A GENERALIZATION OF THE DEMPSTER-SHAFER THEORY

J. W. Guan, D. A. Bell
Department of Information Systems
University of Ulster at Jordanstown
BT37 OQB, Northern Ireland, U.K.
E-mail: CBFM23@UK.AC.ULSTER.UJVAX

Abstract

The Dempster-Shafer theory gives a solid basis for reasoning applications characterized by uncertainty. A key feature of the theory is that propositions are represented as subsets of a set which represents a *hypothesis space*. This power set along with the set operations is a Boolean algebra. Can we generalize the theory to cover arbitrary Boolean algebras ? We show that the answer is yes. The theory then covers, for example, infinite sets.

The practical advantages of generalization are that increased flexibility of representation is allowed and that the performance of evidence accumulation can be enhanced.

In a previous paper we generalized the Dempster-Shafer orthogonal sum operation to support practical evidence pooling.

In the present paper we provide the theoretical underpinning of that procedure, by systematically considering familiar evidential functions in turn. For each we present a "weaker form" and we look at the relationships between these variations of the functions. The relationships are not so strong as for the conventional functions. However, when we specialize to the familiar case of subsets, we do indeed get the well-known relationships.

1 Introduction

Uncertainty is a feature of our experience and observation of the world. Finding suitable means of representation and manipulation of uncertainty of information and knowledge [Bell 1992] is a challenge which will have to be met if computerized decision-making based on imperfect input is to be contemplated. An understanding of the effect of uncertainty on evidence appraisal, and ultimately on the behavior and properties of agents is essential. This paper contributes to this understanding and to the practical handling of evidence. It addresses the extension, in both practical and theoretical terms, of a numerical system which enables computer applications to reflect some aspects of uncertainty.

The theory of evidence which originated with Dempster and Shafer underpins a method which has been shown to be a promising tool for making judgements when confronted with uncertainty in numerical evidence. It generalizes Bayesian theory which is itself a popular theory of uncertainty.

The generalization of evidence theory in turn is the subject of this paper. It involves moving away from the standard finite set based derivation of theoretical and computational results underpinning the Dempster-Shafer approach. Conventionally propositions are represented as subsets of a collection of all possible values of a target variable. This particular representation of the *hypothesis space,* is not the only way to represent propositions. Most obviously we can think of leaving the propositions as they are, avoiding their transformation into subsets. This is of immediate interest in reasoning applications, because propositions are familiar and can be used to represent arguments, hypotheses, etc.

If this were done we would still be dealing with a structure which has an important similarity to the previous space — both are Boolean algebras. This leads to the question: can we generalize evidence theory to general Boolean algebras ? If we can, this allows us to choose a representation — we can use subsets, propositions, and other means to represent hypotheses and their relationships, as appropriate, in the hypothesis space. It can also allow us to establish a theory which covers infinite hypothesis spaces and evidence spaces by this extension.

This representational and theoretical advantage of the generalization is our focus of attention in this paper. However we have argued elsewhere [Guan & Bell 1993a], and supported our arguments by defining operations, that many applications can achieve improved performance through using more appropriate representations. We demonstrated that by generalizing the *orthogonal sum* operation so that hypotheses could be represented directly as propositions, such a performance enhancement could accrue. Using this representation, all the subsets of the hypothesis space Θ, i.e., $2^{|\Theta|}$ subsets, need not be considered (as they would in standard Dempster-Shafer theory). By focusing on relevant propositions only, the time complexity may be reduced to well below the previous $O(2^{|\Theta|})$ time.

To these advantages of representational and manipulative flexibility and efficiency for applications, we can add the advantage of developing a theory which covers *infinite* hypothesis spaces and evidence spaces, and extending our understanding of the Dempster-Shafer technique.

In section 2 we define evidential functions which are based on Boolean algebras. In particular we introduce weak versions of Bayesian functions, belief functions, and other evidential functions. We establish relationships between these weak functions and the familiar corresponding functions from evidence theory, showing that the results for power sets do not carry over to Boolean algebras in the general case. In section 3 we discuss nested evidential functions. We show in section 4 that we can obtain familiar relations for the particular case of the power set. The well-known inversions between the most conspicuous evidential functions are derived. Then we complete the paper by summarizing the relationships between the weak evidential functions obtained.

2 Evidential functions

In evidence theory, evidence is described in terms of evidential functions. There are several functions commonly used in the theory — mass functions, belief functions, commonality functions, doubt functions, and plausibility functions. Normally they are defined over finite sets. Here we generalize evidential functions to Boolean algebras. The significance of this is that conventional evidential functions are defined over the power set of a frame of discernment, but Boolean algebras include other interesting spaces, such as the space of propositions.

Let $< \mathcal{X}, \cup, \cap, ', \Phi, \Psi >$ be a Boolean algebra, where $\mathcal{X}$ is the *space of discernment*; $\cup$ is the union operation; $\cap$ is the intersection operation; $'$ is the negation operation; Φ is the zero element, the least element in $\mathcal{X}$ under the partial ordering relation $\subseteq$; Ψ is the identity element, the greatest element in $\mathcal{X}$ under $\subseteq$.

Let $[0, 1]$ denote the unit interval of real numbers.

A function $bay : \mathcal{X} \to [0, 1]$ is called a *weak Bayesian (probabilistic) function* if y1) $bay(\Phi) = 0$, y2) $bay(\Psi) = 1$, y3) $bay(A \cup B) = bay(A) + bay(B)$ whenever $A \cap B = \Phi$.

THEOREM 1. Let bay be a weak Bayesian function. Then (1) $bay(A) + bay(A') = 1$. (2) $bay(A) < bay(B)$ when $A \subseteq B$; i.e., bay is non-decreasing.

A function $bay : \mathcal{X}' \to [0, 1]$ is called a *Bayesian function* if y1) $bay(\Phi) = 0$, y2) $bay(\Psi) = 1$, y3' $bay(A \cup B) = bay(A) + bay(B) - bay(A \cap B)$.

A Bayesian function is a weak Bayesian function since y1), y2), and y3' imply y1), y2), and y3).

THEOREM 2. Suppose bay is a function $bay : \mathcal{X} \to [0, 1]$ and it satisfies y1) $bay(\Phi) = 0$, y2) $bay(\Psi) = 1$.

Then the following assertions are all equivalent: y3. bay is a Bayesian function. y3' $bay(A \cup B) = bay(A) + bay(B) - bay(A \cap B)$. y3" For any collection $A_1, A_2, ..., A_n$ ($n \geq 1$) of subsets of Ψ, $bay(A_1 \cup A_2 \cup ... \cup A_n) = \sum_{I \subseteq \{1,2,...,n\}, I \neq \emptyset} (-1)^{|I|+1} bay(\cap_{i \in I} A_i)$.

A function $bel : \mathcal{X} \to [0, 1]$ is called a *weak belief function* if it satisfies b1) $bel(\Phi) = 0$, b2) $bel(\Psi) = 1$, b3) for any collection $A_1, A_2, ..., A_n$ ($n \geq 1$) of subsets of Ψ, $bel(A_1 \cup A_2 \cup ... \cup A_n) \geq \sum_{I \subseteq \{1,2,...,n\}, I \neq \emptyset} (-1)^{|I|+1} bel(\cap_{i \in I} A_i)$.

THEOREM 3. Let bel be a weak belief function. Then (1) We have $bel(A \cup B) \geq bel(A) + bel(B) - bel(A \cap B)$ for all $A, B \in \mathcal{X}$. (2) bel is non-decreasing; i.e., $bel(A) \leq bel(B)$ when $A \subseteq B$. (3) $bel(X) + bel(X') \leq 1$.

Given a weak belief function bel, the function $dou(A) = bel(A')$ is called a *doubt function* and the function $pls(A) = 1 - dou(A) = 1 - bel(A')$ is called a *plausibility function*. It is easy to see that $dou(\Phi) = 1, dou(\Psi) = 0; pls(\Phi) = 0, pls(\Psi) = 1$.

THEOREM 4. A doubt function dou is non-increasing; i.e., $A \subset B$ implies $dou(A) \geq dou(B)$.

THEOREM 5. (1) A plausibility function pls is non-decreasing; i.e., $A \subseteq B$ implies $pls(A) \leq pls(B)$. (2) $pls(X) + pls(X') \geq 1, bel(X) \leq pls(X)$ for all $X \in \mathcal{X}$.

A function $m : \mathcal{X} \to [0, 1]$ is called a *mass function* if it has non-zero value only at a *finite* number F of elements $A_1, A_2, ..., A_F$ in $\mathcal{X}$ (i.e., subsets $\subseteq \Psi$) and it satisfies m1) $m(\Phi) = 0$, m2) $\sum_{X \subseteq \Psi} m(X) = \sum_{i=1}^{F} m(A_i) = 1$.

Notice that $X \subseteq \Psi$ if and only if $X \in \mathcal{X}$. Also, notice that m has non-zero value only at a finite number of subsets $X \subseteq \Psi$, so $\sum_{X \subseteq \Psi}$ makes sense.

A subset A of a frame $\mathcal{X}$ is called a *focal element* of a mass function m over $\mathcal{X}$ if $m(A) > 0$. Thus, the number of focal elements of a mass function is finite. The union C of all the focal elements of a mass function is called its *core*: $C = \cup_{X \in \mathcal{X}, m(X) > 0} X$, where $\cup_{X \in \mathcal{X}}$ makes sense as the number of focal elements of a mass function is finite.

A function bel on $\mathcal{X}$ is called a *belief function* if it can be expressed in terms of a mass function m: $bel(A) = \sum_{A \subseteq X} m(X)$ for all $A \in \mathcal{X}$.

THEOREM 6. Let m be a mass function. Then the function bel defined by the following expression $bel(A) = \sum_{X \subseteq A} m(X)$ for all $A \subseteq \Psi$ is a weak belief function.

A function $com : \mathcal{X} \to [0, 1]$ is a *commonality function* if there is a mass function m such that $com(A) = \sum_{A \subseteq X} m(X)$ for all $A \subseteq \Psi$.

3 The nested evidential functions

A function ben on $\mathcal{X}$ such that $ben : \mathcal{X} \to [0, 1]$ is said to be a *weak nested belief function* if it satisfies n1) $ben(\Phi) = 0$; n2) $ben(\Psi) = 1$; n3) $ben(A \cap B) = \min\{ben(A), ben(B)\}$ for all $A, B \subseteq \Psi$.

THEOREM 7. Let ben be a weak nested belief function on $< \mathcal{X}, \cup, \cap, ', \Phi, \Psi >$; i.e., ben is a function such that $ben : \mathcal{X} \to [0, 1]$ satisfying n1) $ben(\Phi) = 0$; n2) $ben(\Psi) = 1$; n3) $ben(A \cap B) = \min\{ben(A), ben(B)\}$ for all $A, B \subseteq \Psi$. Then ben is a weak belief function.

THEOREM 8. Let $< \mathcal{X}, \cup, \cap, ', \Phi, \Psi >$ be a Boolean algebra. Let ben, pls, dou be functions $\mathcal{X} \to [0, 1]$, and let $pls(A) = 1 - ben(A'), ben(A) = 1 - pls(A'); dou(A) = ben(A'), ben(A) = dou(A')$ for all $A \subseteq \Psi$. Suppose ben satisfies n1) $ben(\Phi) = 0$; n2) $ben(\Psi) = 1$. Then the following assertions are equivalent.

n3. ben is a weakly nested belief function.

n3) $ben(A \cap B) = \min\{ben(A), ben(B)\}$ for all $A, B \subseteq \Psi$.

p3) $pls(A \cup B) = \max\{pls(A), pls(B)\}$ for all $A, B \subseteq \Psi$.

d3) $dou(A \cup B) = \min\{dou(A), dou(B)\}$ for all $A, B \subseteq \Psi$.

THEOREM 9. Let $< \mathcal{X}, \cup, \cap, ', \Phi, \Psi >$ be a Boolean algebra.

Let *ben* be a weak nested belief function on $\mathcal{X}$, and let *pls* be its plausibility function on $\mathcal{X}$: $pls(A) = 1 - ben(A'), ben(A) = 1 - pls(A')$ for all $A \subseteq \Psi$.

Then we have the following.

(1) $\min\{ben(A), ben(A')\} = 0$ for all $A \subseteq \Psi$.

(2) $\max\{pls(A), pls(A')\} = 1$ for all $A \subseteq \Psi$.

(3) $pls(A) < 1$ implies $ben(A) = 0$ for all $A \subseteq \Psi$; i.e.,

$$interval(A) = [ben(A), pls(A)] = [0, pls(A)]$$

when $pls(A) < 1$;

$$ignorance(A) = pls(A) - ben(A) = pls(A) - 0 = pls(A)$$

when $pls(A) < 1$ for all $A \subseteq \Psi$.

(4) $ben(A) > 0$ implies $pls(A) = 1$ for all $A \subseteq \Psi$; i.e.,

$$interval(A) = [ben(A), pls(A)] = [ben(A), 1]$$

when $ben(A) > 0$;

$$ignorance(A) = pls(A) - ben(A) = 1 - ben(A)$$

when $ben(A) > 0$ for all $A \subseteq \Psi$.

Define $bel(X|B) = \frac{bel(X \cup B') - bel(B')}{1 - bel(B')}$ for all $X \subseteq \Psi$ and all $B \subseteq \Psi, bel(B') < 1$. Also, define $pls(X|B) = 1 - bel(X'|B), bel(X|B) = 1 - pls(X'|B)$ for all $X \subseteq \Psi$ and all $B \subseteq \Psi, bel(B') < 1, pls(B) > 0$.

THEOREM 10. Let $< \mathcal{X}, \cup, \cap, ', \Phi, \Psi >$ be a Boolean algebra. Let bel, pls be functions $\mathcal{X} \to [0,1]$, and $pls(A) = 1 - bel(A'), bel(A) = 1 - pls(A')$ for all $A \subseteq \Psi$. Suppose *bel* satisfies n1) $bel(\Phi) = 0$; n2) $bel(\Psi) = 1$; i3) *bel* is non-decreasing; i.e., $A \subseteq B$ implies $bel(A) \leq bel(B)$ for all $A, B \subseteq \Psi$. Then, the following assertions are all equivalent.

(1) For function *bel*, $\min\{bel(A|B), bel(A'|B)\} = 0$ for all $A \subseteq \Psi$ and all $B \subseteq \Psi$ such that $bel(B') < 1$.

(2) For function *pls*, $\max\{pls(A|B), pls(A'|B)\} = 1$ for all $A \subseteq \Psi$ and all $B \subseteq \Psi$ such that $pls(B) > 0$.

(3) Again $\max\{pls(A \cap B), pls(A' \cap B)\} = pls(B)$ for all $A \subseteq \Psi$ and all $B \subseteq \Psi$ such that $pls(B) > 0$.

(4) And, $\max\{pls(A \cap B), pls(A' \cap B)\} = pls(B)$ for all $A, B \subseteq \Psi$.

(5) Also, $pls(D \cup E) = \max\{pls(D), pls(E)\}$ for all $D, E \subseteq \Psi$.

(6) *bel* is a weak nested belief function.

Suppose m is a mass function on the Boolean algebra $< \mathcal{X}, \cup, \cap, ', \Phi, \Psi >$. m is said to be *nested* if its focal elements $A_i; i = 1, 2, ..., F; F > 0$ can be arranged into an increasing chain of supersets: $\Phi \subset A_1 \subset A_2 \subset ... \subset A_F \subseteq \Psi$, where $m(A_1), m(A_2), ..., m(A_F) > 0; \sum_{i=1}^{F} m(A_i) = 1$.

We also say the belief function, plausibility function, and doubt function *bel, pls, dou* are *nested* when the mass function m is nested.

THEOREM 11. Let $< \mathcal{X}, \cup, \cap, ', \Phi, \Psi >$ be a Boolean algebra. Let m be a mass function on $\mathcal{X}$, and let *bel* be its belief function: $bel(A) = \sum_{X \subseteq A} m(X)$ for all $A \subseteq \Psi$. Suppose m is nested.

Then *bel* is a weak nested belief function. That is, (by definition) *bel* satisfies n1) $bel(\Phi) = 0$; n2) $bel(\Psi) = 1$; n3) $bel(A \cap B) = \min\{bel(A), bel(B)\}$ for all $A, B \subseteq \Psi$.

4 THE SPECIAL CASE WHERE SETS ARE FINITE

In this section we show that the generalized theory presented in earlier sections reduces to familiar results when we are dealing with finite sets and subsets.

Let $< \mathcal{X}, \cup, \cap, ', \Phi, \Psi >$ be a Boolean algebra as before. Earlier, we showed that

$$\{ \text{ Bayesian functions } \} \subseteq \{ \text{ Weak Bayesian functions } \};$$

$$\{ \text{ Bayesian functions } \}, \{ \text{ Weak nested belief functions } \},$$

$$\{ \text{ Belief functions } \} \subseteq \{ \text{ Weak belief functions } \};$$

$$\{ \text{ Nested belief functions } \}$$

$$\subseteq \{ \text{ Weak nested belief functions } \}.$$

In this section we will prove that for finite set 2^Θ these $\subseteq$ relations become equalities:

$$\{ \text{ Bayesian functions } \} = \{ \text{ Weak Bayesian functions } \};$$

$$\{ \text{ Belief functions } \} = \{ \text{ Weak belief functions } \},$$

$$\{ \text{ Nested belief functions } \}$$

$$= \{ \text{ Weak nested belief functions } \}.$$

4.1 BAYESIAN FUNCTIONS

This subsection shows that Bayesian functions and weak Bayesian functions are identical for finite sets.

THEOREM 12. Let $< \mathcal{X}, \cup, \cap, ', \Phi, \Psi > = < 2^\Theta, \cup, \cap, ', \emptyset, \Theta >$, where Θ is a finite set. Then all weak Bayesian functions are Bayesian functions.

PROOF. Suppose *bay* is a function $bay : 2^\Theta \to [0,1]$ and it satisfies y1) $bay(\emptyset) = 0$, y2) $bay(\Theta) = 1$. Consider the following assertions:

y3) *bay* is weak Bayesian; i.e., $bay(A \cup B) = bay(A) + bay(B)$ whenever $A \cap B = \emptyset$.

y3} $bay(A) = \sum_{x \in A} bay(\{x\})$ for all $A \subseteq \Theta$.

y3' *bay* is Bayesian; i.e., $bay(A \cup B) = bay(A) + bay(B) - bay(A \cap B)$.

We prove this theorem as follows: y3) implies y3}, y3} implies y3'.

(i) y3) implies y3}. If $A = \emptyset$, then y3} $bay(\emptyset) = \sum_{x \in \emptyset} bay(\{x\})$ is true from y1) $bay(\emptyset) = 0$ and from convention

$$\sum_{x \in \emptyset} bay(\{x\}) = 0.$$

If $A \neq \emptyset$, then we may write $A = \{a_1, a_2, ..., a_n\}$, where $a_1, a_2, ..., a_n$ are distinct and $n \geq 1$. Applying y3) repeatedly, we find that $bay(A) = bay(\{a_1, a_2, ..., a_n\}) = bay(\{a_1\}) + bay(\{a_2, ..., a_n\}) = bay(\{a_1\}) + bay(\{a_2\}) + bay(\{a_3, ..., a_n\}) = ... = bay(\{a_1\}) + bay(\{a_2\}) + ... + bay(\{a_n\}) = \sum_{x \in A} bay(\{x\})$ since $\{a_1\} \cap \{a_2, ..., a_n\} = \emptyset, \{a_2\} \cap \{a_3, ..., a_n\} = \emptyset, ..., \{a_{n-1}\} \cap \{a_n\} = \emptyset$. Thus, y3) implies y3}.

(ii) y3} implies y3'. From y3}, $bay(A \cup B) = \sum_{x \in A \cup B} bay(\{x\}) = \sum_{x \in A} bay(\{x\}) + \sum_{x \in B} bay(\{x\}) - \sum_{x \in A \cap B} bay(\{x\}) = bay(A) + bay(B) - bay(A \cap B)$. So y3} implies y3'. **QED**

4.2 BELIEF FUNCTIONS

In order to prove that in the case where $\mathcal{X}$ is the power set of a finite set Θ, all weak belief functions are belief functions, we need the following Mobius inversion.

MOBIUS INVERSION. Suppose Θ is a finite set, f and g are functions on 2^Θ. Then $f(A) = \sum_{X \subseteq A} g(X)$ for all $A \subseteq \Theta$ if and only if $g(A) = \sum_{X \subseteq A} (-1)^{|A-X|} f(X)$ for all $A \subseteq \Theta$.

As usual, if m is a mass function, then the function bel defined by (BM) $bel(A) = \sum_{X \subseteq A} m(X)$ for all $A \subseteq \Theta$ is said to be the *belief function* of m.

THEOREM 13. When Θ is a finite set, all weak belief functions are belief functions. That is, if bel is a weak belief function, then there exists a mass function m such that the function bel can be given by m as follows: (BM) $bel(A) = \sum_{X \subseteq A} m(X)$ for all $A \subseteq \Theta$.

PROOF. Suppose bel is a weak belief function; i.e., bel is a function $bel : 2^\Theta \rightarrow [0,1]$ satisfying the three conditions b1), b2), b3).

Define a function m on 2^Θ by (MB) $m(A) = \sum_{X \subseteq A} (-1)^{|A-X|} bel(X)$ for all $A \subseteq \Theta$.

Then we have (BM) $bel(A) = \sum_{X \subseteq A} m(X)$ for all $A \subseteq \Theta$ and bel can be expressed in terms of m by Mobius inversion.

In order to prove that m is a mass function, we need to establish m1), m2), and $m : 2^\Theta \rightarrow [0,1]$, i.e., $0 \leq m(A) \leq 1$ for all $A \subseteq \Theta$.

From (MB) and b1) we have m1) $m(\emptyset) = \sum_{X \subseteq \emptyset} (-1)^{|\emptyset - X|} bel(X) = (-1)^0 bel(\emptyset) = 0$, and from (BM) and b2) we have m2) $\sum_{X \subseteq \Theta} m(X) = bel(\Theta) = 1$.

Now we only need to show that $0 \leq m(A) \leq 1$ for all $A \subseteq \Theta$. Let us prove first that $m(A) \geq 0$ for all $A \neq \emptyset$, $A \subset \Theta$.

If $|A| > 1$, then let $A = \{a_1, a_2, ..., a_n\}$, where $n > 1$ and $a_1, a_2, ..., a_n$ are distinct. Denote $A_i = A - \{a_i\}$ for $i = 1, 2, ..., n$, so that $A_1, A_2, ..., A_n$ are precisely the subsets of A that omit exactly one respective element $a_1, a_2, ..., a_n$ of A and $A = A_1 \cup A_2 \cup ... \cup A_n$. Then every proper subset X of A can be uniquely expressed as an intersection of the A_i. That is, for every $X \subset A$, there is an $I \subseteq \{1, 2, ..., n\}, I \neq \emptyset$ such that $X = \cap_{i \in I} A_i$ and $A - X = \{a_i | i \in I\}$.

Conversely, for every $I \subseteq \{1, 2, ..., n\}, I \neq \emptyset$ there is an $X \subset A$ such that $X = \cap_{i \in I} A_i$ and $A - X = \{a_i | i \in I\}$.

In other words, there is a one-to-one correspondence between the proper subsets X of Θ and the non-empty subsets I of $\{1, 2, ..., n\}$ such that $X = \cap_{i \in I} A_i$ and $A - X = \{a_i | i \in I\}$.

Thus by b3) $m(A) = \sum_{X \subseteq A} (-1)^{|A-X|} bel(X) = \sum_{X = A} (-1)^{|A-X|} bel(X) + \sum_{X \subset A} (-1)^{|A-X|} bel(X) = (-1)^{|\emptyset|} bel(A) + \sum_{I \subseteq \{1,2,...,n\}, I \neq \emptyset} (-1)^{|\{a_i | i \in I\}|} bel(\cap_{i \in I} A_i) = bel(A) + \sum_{I \subseteq \{1,2,...,n\}, I \neq \emptyset} (-1)^{|I|} bel(\cap_{i \in I} A_i) = bel(A_1 \cup A_2 \cup ... \cup A_n) - \sum_{I \subseteq \{1,2,...,n\}, I \neq \emptyset} (-1)^{|I|+1} bel(\cap_{i \in I} A_i) \geq 0$.

And if $A = \{a\}$, then from (MB) $m(\{a\}) = \sum_{X \subseteq \{a\}} (-1)^{|\{a\} - X|} bel(X) = (-1)^{|\{a\}|} bel(\emptyset) + (-1)^{|\emptyset|} bel(\{a\}) = bel(\{a\}) \geq 0$ (because $bel(\emptyset) = 0$ by b1) and bel is a function $bel : 2^\Theta \rightarrow [0,1]$ therefore $bel(\{a\}) \geq 0$).

Summarizing, we have $m(A) \geq 0$ for all $A \neq \emptyset$, $A \subset \Theta$ and from m1), we have $m(A) \geq 0$ for all $A \subseteq \Theta$.

Also from m2), we have $\sum_{X \subseteq \Theta} m(X) = 1$. So $m(X) \leq 1$ for all subsets $X \subseteq \Theta$, and $0 \leq m(X) \leq 1$ for all subsets $X \subseteq \Theta$, which is what we set out to show. **QED**

4.3 NESTED BELIEF FUNCTIONS

Let $< \mathcal{X}, \cup, \cap, ', \Phi, \Psi > = < 2^\Theta, \cup, \cap, ', \emptyset, \Theta >$ as before. We can prove that weak nested belief functions are nested belief functions.

We have $pls(A) = \sum_{X \cap A \neq \emptyset} m(X)$ for all $A \subseteq \Theta$: $pls(A) = 1 - bel(\bar{A}) = 1 - \sum_{X \subseteq \bar{A}} m(X) = 1 - \sum_{X \cap A = \emptyset} m(X) = \sum_{X \subseteq \Theta} m(X) - \sum_{X \cap A = \emptyset} m(X) = \sum_{X \cap A \neq \emptyset} m(X)$.

Also, we have the following *pls-com* inversion (theorem 2.4.5, p.59, Vol.1, Guan & Bell, 1991): $pls(\emptyset) = 0, pls(A) = \sum_{X \subseteq A, X \neq \emptyset} (-1)^{|X|+1} com(X)$ for all non-empty $A \subseteq \Theta$, and $com(\emptyset) = 1, com(A) = \sum_{X \subseteq A, X \neq \emptyset} (-1)^{|X|+1} pls(X)$ for all non-empty $A \subseteq \Theta$. Now we can prove the following.

THEOREM 14. Let $< \mathcal{X}, \cup, \cap, ', \Phi, \Psi > = < 2^\Theta, \cup, \cap, ', \emptyset, \Theta >$, where Θ is a finite set. Let m be a mass function on 2^Θ and let bel, pls, com be its corresponding belief function, plausibility function, and commonality function, respectively. Then the following assertions are all equivalent.

(1) bel is nested.

(2) bel is weakly nested; i.e.,

$$bel(A \cap B) = \min\{bel(A), bel(B)\}$$

for all $A, B \subseteq \Theta$.

(3) $pls(A \cup B) = \max\{pls(A), pls(B)\}$ for all $A, B \subseteq \Theta$.

(4) $pls(A) = \max_{x \in A}\{pls(\{x\})\}$ for all non-empty $A \subseteq \Theta$.

(5) $com(A) = \min_{x \in A}\{com(\{x\})\}$ for all non-empty $A \subseteq \Theta$.

PROOF. We proceed as follows: (1) implies (2), (2) implies (3), (3) implies (4), (4) implies (5), (5) implies (1).

(i) We prove (1) implies (2). Since m is a nested mass function on 2^Θ, its F focal elements can be arranged into an increasing chain of supersets: $\emptyset \subset A_1 \subset A_2 \subset ... \subset A_F \subseteq \Theta, F > 0$ where $m(A_1), m(A_2), ..., m(A_F) > 0; \sum_{i=1}^F m(A_i) = 1$. We want to prove that $bel(A \cap B) = \min\{bel(A), bel(B)\}$ for all $A, B \subseteq \Theta$.

Denote $A_0 = \emptyset$. Given two subsets $A, B \subseteq \Theta$, let $l \geq 0$ be the maximum integer such that $A_l \subseteq A$: $A_0 \subset A_1 \subset A_2 \subset ... \subset A_l \subseteq A, A_{l+1} \not\subseteq A$; i.e., $l = \max\{i | 0 \leq i \leq F, A_i \subseteq A\}$; and let $n \geq 0$ be the maximum integer such that $A_n \subseteq B$: $A_0 \subset A_1 \subset A_2 \subset ... \subset A_n \subseteq B, A_{n+1} \not\subseteq B$; i.e., $n = \max\{i | 0 \leq i \leq F, A_i \subseteq A\}$. Then

1. $A_i \subseteq A$ for $i = 0, 1, 2, ..., F$ if and only if $0 \leq i \leq l$.
2. $A_i \subseteq B$ for $i = 0, 1, 2, ..., F$ if and only if $0 \leq i \leq n$.
3. $A_i \subseteq A \cap B$ for $i = 0, 1, 2, ..., F$ if and only if $0 \leq i \leq \min\{l, n\}$.

Therefore we find that $bel(A \cap B) = \sum_{X \subseteq A \cap B} m(X) = \sum_{A_i \subseteq A \cap B; i=1,2,...,F} m(A_i)$ since $A_i; i = 1, 2, ..., F$ are all the focal elements of m at which m has positive values, so $bel(A \cap B)$

$$= \sum_{A_i \subseteq A \cap B; i=0,1,2,...,F} m(A_i) = \sum_{A_i, 0 \le i \le \min\{l,n\}} m(A_i)$$

$$= \sum_{i=0,1,2,...,\min\{l,n\}} m(A_i)$$

since $m(A_0) = m(\emptyset) = 0$ and by 3 above,

$$= \min\{ \sum_{i=0,1,2,...,l} m(A_i), \sum_{i=0,1,2,...,n} m(A_i)\}$$

$$= \min\{ \sum_{A_i, 0 \le i \le l} m(A_i), \sum_{A_i, 0 \le i \le n} m(A_i)\}$$

$$= \min\{ \sum_{A_i \subseteq A; i=0,1,2,...,F} m(A_i), \sum_{A_i \subseteq B; i=0,1,2,...,F} m(A_i)\}$$

by 1 and 2 above,

$$= \min\{ \sum_{A_i \subseteq A; i=1,2,...,F} m(A_i), \sum_{A_i \subseteq B; i=1,2,...,F} m(A_i)\}$$

since $m(A_0) = m(\emptyset) = 0$,

$$= \min\{ \sum_{X \subseteq A} m(X), \sum_{X \subseteq B} m(X)\} = \min\{bel(A), bel(B)\}$$

again since $A_i; i = 1, 2, ..., F$ are all the focal elements of m at which m has positive values.

(ii) (2) implies (3). Suppose $bel(A \cap B) = \min\{bel(A), bel(B)\}$ for all $A, B \subseteq \Theta$. We want to prove that $pls(A \cup B) = \max\{pls(A), pls(B)\}$ for all $A, B \subseteq \Theta$.

Indeed, we find that $pls(A \cup B) = 1 - bel(\overline{A \cup B}) = 1 - bel(\bar{A} \cap \bar{B}) = 1 - \min\{bel(\bar{A}), bel(\bar{B})\} = \max\{1 - bel(\bar{A}), 1 - bel(\bar{B})\} = \max\{pls(A), pls(B)\}$ for all $A, B \subseteq \Theta$.

(iii) (3) implies (4). Suppose (3) holds: $pls(A \cup B) = \max\{pls(A), pls(B)\}$ for all $A, B \subseteq \Theta$. We want to prove the assertion $pls(X) = \max_{x \in X}\{pls(\{x\})\}$ for all non-empty $X \subseteq \Theta$.

Use mathematical induction on $|X|$. When $|X| = 1$: $X = \{x\}, x \in \Theta$ the assertion is true since $pls(\{x\}) = \max_{x \in \{x\}}\{pls(\{x\})\} = \max\{pls(\{x\})\} = pls(\{x\})$ for all $x \in \Theta$.

Suppose the assertion is true when $|X| = n$ and now $|X| = n+1$, $X = \{x_1, x_2, ..., x_n, x_{n+1}\}$. Then we find that $pls(X) = \max\{pls(\{x_1, x_2, ..., x_n\}), pls(\{x_{n+1}\})\}$ from (3), $= \max\{\max\{pls(\{x_1\}), pls(\{x_2\}), ..., pls(\{x_n\})\}, pls(\{x_{n+1}\})\}$ by induction hypothesis since $|\{x_1, x_2, ..., x_n\}| = n$,

$$= \max\{pls(\{x_1\}), pls(\{x_2\}), ..., pls(\{x_n\}), pls(\{x_{n+1}\})\}$$

$= \max_{x \in X}\{pls(\{x\})\}$. That is, the assertion is true when $|X| = n+1$ and the mathematical induction is completed. The assertion is proved.

(iv) (4) implies (5). Suppose assertion (4) holds: $pls(X) = \max_{x \in X}\{pls(\{x\})\}$ for all non-empty $X \subseteq \Theta$. We want to prove that $com(X) = \min_{x \in X}\{com(\{x\})\}$ for all non-empty $X \subseteq \Theta$.

Notice that the plausibility function pls and the commonality function com are equal at all singletons: $com(\{x\}) = pls(\{x\})$ for all $x \in \Theta$ since $com(\{x\}) = \sum_{\{x\} \subseteq X} m(X) = \sum_{\{x\} \cap X \neq \emptyset} m(X) = pls(\{x\})$.

Let $X = \{x_1, x_2, ..., x_n\}$ and let $pls(\{x_1\}) \le pls(\{x_2\}) \le ... \le pls(\{x_n\})$. Then

1. Non-empty subsets Y of X are in one-to-one correspondence to non-empty subsets I of $\{1, 2, ..., n\}$: $Y \subseteq X, Y \neq \emptyset, Y = \{x_i | i \in I\} \longleftrightarrow I \subseteq \{1, 2, ..., n\}, I \neq \emptyset$.
2. We have

$$\max_{i \in I}\{pls(\{x_i\})\} = \sum_{i = \max\{j | j \in I\}} pls(\{x_i\}).$$

3. We have $i = \max\{j | j \in I\}$ if and only if $I \subseteq \{1, 2, ..., i\}$ and $i \in I$.
4. We have $pls(\{x_1\}) = \min_{x \in X}\{pls(\{x\})\}$.

So we find that $com(X)$

$= \sum_{Y \subseteq X, Y \neq \emptyset} (-1)^{|Y|+1} pls(Y)$ by the pls-com inversion,

$$= \sum_{I \subseteq \{1,2,...,n\}, I \neq \emptyset} (-1)^{|I|+1} pls(\{x_i | i \in I\})$$

$$= \sum_{I \subseteq \{1,2,...,n\}, I \neq \emptyset} (-1)^{|I|+1} \max_{i \in I}\{pls(\{x_i\})\}$$

by assertion (4),

$$= \sum_{I \subseteq \{1,2,...,n\}, I \neq \emptyset} (-1)^{|I|+1} \sum_{i = \max\{j | j \in I\}} pls(\{x_i\})$$

$$= \sum_{i \in \{1,2,...,n\}} pls(\{x_i\}) \sum_{I \subseteq \{1,2,...,i\}, i \in I} (-1)^{|I|+1}$$

$$= \sum_{i \in \{1,2,...,n\}} pls(\{x_i\}) \sum_{I = J \cup \{i\}, J \subseteq \{1,2,...,i-1\}} (-1)^{|J \cup \{i\}|+1}$$

$$= \sum_{i \in \{1,2,...,n\}} pls(\{x_i\}) \sum_{J \subseteq \{1,2,...,i-1\}} (-1)^{|J|} = pls(\{x_1\})$$

by the subset formula (formula 1.2.1, p.11, Vol.1, Guan & Bell, 1991),

$$\sum_{J \subseteq \{1,2,...,i-1\}} (-1)^{|J|} = \begin{cases} 1 & \text{when } i = 1; i.e., \{1, 2, ..., i-1\} = \emptyset \\ 0 & \text{otherwise.} \end{cases}$$

Thereby we conclude that assertion (5) holds: $com(X) = pls(\{x_1\}) = \min_{x \in X}\{pls(\{x\})\} = \min_{x \in X}\{com(\{x\})\}$ for all non-empty $X \subseteq \Theta$.

(v) (5) implies (1). We want to prove that $com(X) = \min_{x \in X}\{com(\{x\})\}$ for all non-empty $X \subseteq \Theta$ implies that m is nested; i.e., its F focal elements can be arranged into an increasing chain of supersets: $\emptyset \subset A_1 \subset A_2 \subset ... \subset A_F, F > 0$ where $m(A_1), m(A_2), ..., m(A_F) > 0; \sum_{i=1}^{F} m(A_i) = 1$.

Suppose m is not nested; i.e., there exist two focal elements $A, B \subseteq \Theta$ such that $A \not\subseteq B, B \not\subseteq A; m(A), m(B) > 0$. So there exist two elements $a, b \in \Theta$ such that $a \in A, a \notin B; b \in B, b \notin A$. We want to prove that assertion (5) does not hold; i.e., there exists a non-empty $D \subseteq \Theta$ such that $com(D) \neq \min_{x \in D}\{com(\{x\})\}$.

Let $D = \{a, b\}$. It suffices to prove that $com(\{a, b\}) < com(\{a\}), com(\{b\})$.

Indeed, we find that $com(\{a\}) = \sum_{X \subseteq \Theta, X \supseteq \{a\}} m(X) \geq m(A) + \sum_{X \subseteq \Theta, X \supseteq \{a,b\}} m(X)$ since $\{X \subseteq \Theta | X \supseteq \{a\}\} \supseteq \{X \subseteq \Theta | X \supseteq \{a, b\}\}, A \supseteq \{a\}, A \not\supseteq \{a, b\}$ Thus, we have $com(\{a\}) \geq m(A) + \sum_{X \subseteq \Theta, X \supseteq \{a,b\}} m(X) > \sum_{X \subseteq \Theta, X \supseteq \{a,b\}} m(X) = com(\{a, b\})$ since the focal element A makes $m(A) > 0$.

Also, we find that $com(\{b\}) = \sum_{X \subseteq \Theta, X \supseteq \{b\}} m(X) \geq m(B) + \sum_{X \subseteq \Theta, X \supseteq \{a,b\}} m(X)$ since $\{X \subseteq \Theta | X \supseteq \{b\}\} \supseteq \{X \subseteq \Theta | X \supseteq \{a, b\}\}, B \supseteq \{b\}, B \not\supseteq \{a, b\}$. Thus, we have $com(\{b\}) \geq m(B) + \sum_{X \subseteq \Theta, X \supseteq \{a,b\}} m(X) > \sum_{X \subseteq \Theta, X \supseteq \{a,b\}} m(X) = com(\{a, b\})$ since the focal element B makes $m(B) > 0$. **QED**

5 SUMMARY

Most important spaces in artificial intelligence are Boolean algebras, for example, power sets and proposition sets. The Dempster-Shafer theory originally addressed only the power sets. This paper generalizes the theory to Boolean algebras.

We investigate all the most important kinds of belief functions on an algebra to enable us to choose the most suitable belief function to represent evidence, according to the particular situation presented. The generalization enables us to choose the most suitable algebra to represent knowledge and reason efficiently.

We introduce weak Bayesian (probabilistic) functions, Bayesian (probabilistic) functions, weak belief functions, and belief functions. We show that Bayesian functions are weak Bayesian functions; and Bayesian functions and belief functions are weak belief functions.

Mass functions, commonality functions, plausibility functions, and doubt functions are also introduced.

In the case where $\mathcal{X} = 2^{\Theta}$ and Θ is a finite set, we show that weak Bayesian functions are Bayesian functions and vice versa. Moreover, weak belief functions are then belief functions and vice versa, and weak nested belief functions are nested belief functions and vice versa.

References

[Barnett, 198l] Barnett, J. A. 1981, " Computational methods for a mathematical theory of evidence ", Proc. IJCAI-81, 1981, 868-875.

[Bell et a/., 1992] Bell, D. A.; Guan, J. W.; Lee, S. K. 1993, " Efficient generalized union and projection operations for pooling uncertain and imprecise information ", Informatics Technical Reports, Univ. of Ulster.

[Gordon and Shortliffc, 1985] Gordon, J. ; Shortliffc, E. H. 1985 ," A method for managing evidential reasoning in a hierarchical hypothesis space ", Artificial Intelligence, 26(1985) 323-357.

[Guan et al., 1985] Guan, J. W. ; Xu, Y. and others, 1985, " Model expert system MES ", Proc. of IJCAI-85, 397-399.

[Guan and Lesser, 1987a] Guan, J. W. ; Lesser, V. R. 1987a, " The computational formulae of evidence combination scheme in a hierarchical hypothesis space ", Proc. of the 2nd Int. Conf. on Computers and Applications, Peking, 1987.

[Guan and Lesser, 1987b] Guan, J. W. ; Lesser, V R. 1987b, " On the evidence combination scheme in a hierarchical hypothesis space ", Proceedings of TENCON 87, Seoul, IEEE Region 10 Conference 1987.

[Guan et al., 1989] Guan, J. W ; Pavlin, J.; Lesser, V. R. 1989, " Combining evidence in the extended Dempster-Shafer theory ",Proc. of the 2nd Irish Con/. on AI and Cognitive Science, Dublin, 1989. Smeaton and McDermott (Eds.) Springer-Verlag, 1990, 163-178.

[Guan et al., 1991] Guan, J. W. ; Bell, D. A. ; Lesser, V. R. 1991, " Evidential reasoning and rule strengths in expert systems ", In McTear, M. F. ; Creany, N. (eds) Artificial Intelligence and Cognitive Science ' 90, pp. 378-390, 1991.

[Guan et al, 1990] Guan, J. W. ; Bell, D. A. ; Pavlin, J. ; Lesser, V. R. 1990, " The Dempster-Shafcr theory and rule strengths in expert systems ", IEE Colloquium on " Reasoning under Uncertainty ", 1990.

[Guan and Bell, 199l] Guan, J. W. ; Bell, D. A. 1991, " Evidence theory and its applications ", Volume 1, Studies in Computer Science and Artificial Intelligence 7, Elsevier, North-Holland.

[Guan and Bell, 1992] Guan, J. W. ; Bell, D. A. 1992, " Evidence theory and its applications ", Volume 2, Studies in Computer Science and Artificial Intelligence 8, Elsevier, North-Holland.

[Guan and Bell, 1993a] Guan, J. W. ; Bell, D. A. 1993, " Generalizing the Dempster-Shafer rule of combination to Boolean algebras ", Proc. of IEEE/ACM Int. Conf. on Developing and Managing Intelligent Systems Projects, Washington, March 1993.

[Guan and Bell, 1993b] Guan, J. W. ; Bell, D. A. 1993, " On some kinds of belief functions in Expert control ", Proc. of IFSICC93, USA, March 1993.

[Prade, 1985] Prade, H. 1985, " A computational approach to approximate and plausible reasoning with applications to expert systems ", IEEE Trans. Vol. PAMI-1, 260-283.

[Shafer, 1976] Shafer, G. 1976, " A mathematical theory of evidence ", Princeton University Press, Princeton, New Jersey, 1976.

[Shortliffe, 1976] Shortliffc, E. H. 1976, " Computer-based medical consultations: MYCIN ", Elsevier Publishing Company, New York.

[Shortliffe and Buchanan, 1975] Shortliffe, E. H. ; Buchanan, B. G. 1975, " A model of inexact reasoning in medicine ", Mathematical Biosciences, 23, 1975, 351-379.

[Yen, 1989] Yen, J. 1989, " GERT1S: A Dempster-Shafer approach to diagnosing hierarchical hypotheses ", CACM 5 vol. 32, 1989, pp.573-585.