Belief Functions on Distributive Lattices[*]

Chunlai Zhou

Department of Computer Science and Technology
School of Information, Renmin University of China
Beijing, 100872, CHINA
chunlai.zhou@gmail.com

Abstract

The Dempster-Shafer theory of belief functions is an important approach to deal with uncertainty in AI. In the theory, belief functions are defined on *Boolean* algebras of events. In many applications of belief functions in real world problems, however, the objects that we manipulate is no more a Boolean algebra but a *distributive lattice*. In this paper, we extend the Dempster-Shafer theory to the setting of distributive lattices, which has a mathematical theory as attractive as in that of Boolean algebras. Moreover, we apply this more general theory to a simple epistemic logic the first-degree-entailment fragment of relevance logic **R**, provide a sound and complete axiomatization for reasoning about belief functions for this logic and show that the complexity of the satisfiability problem of a belief formula with respect to the class of the corresponding Dempster-Shafer structures is NP-complete.

1 Introduction

Dealing with uncertainty is a fundamental issue for Artificial Intelligence (Halpern 2005). Numerous approaches have been proposed, including Dempster-Shafer theory of belief functions (Shafer 1976). Ever since the pioneering works by Dempster and Shafer, belief functions have become a standard tool in Artificial Intelligence for knowledge representation and decision-making.

Dempster-Shafer belief functions on a finite frame of discernment S are defined on the power set of S, which is a *Boolean algebra*. They have an attractive mathematical theory and many intuitively appealing properties. Belief functions satisfy the three axioms which generalize the Kolmogorov axioms for probability functions. Interestingly enough, they can also be characterized in terms of *mass functions* m. Intuitively, for a subset event A, $m(A)$ measures the belief that an agent commits *exactly* to A, not the total belief that an agent commits to A. Shafer (Shafer 1976) showed that an agent's belief in A is the sum of the masses he has assigned to all the subsets of A. This characterization of belief functions through mass functions is simply

an example of the well-known Inclusion-Exclusion principle in Enumerative Combinatorics (Stanley 1997) and hence has a strong combinatorial flavor. In this theory, mass functions are recognized as Möbius transforms of belief functions. Moreover, Dempster-Shafer theory of belief functions is closely related to classical probability theory. On one hand, a belief function on S is a probability function (also called *Bayesian belief function*) if and only if its corresponding mass function assigns *positive* weights only to singletons. On the other hand, a belief function as a generalized probability function can be regarded as *the inner probability measure* induced by some probability function on some Boolean algebra (Fagin and Halpern 1991), which implies that it represents *partial belief* and hence is the *restriction* of some probability function in the language of (Shafer 1976). There is an immediate payoff to this view of Dempster-Shafer belief functions: a logic for reasoning about belief functions can be obtained from that for inner probability measures (Fagin, Halpern, and Megiddo 1990).

As shown by Grabisch (Grabisch 2009), the theory of belief functions can be transposed in *general lattice setting*. This generalized theory has been applied to many objects in real world problems that may not form a Boolean algebra. Let us give some examples: set-valued variables in *multi-label classification*(Denoeux, Younes, and Abdallah 2010) (Zhang and Zhou 2007), the set of partitions in *ensemble clustering* (Masson and Denoeux 2011) and *bi-capacities* in cooperative game theory (Grabisch and Labreuche 2005). Because of its generality, however, Grabisch's theory loses many intuitively appealing properties. For example, since a lattice does not necessarily admit a probability function (Grabisch 2009), belief functions in general lattice settings fail to maintain a close connection with classical probability theory and therefore lack many of the desirable properties associated with this theory as in Dempster-Shafer theory (Fagin and Halpern 1991).

An optimal balance between utility and elegance of a theory of belief functions is achieved for *distributive lattices*, which is the main contribution of this paper. Not only does our approach for distributive lattices yields a mathematical theory as appealing as Dempster-Shafer theory, but also its applications extend to many *non-classical* formalisms of structures in Artificial Intelligence (quantum theory (Ying 2010) is one of very few important exceptions).

In this paper, we extend Dempster-Shafer theory to the setting of distributive lattices. The main difficulty in the extension is how to characterize the class of belief functions *without* reference to mass assignments. Birkhoff's fundamental theorem for finite distributive lattices solves this problem. Through this characterization, many fundamental properties of belief functions in the Boolean case are also preserved in distributive lattices. We show that, for any lattice (not necessarily distributive), a capacity is totally monotone iff its Möbius transform is non-negative, which answers an open question raised in (Grabisch 2009). Moreover, we prove that belief functions on distributive lattices can be viewed as generalized probability functions. A belief function Bel on a finite distributive lattice L is a probability function if and only if all *focal* elements i.e. those elements with positive weights assigned by its mass function are *join-irreducible* in L and Bel is the inner measure induced by some probability function. As an application, we apply the theory to a simple *non-Boolean* epistemic logic the first-degree-entailment fragment of the relevance logic **R** (Anderson and Belnap 1975), which is used to deal with the famous *logical-omniscience problem* in the foundations of Knowledge Representation (Fagin, Halpern, and Vardi 1995) (Levesque 1984), and used for reasoning in the presence of inconsistency in knowledge base systems (Lin 1996). A sound and complete axiomatization is provided for reasoning about belief functions for first degree entailments, and finally the complexity of the satisfiability problem of a belief formula with respect to the class of the corresponding Dempster-Shafer structures is shown to be NP-complete.

2 Belief function on lattices

All posets and lattices occurring in this paper are supposed to be finite. The reader is referred to (Davey and Priestley 2002) for all unexplained lattice-theoretical notation and terminology in this paper. Let $(L, \leq)$ be a poset having a bottom element $\perp$ and a top one $\top$ and $\mathbb{R}$ be the real field. For any function f on $(L, \leq)$, the *Möbius transform* of f is the function $m : L \to \mathbb{R}$ defined as the solution of the equation

$$f(x) = \sum_{y \leq x} m(y).$$

m is also called the *mass function* or *mass assignment* of f. The expression of m is obtained through *the Möbius function* $\mu : L^2 \to \mathbb{R}$ by

$$m(x) = \sum_{y \leq x} \mu(y, x) f(y)$$

where μ is defined inductively by

$$\mu(x, y) = \begin{cases} 1 & \text{if } x = y, \\ -\sum_{x \leq t \leq y} \mu(x, t) & \text{if } x < y, \\ 0 & \text{if } x > y. \end{cases}$$

Note that μ solely depends on L. And *the co-Möbius transform* of f is defined dually:

$$q(x) = \sum_{y \geq x} m(y), x \in L.$$

Definition 2.1 Given a lattice $\langle L, \leq \rangle$, a function f on L is called a *capacity* if it satisfies the following three conditions:

1. $f(\perp) = 0$;
2. $f(\top) = 1$;
3. $x \leq y$ implies $f(x) \leq f(y)$.

A function $bel : L \to [0, 1]$ is called a *belief function* if $bel(\top) = 1, bel(\perp) = 0$ and its Möbius transform is nonnegative. Its co-Möbius transform $q : L \to [0, 1]$ is called the *commonality function* associated to bel.

◁

Note that any belief function is a monotonic function by non-negativity of m, and hence a capacity.

Definition 2.2 Given a lattice $\langle L, \leq \rangle$, a function f on L is called a *k-monotone* whenever for each $(x_1, \cdots, x_k) \in L^k$, we have

$(*) : f(\bigvee_{1 \leq i \leq k} x_i) \geq \sum_{J \subset K, J \neq \emptyset} (-1)^{|J|+1} f(\bigwedge_{j \in J} x_j)$

A capacity is *totally monotone* if it is k-monotone for every $k \geq 2$. A k-monotone function f is called a *k-valuation* if the above inequality degenerates into the following equality:

$(*) : f(\bigvee_{1 \leq i \leq k} x_i) = \sum_{J \subset K, J \neq \emptyset} (-1)^{|J|+1} f(\bigwedge_{j \in J} x_j)$

It is an ∞-*valuation* if it is a k-valuation for each integer k. f is called a *probability function* if it is both a capacity and an ∞-valuation.

◁

The following proposition (Barthelemy 2000) tells us that there is a close relation between belief functions and totally monotone capacities.

Proposition 2.3 *Let* $f : L \to [0, 1]$ *be a capacity and* m *be its Möbius transform. If* f *is a belief function, then it is totally monotone.*

Shafer (Shafer 1976) proved that the converse is also true for any belief function on *Boolean algebras*. We show that actually it holds *generally* for any lattice, which answers an open question raised in (Grabisch 2009).

Theorem 2.4 *Let* L *be a lattice and* $f : L \to [0, 1]$ *be a capacity on* L *and* m *be its Möbius transform. The following two statements are equivalent:*

- *m is nonnegative;*
- *f is totally monotone.*

The proof is based on a structurally inductive construction of the Möbius transform m from f: $m(a) = f(a) - \sum_{b < a} m(b)$.

3 Belief functions on distributive lattices

Given a poset P, $J(P)$ denotes the lattice of order ideals of P with the ordinary union and intersection (as subsets of P). So $J(P)$ is distributive.

Proposition 3.1 *(Birkhoff's fundamental theorem for finite distributive lattices) Let* L *be a finite distributive lattice. Then there is a unique (up to isomorphism) finite poset* P *for which* $L \cong J(P)$.

The following two propositions provide formulas for Möbius functions and Möbius transforms in distributive lattices.

Theorem 3.2 *The Möbius function of the distributive lattice* $L = J(P)$ *is: for any* $I, I' \in J(P)$,

$$\mu(I, I') = \begin{cases} (-1)^{|I' \setminus I|} & \text{if } [I, I'] \text{ is a Boolean algebra,} \\ 0 & \text{otherwise.} \end{cases}$$

where $[I, I']$ *denotes the interval* $\{K \in J(P) : I \subseteq K \subseteq I'\}$.

Theorem 3.3 *Let* $L = J(P)$ *be a distributive lattice for some poset* P. *Suppose* $Bel : L \to [0, 1]$ *is the belief function given by the mass assignment* $m : L \to [0, 1]$. *Then*

$$m(A) = \sum_{\substack{[B,A] \ is \ a \\ Boolean \ algebra}} (-1)^{|A \setminus B|} Bel(B)$$

for all $A \in J(P)$.

From the above two theorems, we prove the following proposition along the same line as Theorem 2.1 in (Shafer 1976). It is also the distributive version of Theorem 2.4.

Corollary 3.4 *Given a distributive lattice* L, *a capacity* $Bel : L \to [0, 1]$ *is a belief function iff it is totally monotonic.*

It is easy to see that the intersection of Boolean algebras is still a Boolean algebra. For any two lattices L_1 and L_2, $L_1 \preceq L_2$ denotes that L_1 can be embedded into L_2.

Definition 3.5 A Boolean algebra B_L is *generated* by the distributive lattice L if B_L is the smallest Boolean algebra into which L can be embedded in the sense that

$$B_L = \bigcap \{B : B \text{ is a Boolean algebra and } L \preceq B\}.$$

$\triangleleft$

If $L = J(P)$ is a distributive lattice for some poset P, then the Boolean algebra B_L generated by L is the powerset of P with the usual set operations (Priestley 1970). However, it is only true in the finite case although Definition 3.5 also applies to infinite distributive lattices.

Lemma 3.6 *Let* L *be a distributive lattice and* B_L *be the Boolean algebra generated by* L.

1. *If* μ *is a probability function on* L, *then* μ *has a unique extension of probability function on* B_L;
2. *If* μ *is a probability function on* B_L, *then the restriction of* μ *into* L *is also a probability function on* L.

Lemma 3.7 *Let* L *be a distributive lattice and* B_L *be the Boolean algebra generated by* L.

1. *Any belief function* Bel *on* L *can be conservatively extended to a belief function* μ *on* B_L *in the sense that, for any* $x \in L$, $Bel(x) = \mu(x)$.

2. *For any belief function* Bel *on* B_L, *the restriction* $Bel \restriction_L$ *of* Bel *into* L *is a belief function on* L.

Shafer showed that a belief function is a probability function if and only if all of its focal elements are singletons (Shafer 1976). We have a similar property for distributive lattices.

Definition 3.8 Let Bel be a belief function on a distributive lattice D with the sub-poset I_D of join-irreducibles in D. The *inner belief function* Bel° of Bel is defined as follows:

$$Bel^\circ(a) = \sum_{x \in I_D, x \leq a} m(a).$$

In other words, $Bel^\circ(a)$ is the sum of all mass assignments on join-irreducibles which are less than or equal to a. $\triangleleft$

Bel° is a probability function on D. In some sense, the converse of this proposition also holds.

Theorem 3.9 *A belief function* Bel *on a distributive lattice* D *is a probability function iff its mass assignment* m *is given by* $m(c) = Bel(c)$ *for any join-irreducibles* c *in* D *and* $m(a) = 0$ *for all* join-reducible *elements* $a \in D$.

We have provided the condition when a belief function is a probability function. In the remainder of this section, we explore one perspective from which a belief function is regarded as a *generalized* probability function.

The notion of (non)measurability in measure theory is a desirable feature in reasoning with probabilities (Fagin and Halpern 1991) and can be generalized to the setting of distributive lattices. Let L be a distributive lattice and Bel be a belief function on L. Define

$$L_{Bel} := \{a \in L : Bel(a \vee b) = \\ Bel(a) + Bel(b) - Bel(a \wedge b) \text{ for every } b \in L\}.$$

Every element $a \in L_{Bel}$ is called Bel-*measurable* or simply measurable when the context is clear. From the main theorem in (Smiley 1940), we prove

Theorem 3.10 L_{Bel} *is a sublattice of* L *and hence is distributive. Moreover,* $Bel \restriction_{L_{Bel}}$ *is a probability function on* L_{Bel}.

Let L' be a sublattice of L and hence is distributive. If μ is a probability measure on L', then, for each element $x \in L$, we define

- $\mu_*(x) = sup\{\mu(y) : y \in L', y \leq x\}$;
- $\mu^*(x) = inf\{\mu(y) : y \in L', y \geq x\}$

μ_* and μ^* are called *inner and outer probability functions induced by* μ on L respectively. Such defined μ_* is a belief function on L and is called *canonical* in the language of (Shafer 1979) in the sense that it gives each element $a \in L$ the minimal degree of belief that is compelled by μ.

Corollary 3.11 *For the above belief function* Bel *on* L, $Bel \geq (Bel \restriction_{L_{Bel}})_*$ *in the sense that* $Bel(a) \geq (Bel \restriction_{L_{Bel}})_*(a)$ *for all* $a \in L$.

Although a belief function on a distributive lattice L is not generally an inner probability function on L, we show that from a certain perspective it is an inner probability function on some *expanded* distributive lattice.

The language Φ_0 is defined inductively as follows:

$$\phi := \bot \mid \top \mid p \mid \phi \wedge \phi \mid \phi \vee \phi$$

where p is a propositional letter.

Definition 3.12 A *probability structure* is a tuple $M = \langle L, L', \mu, v \rangle$ where both L and L' are distributive lattices, L' is a sublattice of L, μ is a probability function on L', μ_* is the inner probability function induced by μ, and v associates with each p with an element in L. v can be easily extended to a homomorphism from Φ_0 to L.

$\triangleleft$

Definition 3.13 A *DS structure* is a tuple $D = \langle L, Bel, v \rangle$ where L is a distributive lattice, Bel is a belief function on L and v maps each propositional letter p to an element in L. Similarly, v can be easily extended to a homomorphism from Φ_0 to L.

$\triangleleft$

We call a probability structure $M = \langle L, L', \mu, v_M \rangle$ and a DS structure $D = \langle L_D, Bel, v_D \rangle$ *equivalent* if

for any formula $\phi \in \Phi_0$, $\mu_*(v_M(\phi)) = Bel(v_D(\phi))$.

Theorem 3.14 *For any probability structure $M = \langle L, L', \mu, v_M \rangle$, there is an equivalent DS-structure. Moreover, for every DS structure $D = \langle L_D, Bel, v_D \rangle$, there is an equivalent probability structure $M = \langle L, L', \mu, v_M \rangle$.*

By summarizing the results in the above theorem, we conclude that belief functions and inner probability functions are equivalent on distributive lattices if we view them both as functions on *formulas* rather on sets. The following theorem is simply a corollary of Theorem 3.14 which says that each belief function on distributive lattices is the *restriction* of some probability function in the language of (Shafer 1976).

Corollary 3.15 *Given a belief function Bel defined on a distributive lattice L_D, there are a distributive lattice L, a probability measure μ on a sublattice L' of L and a surjective homomorphism $f : L \to L_D$ such that, for each $x \in L_D$, we have $Bel(x) = \mu_*(f^{-1}(x))$.*

Just as in (Fagin and Halpern 1991), there is an immediate payoff to this view of belief functions as generalized probability functions: a logic for reasoning about belief functions for the first degree entailments in the last section is obtained from that for inner probability functions (Fagin, Halpern, and Megiddo 1990).

4 Belief functions on de Morgan lattices

An important issue for belief functions, each of which can be viewed as representing a distinct body of evidence, is how to combine them to obtain a new belief function that reflects the combined evidence. A way of doing so is provided by Dempster's rule of combination (Shafer 1976). The definition of rule of combination is usually given in terms of mass functions.

In this section, we first address the combination of belief functions on distributive lattices and then apply this result to define belief functions on de Morgan lattices through a duality theorem for finite de Morgan lattices which is similar to Birkhoff's theorem for finite distributive lattices. A belief function is defined on a de Morgan lattice through the combination of two district beliefs, one of which accounts for "true" facts of a knowledge base and the other for "false" facts. Note that de Morgan lattices provide the algebraic semantics for the relevance logic **R** (Dunn 1986).

Definition 4.1 Let Bel_1, Bel_2 be two belief functions on a distributive lattice L with the corresponding mass assignments m_1 and m_2, respectively. Let $m_1 \oplus m_2$ be a function on L defined as $(m_1 \oplus m_2)(a) = c \sum_{b \wedge c = a} m_1(b)m_2(c)$ for any $a \in L$ where $c = (\sum_{b \wedge c \neq \bot} m_1(b)m_2(c))^{-1}$. It is easy to see that so defined $m_1 \oplus m_2$ is a belief function on L and c is actually the normalizing constant. If there is no pair $b, c \in L$ such that $b \wedge c \neq \bot$ and $m_1(b)m_2(c) > 0$, then we can not find such a constant c and hence $m_1 \oplus m_2$ is undefined. $Bel_1 \oplus Bel_2$ denotes the combined belief function which corresponds to $m_1 \oplus m_2$. $\triangleleft$

Next we apply the above combination rule to define belief functions on de Morgan lattices. We need the following duality theorem for finite de Morgan lattices which is adapted from (Dunn 1986), (Urquhart 1979) and (Priestley 1970):

Theorem 4.2 *Any finite de Morgan lattice D can be represented as the lattice $J(P_D)$ of order ideals in the sub-poset P_D of join-irreducibles with an order-reversing involution g. And there is a one-to-one correspondence between de Morgan lattices and posets with order-reversing involutions.*

Given a de Morgan lattice D, we know from the representation theorem about finite de Morgan lattice that it can be represented as a concrete lattice $J(P)$ of all order ideals in P for some poset P with an order-reversing involution g. Let $E(P)$ be the concrete lattice of order filters in P with the usual set operations. It is easy to see that $E(P)$ is distributive and g is a *dual* isomorphism between lattices $J(P)$ and $E(P)$. Based on the poset P, another concrete lattice $D(P)$ on $\{(I, F) : I \in J(P), F \in E(P))\}$ is defined as follows:

- $(I_1, F_1) \wedge (I_2, F_2) = (I_1 \cap I_2, F_1 \cap F_2)$;
- $(I_1, F_1) \vee (I_2, F_2) = (I_1 \cup I_2, F_1 \cup F_2)$;
- $\sim (I, F) = (F^c, I^c)$ where F^c and I^c are the complements of F and I with respect to P, respectively.

Such defined lattice $D(P)$ is a de Morgan lattice but not necessarily Boolean. $D_r(P) := \{(I, F) : I \in J(P), F = $

$g(I)\}$ is a sublattice of $D(P)$. Moreover, it is also a concrete representation of the distributive lattice D, i.e., D is isomorphic to $D_r(P)$ (Dunn 1986).

Let Bel_1, Bel_2 be two belief functions on distributive lattices $J(P)$ and $E(P)$, respectively. They can also be regarded as belief functions on the following two sub-lattices of $D(P)$:

- $D_1(P) = \{(I, F) : I \in J(P), F = P\}$;
- $D_2(P) = \{(I, F) : I = P, F \in E(P)\}$.

becasue these two distributive lattices $D_1(P)$ and $D_2(P)$ are isomorphic to $J(P)$ and $E(P)$, respectively. Elements in $D_1(P)$ ($D_2(P)$) are called *t-grounded* (*f-grounded*) in $D(P)$ in the language of (Ginsberg 1988). So we may also regard Bel_1 and Bel_2 as belief functions on these two sub-lattices of $D(P)$. Intuitively, Bel_1 is the support function for the true facts in a knowledge base system while Bel_2 is that for the false facts. Let m_1 and m_2 be their corresponding mass assignments. Bel_1^* is defined to be the extension of Bel_1 as follows: for any $a \in D(P)$,

$$m_1^*(a) = \begin{cases} m_1(a) & \text{if } a \in D_1(P) \\ 0 & \text{otherwise.} \end{cases}$$

Bel_2^* is defined similarly through the extension m_2^* of m_2 to $D(P)$. Now we define a combined belief function of Bel_1^* and Bel_2^*. For, any $(I, F) \in D(P)$,

$$(m_1^* \oplus m_2^*)(I, F) = \sum_{A \wedge B = (I,F)} m_1^*(A) m_2^*(B)$$
$$= m_1^*(I, P) m_2^*(P, F)$$
$$= m_1(I) m_2(F)$$

Since $\sum_{I \subseteq P, F \subseteq P}(m_1^* \oplus m_2^*)(I, F) = 1$, $m_1^* \oplus m_2^*$ is the mass assignment for the combined belief function $Bel_1^* \oplus Bel_2^*$, which is simply written as Bel^*.

In order to define the combined belief function on $D_r(P)$, which is isomorphic to D, through the two individual belief functions Bel_1 and Bel_2, we only need the following normalizing constant:

$$c_D := \left(\sum_{I \subseteq P} m_1(I) m_2(g(I))\right)^{-1}.$$

Note that the above argument can be reversed to show that any belief function on the de Morgan lattice D *is* the combination of two "independent" belief functions in which one accounts for "true" propositions and the other for "false" ones. The main purpose of this section is to provide semantics for the following section.

5 Reasoning about belief functions for first-degree entailments

In this section, we first present the syntax and semantics of first degree entailments. Here we choose the Routley semantics with the well-known Routley star dealing with negation (Routley and Routley 1972). Next we define belief structures for this semantics through belief functions on de Morgan lattices in Section 4 and show that the consequence relation with respect to the class of belief structures is the same as

that with respect to the Routley semantics. In addition, we extend the above language to include belief formulas which can express the linear relations of different beliefs and give a sound and complete axiomatization of validity in the class of belief structures.

The syntax of first degree entailments is very similar to that of propositional logic. We start with a fixed *finite* set $P := \{p_1, p_2, \cdots, p_n\}$ of propositional letters, which can be thought of as corresponding to basic "events". A formula ϕ is formed by the following syntax:

$$\phi := \bot \mid \top \mid p \mid \sim \phi \mid \phi_1 \wedge \phi_2 \mid \phi_1 \vee \phi_2$$

The following is the deductive system R_{fde} which is the well-known first-degree entailment fragment of the relevance logic **R** (Dunn 1986). Without further notice, $\vdash$ denotes $\vdash_{R_{fde}}$. The bi-entailment $\phi \dashv\vdash \gamma$ is short for $\phi \vdash \gamma$ and $\gamma \vdash \phi$.

Axioms:

$$\phi \vdash \phi$$
$$\phi \wedge \psi \vdash \phi, \qquad \phi \wedge \psi \vdash \psi$$
$$\phi \vdash \phi \vee \psi, \qquad \psi \vdash \phi \vee \psi$$
$$\phi \wedge (\psi \vee \gamma) \vdash (\phi \vee \psi) \wedge (\psi \vee \gamma) \qquad \phi \dashv\vdash \sim\sim \phi$$
$$\sim (\phi \wedge \psi) \dashv\vdash \sim \phi \vee \sim \psi$$
$$\sim (\phi \vee \psi) \dashv\vdash \sim \phi \wedge \sim \psi$$

Rules:

- From $\phi \vdash \psi$ and $\psi \vdash \gamma$, infer $\phi \vdash \gamma$ (Transitivity)
- From $\phi \vdash \psi$ and $\phi \vdash \gamma$, infer $\phi \vdash \psi \wedge \gamma$ ($\wedge$-introduction)
- From $\phi \vdash \gamma$ and $\psi \vdash \gamma$, infer $\phi \vee \psi \vdash \gamma$ ($\vee$-introduction)
- From $\phi \vdash \psi$, infer $\sim \psi \vdash \sim \phi$ (Contraposition)

Definition 5.1 A *Routley structure* is a tuple $S = \langle S, \leq, g, v \rangle$ where

- $\langle S, \leq, g \rangle$ is a poset with an order-reversing involution g;
- v is a valuation on the set of propositional letters: $v(s)(p) \in \{true, false\}$ for all p which satisfies the following *persistency* condition: for any $s_1, s_2 \in S$ and propositional letter p,

  if $s_1 \leq s_2$ and $v(s_1)(p) = true$, then $v(s_2)(p) = true$.

It is called a *Boolean structure* if $\leq$ is the identity relation and g is the identity function on S. $\lhd$

A satisfaction relation $\models$ between states and formulas is defined inductively as follows:

- $S, s \models p$ if $v(s)(p) = true$;
- $S, s \models \phi_1 \wedge \phi_2$ if $S, s \models \phi_1$ and $S, s \models \phi_2$;
- $S, s \models \phi_1 \vee \phi_2$ if $S, s \models \phi_1$ or $S, s \models \phi_2$;
- $S, s \models \sim \phi$ if $S, g(s) \not\models \phi$.

Definition 5.2 A formula ϕ *logically implies* a formula ψ (denoted as $\phi \models \psi$) if, for any Routley structure $S = \langle S, g, v \rangle$, $S, s \models \phi$ implies $S, s \models \psi$. $\lhd$

Actually the logical implication relation in the class of Routley structures coincides with the above consequence relation $\vdash$ (Dunn 1986) and the complexity of logical implication is the same as that in standard propositional logic (Fagin, Halpern, and Vardi 1995).

Now we add belief functions to Routley structures just as Shafer did to Boolean structures (Shafer 1976).

Definition 5.3 A *Dempster-Shafer structure* B (*DS-structure for short*) for the first degree entailments is a tuple $\langle S, \leq, g, v, Bel \rangle$ where

- S $= \langle S, \leq, g, v \rangle$ is a Routley structure;
- Bel_1 is a belief function on $J(S) = \{I \subseteq S : I$ is an ideal in $\langle S, \leq \rangle\}$ and Bel_2 a belief function on $E(S) = \{F \subseteq S : F$ is a filter in $\langle S, \leq \rangle\}$;
- $Bel = Bel_1 \oplus Bel_2$ which is the combination of the two belief functions Bel_1 and Bel_2 on the de Morgan lattice $D_r(S)(= \{(I, g(I)) : I \in J(S)\})$.

$\triangleleft$

For the Routley structure $\langle S, \leq, g, v \rangle$, $[[\phi]]_S$ denotes the set of states where ϕ is satisfied, i.e., $[[\phi]]_S = \{s \in S : S, s \models \phi\}$, which is a filter in S. So, for any formula ϕ, $[[\phi]]_B := (g([[\phi]]_S), ([[\phi]]_S)) \in D(S)$. The belief in the formula ϕ is defined to be $Bel([[\phi]]_B)$, which is equal to $(Bel_1 \oplus Bel_2)(g([[\phi]]_S), [[\phi]]_S)$.

Definition 5.4 For any two formulas ϕ and ψ in Φ, ϕ *probabilistically entails* ψ (denoted as $\phi \models_{DS} \psi$) if, for any DS-structure $B = \langle S, \leq, g, v, Bel \rangle$, $Bel([[\phi]]_B) \leq Bel([[\psi]]_B)$.

$\triangleleft$

The following theorem tells us that the deductive system $\vdash$ of first degree entailments provides a sound and complete system for both logical implication but also probabilistic entailment.

Theorem 5.5 *For any formulas ϕ and ψ in Φ, $\phi \vdash \psi$ if and only if $\phi \models_{DS} \psi$.*

For the above given set Φ of formulas, a *term* is an expression of the form $a_1 bel(\phi_1) + a_2 bel(\phi_2) + \cdots + a_k bel(\phi_k)$, where $a_1, a_2, \cdots, a_k$ are integers and $\phi_1, \phi_2, \cdots, \phi_k$ are formulas in Φ. *A basic belief formula* is one of the form $t \geq b$, where t is a term and b is an integer. A *belief formula* is a Boolean combination of basic belief formulas. We can always allow rational numbers in our formulas as abbreviations for the formula that would be obtained by clearing the dominator. And other derived relations $=, \leq, <$ and $>$ can be defined as usual.

Given a DS-structure $B = \langle S, \leq, g, v, Bel \rangle$ and a basic belief formula $f := a_1 bel(\phi_1) + a_2 bel_2(\phi_2) + \cdots + a_k bel(\phi_k) \geq b$, B *satisfies* f (denoted as $B \models f$) if $a_1 Bel([[\phi_1]]_B) + a_2 Bel([[\phi_2]]_B) + \cdots + a_k Bel([[\phi_k]]_B) \geq b$. We then extend the above $\models$ in the obvious way to all belief formulas. Let $\mathcal{B}$ be a class of Dempster-Shafer structures. A belief formula f' is *satisfiable with respect to $\mathcal{B}$* if it is satisfied in some $B \in \mathcal{B}$. It is *valid with respect to $\mathcal{B}$* if $B \models f$ for all $B \in \mathcal{B}$.

The axiomatization $\mathcal{B}_{fde}$ of reasoning about belief functions for first degree entailments consists of three parts: the first-degree entailments, reasoning about linear inequalities and reasoning about belief functions.

1. First-degree entailments

   - The complete system $\vdash$ of first degree entailments;

2. Reasoning about linear inequalities

   (a) $a_1 bel(\phi_1) + a_2 bel(\phi_2) + \cdots + a_k bel(\phi_k) \geq b$ iff $a_1 bel(\phi_1) + a_2 bel(\phi_2) + \cdots + a_k bel(\phi_k) + 0bel(\phi_{k+1}) \geq b$;

   (b) $a_1 bel(\phi_1) + a_2 bel(\phi_2) + \cdots + a_k bel(\phi_k) \geq b$ iff $a_{j_1} bel(\phi_{j_1}) + a_{j_2} bel(\phi_{j_2}) + \cdots + a_{j_k} bel(\phi_{j_k}) \geq b$ where $j_1, j_2, \cdots, j_k$ is a permutation of $1, 2, \cdots, k$;

   (c) $a_1 bel(\phi_1) + a_2 bel(\phi_2) + \cdots + a_k bel(\phi_k) \geq b$ iff $ca_1 bel(\phi_1) + ca_2 bel(\phi_2) + \cdots + ca_k bel(\phi_k) \geq cb$ where $c > 0$;

   (d) $(a_1 + a_1')bel(\phi_1) + (a_2 + a_2')bel(\phi_2) + \cdots + (a_k + a_k')bel(\phi_k) \geq b + b'$ if $a_1 bel(\phi_1) + a_2 bel(\phi_2) + \cdots + a_k bel(\phi_k) \geq b$ and $a_1' bel(\phi_1) + a_2' bel(\phi_2) + \cdots + a_k' bel(\phi_k) \geq b'$;

   (e) Either $t \geq b$ or $t \leq b$ where t is a term;

   (f) $t \geq b$ implies $t > b'$ where t is a term and $b' < b$.

3. Reasoning about belief functions

   (a) $bel(\phi) \geq 0$ for all formulas $\phi \in \Phi$;

   (b) $bel(\top) = 1$

   (c) $bel(\bot) = 0$;

   (d) $bel(\phi_1 \vee \phi_2 \vee \cdots \vee \phi_n) \geq \sum_{I \subseteq \{1, 2, \cdots, n\}} (-1)^{|I|+1} bel(\wedge_{i \in I} \phi_i)$;

   (e) $bel(\phi) \leq bel(\psi)$ if $\phi \vdash \psi$.

Item 3e (also Theorem 5.5) connects first degree entailments to reasoning about belief functions.

Theorem 5.6 *$\mathcal{B}_{fde}$ is a sound and complete axiomatization of belief formulas with respect to the class of DS-structures.*

Theorem 5.7 *The time complexity of deciding whether a belief formula is satisfiable with respect to the class of DS-structures is NP-complete.*

6 Conclusion

The main contribution of this paper is the extension of Dempster-Shafer theory of belief functions on Boolean algebras to the setting of distributive lattices and show that many intuitively appealing properties in the theory are transposed to this more general case. As an application, we apply this general theory to a non-classical formalism in the foundations of Knowledge Representation the first-degree-entailment of the relevance logic **R**.

This paper is a theoretical framework for our ongoing project to apply belief functions to multi-valued reasoning in AI (Ginsberg 1988), decision-making with bipolar information (Dubois and Prade 2008) and adding belief annotations to databases (Gatterbauer et al. 2009) with incomplete and/or inconsistent knowledge where structures of interest

are usually assumed to be no more Boolean but distributive. Our immediate task is to investigate information fusion in belief functions for distributive *bilattice reasoning* (Ginsberg 1988). Another ongoing further project, which is probably more of theoretical interest, is to develop a theory of belief functions for quantum structures (not necessarily distributive) which would relate quantum computation and AI (Ying 2010).

Acknowledgements

References

Anderson, B., and Belnap, N. 1975. *Entailment: the logic of relevance and necessity*. New Jersey: Princeton University Press.

Barthelemy, J.-P. 2000. Monotone functions on finite lattices: An ordinal approach to capacities, belief and necessity functions. In Fodor, J.; Baets, B.; and Perny, P., eds., *Preferences and Decisions under Incomplete Knowledge*. 195–208. Physica Verlag.

Davey, B., and Priestley, H. 2002. *Introduction to Lattices and Order*. Cambridge University Press, second edition.

Denoeux, T.; Younes, Z.; and Abdallah, F. 2010. Representing uncertainty on set-valued variables using belief functions. *Artif. Intell.* 174(7-8):479–499.

Dubois, D., and Prade, H. 2008. An introduction to bipolar representations of information and preference. *Int. J. Intell. Syst.* 23(8):866–877.

Dunn, M. 1986. Relevance logic and entailment. In Gabbay, D., and Guenthner, F., eds., *Handbook of Philosophical Logics*. Dordrecht: D. Reidel Publishing Company. 117–224.

Fagin, R., and Halpern, J. 1991. Uncertainty, belief, and probability. *Computational Intelligence* 7:160–173.

Fagin, R.; Halpern, J.; and Megiddo, N. 1990. A logic for reasoning about probabilities. *Inf. and Comp.* 87:78–128.

Fagin, R.; Halpern, J.; and Vardi, M. 1995. A nonstandard approach to the logical omniscience problem. *Artif. Intell.* 79(2):203–240.

Gatterbauer, W.; Balazinska, M.; Khoussainova, N.; and Suciu, D. 2009. Believe it or not: Adding belief annotations to databases. *PVLDB* 2(1):1–12.

Ginsberg, M. 1988. Multivalued logics: a uniform approach to reasoning in artificial intelligence. *Computational Intelligence* 4:265–316.

Grabisch, M., and Labreuche, C. 2005. Bi-capacities - i: definition, möbius transform and interaction. *Fuzzy Sets and Systems* 151(2):211–236.

Grabisch, M. 2009. Belief functions on lattices. *Int. J. Intell. Syst.* 24(1):76–95.

Halpern, J. 2005. *Reasoning about uncertainty*. MIT Press.

Levesque, H. 1984. A logic of implicit and explicit belief. In *Proc. National Conference on Artificial Intelligence (AAAI'84)*, 198–202.

Lin, J. 1996. A semantics for reasoning consistently in the presence of inconsistency. *Artif. Intell.* 86(1):75–95.

Masson, M., and Denoeux, T. 2011. Ensemble clustering in the belief functions framework. *Int. J. Approx. Reasoning* 52(1):92–109.

Priestley, H. 1970. Representation of distributive lattices by means of ordered stone spaces. *Bull. London Math. Soc.* 2:186–190.

Routley, R., and Routley, V. 1972. Semantics for first degree entailment. *Noûs* 3:335–359.

Shafer, G. 1976. *A Mathematical Theory of Evidence*. Princeton, N.J.: Princeton University Press.

Shafer, G. 1979. Allocation of probability. *The Annals of Probability* 7(5):827–839.

Smiley, M. F. 1940. A note on measure functions in lattice. *Bulletin of American Mathematical Society* 46(4):239–241.

Stanley, R. 1997. *Enumerative Combinatorics, Vol. 1*. Combradige Studies in Advanced Mathematics. Combridge University Press.

Urquhart, A. 1979. Distributive lattices with a dual homomorphic operation. *Studia Logica* 38(2):201–209.

Ying, M. 2010. Quantum computation, quantum theory and ai. *Artif. Intell.* 174(2):162–176.

Zhang, M., and Zhou, Z. 2007. Ml-knn: A lazy learning approach to multi-label learning. *Pattern Recognition* 40(7):2038–2048.