.S. and a Ph.D. in Operations Research from Cornell University in 1975 and 1977, respectively. His doctoral dissertation titled “On Coalition Formation and Game theory” was written with Professor William F. Lucas as his dissertation advisor.

In the Summer of 1984, he attended a six-week Faculty Development Institute on Information Systems at the University of Minnesota organized by the American Association of Collegiate Schools of Business. During 1985–86, he was an Intra-University Visiting Professor in the Department of Computer Science at the University of Kansas. Subsequently, he served as a faculty intern in the MIS Division of Hallmark Cards, Inc., in Kansas City, Missouri, for six months from June to December, 1986.

His research interests are in the areas of Artificial Intelligence and Operations Research. He has published many articles on management of uncertainty in expert systems, decision analysis, optimization, and the mathematical theory of games. His articles have appeared in journals such as the *International Journal of Approximate Reasoning*, *IEEE Expert*, *Operations Research*, *Management Science*, and the *International Journal of Game Theory*. He has received several research grants from the *Database and Expert Systems* program of the National Science Foundation, the *Research Opportunities in Auditing* program of the Peat Marwick Main Foundation, the *Higher Education Academic Development Donations* program of Apple Computer, Inc., and the *General Research Fund* program of the University of Kansas.

He served as a guest-editor for a special 1990 issue of the *International Journal of Approximate Reasoning* on the subject of Dempster-Shafer’s Theory of Belief Functions in Artificial Intelligence. He serves on the editorial board of *Uncertainty in Knowledge Bases: An International Journal*, and as an ad-hoc referee for over 25 journals and conferences in Artificial Intelligence and Operations Research. He also serves on the program committees of several conferences in artificial intelligence.

His teaching interests are in the areas of optimization, game theory, decision analysis, information systems, and uncertain reasoning. He has taught graduate-level courses on linear programming, non-linear programming, game theory, management information systems, decision support systems, and management of uncertainty in expert systems. He has served on doctoral dissertation committees of eighteen Ph.D. students in Management Science, Marketing, Accounting, Economics, Computer Science, and Geography, two as chairman. In Spring 1991, he was awarded the Mentor Award by the doctoral students in the School of Business at the University of Kansas.

SELECTED WORKING PAPERS

Unpublished working papers are available from the respective authors at:

School of Business
University of Kansas
Summerfield Hall
Lawrence, KS 66045-2003 USA

- No. 184. "Propagating Belief Functions with Local Computations," Prakash P. Shenoy and Glenn Shafer, February 1986. Appeared in *IEEE Expert*, **1**(1986), 43-52.
- No. 190. "Propagating Belief Functions in Qualitative Markov Trees," Glenn Shafer, Prakash P. Shenoy, and Khaled Mellouli, June 1987. Appeared in *International Journal of Approximate Reasoning*, **1**(1987), 349-400.
- No. 196. "On the Propagation of Beliefs in Networks Using the Dempster-Shafer Theory of Evidence," Khaled Mellouli, April 1988. Available as Ph.D. dissertation from University Microfilms, Inc., Ann Arbor, MI.
- No. 197. "AUDITOR'S ASSISTANT: A Knowledge Engineering Tool for Audit Decisions," Glenn Shafer, Prakash P. Shenoy, and Rajendra Srivastava, April 1988. Appeared in *Auditing Symposium IX: Proceedings of 1988 Touche Ross/University of Kansas Symposium on Auditing Problems*, 61-84, School of Business, University of Kansas, Lawrence, KS.
- No. 198. "Studies on Finding Hypertree Covers of Hypergraphs," Lianwen Zhang, April 1988.
- No. 199. "An Axiomatic Framework for Bayesian and Belief-Function Propagation," Prakash P. Shenoy and Glenn Shafer, July 1988. Appeared in *Proceedings of the Fourth Workshop on Uncertainty in Artificial Intelligence*, Minneapolis, MN, 307-314.
- No. 200. "Probability Propagation," Glenn Shafer and Prakash P. Shenoy, August 1988. Appeared in *Annals of Mathematics and Artificial Intelligence*, **2**(1990), 327-352.
- No. 201. "Local Computation in Hypertrees," Glenn Shafer and Prakash P. Shenoy, August 1988.
- No. 203. "A Valuation-Based Language for Expert Systems," Prakash P. Shenoy, August 1988. Appeared in *International Journal of Approximate Reasoning*, **3**(1989), 383-411.
- No. 206. "An Evidential Reasoning System," Debra K. Zarley, November 1988.
- No. 208. "Constraint Propagation," Prakash P. Shenoy and Glenn Shafer, November 1988.
- No. 209. "Axioms for Probability and Belief-Function Propagation," Prakash P. Shenoy and Glenn Shafer, November 1988. Appeared in *Uncertainty in Artificial Intelligence*, **4**(1990), 169-198. Reprinted in *Readings in Uncertain Reasoning*, Glenn Shafer and Judea Pearl, eds. (1990), 575-610, Morgan Kaufmann, San Mateo, CA.
- No. 210. "The Unity of Probability," Glenn Shafer, February 1989. Appeared in von Furstenberg, G. M., ed., *Acting Under Uncertainty: Multidisciplinary Conceptions*, 95-126, Kluwer, 1990.
- No. 211. "MacEvidence: A Visual Evidential Language for Knowledge-Based Systems," Yen-Teh Hsia and Prakash P. Shenoy, March 1989. A condensed version of this paper appeared as: Hsia, Y-T. and P. P. Shenoy, "An evidential language for expert systems," in Ras, Z. W., ed., *Methodologies for Intelligent Systems*, **4**(1989), 9-16.
- No. 212. "Audit Risk: A Belief-Function Approach," Rajendra Srivastava, Prakash P. Shenoy and Glenn Shafer, April 1989.

- No. 213. "On Spohn's Rule for Revision of Beliefs," Prakash P. Shenoy, July 1989. Appeared in *International Journal of Approximate Reasoning*, **5**(1991), 149–181.
- No. 216. "Consistency in Valuation-Based Systems," Prakash P. Shenoy, February 1990, revised May 1991.
- No. 218. "Perspectives on the Theory and Applications of Belief Functions," Glenn Shafer, April 1990. Appeared in *International Journal of Approximate Reasoning*, **4**(1990), 323-362.
- No. 220. "Valuation-Based Systems for Bayesian Decision Analysis," Prakash P. Shenoy, April 1990, revised May 1991. To appear in *Operations Research*, **40**(1992).
- No. 221. "Valuation-Based Systems for Discrete Optimization," Prakash P. Shenoy, June 1990. Appeared in Bonissone, P. P., M. Henrion, L. N. Kanal, and J. F. Lemmer, eds., *Uncertainty in Artificial Intelligence*, **6**(1991), 385–400, North-Holland, Amsterdam.
- No. 223. "A New Method for Representing and Solving Bayesian Decision Problems," Prakash P. Shenoy, September 1990, revised February 1991. To appear in 1991 in Hand, D. J., ed., *Artificial Intelligence and Statistics III*, Chapman & Hall, London, England.
- No. 225. "Why Should Statisticians be Interested in Artificial Intelligence?" Glenn Shafer, November, 1990. Appeared in *Proceedings of the Fifth Annual Conference on Making Statistics More Effective in Schools of Business*, 1990, 16–58, Lawrence, KS.
- No. 226. "Valuation-Based Systems: A Framework for Managing Uncertainty in Expert Systems," Prakash P. Shenoy, March, 1991. Revised July 1991. To appear in *Fuzzy Logic for the Management of Uncertainty*, Zadeh, L. A. and J. Kacprzyk, eds. (1991), John Wiley and Sons, New York, NY.
- No. 227. "Valuation Networks, Decision Trees, and Influence Diagrams: A Comparison," Prakash P. Shenoy, June 1991. Revised October 1991.
- No. 228. "The Early Development of Mathematical Probability," Glenn Shafer, August 1991. To appear in Grattan-Guinness, I., ed., *Encyclopedia of the History and Philosophy of the Mathematical Sciences*, Routledge, London.
- No. 229. "What is Probability?," Glenn Shafer, August 1991. To appear in Hoaglin, D. C. and D. S. Moore, eds., *Perspectives on Contemporary Statistics*, Mathematical Association of America.
- No. 230. "Can the Various Meanings of Probability be Reconciled?," Glenn Shafer, August 1991. To appear in Keren, G. and C. Lewis, eds., *Methodological and Quantitative Issues in the Analysis of Psychological Data*, Lawrence Erlbaum, Hillsdale, NJ.
- No. 231. "Response to the Discussion of Belief Functions," Glenn Shafer, August 1991. To appear in *International Journal of Approximate Reasoning* in 1992.
- No. 232. "An Axiomatic Study of Computation in Hypertrees," Glenn Shafer, August 1991.
- No. 233. "Using Possibility Theory in Expert Systems," Prakash P. Shenoy, September 1991. To appear in 1992 in *Fuzzy Sets and Systems*.
- No. 234. "Using Dempster-Shafer's Belief-Function Theory in Expert Systems," Prakash P. Shenoy, September 1991. To appear in 1992 in the *Proceedings of the SPIE Conference on Applications of Artificial Intelligence X*, Orlando, FL.
- No. 235. "Potential Influence Diagrams," Pierre Ndilikilikesha, September 1991.
- No. 236. "Independence in Valuation-Based Systems," Prakash P. Shenoy, September 1991. Revised November 1991.