Information Fusion with Uncertainty Modeled on Topological Event Spaces

Roman Ilin
Sensors Directorate
Air Force Research Laboratory
Wright Patterson AFB, OH, USA
rilin325@gmail.com

Jun Zhang
Department of Psychology & Department of Mathematics
University of Michigan
Ann Arbor, MI, USA
junz@umich.edu

***Abstract*— We investigate probability and belief functions constructed on topological event spaces (without requiring complementation operation as in the definition of Borel sets). Anchored on the Lattice Theory, and making use of the correspondence of distributive lattice and topology, we propose a hierarchical scheme for modeling fusion of evidence based on constructing the lattice of topologies over a given sample space, where each topology encodes context for sensor measurement as specified by the basic probability assignment function. Our approach provides a rigorous mathematical grounding for modeling uncertainty and information fusion based on upper and lower probabilities (such as the Dempster-Shafer model).**

Keywords—Probability, Lattice Theory, Distributive Lattice, Topology, Belief Functions, Sensor Networks

I. Introduction

Classical probability theory prescribes a *sample space*, which is interpreted as the set of possible outcomes of an experiment or possible answers to some question. Denote this ground set (sample space) by Ω where any subset of Ω is called an *event*. Probabilities are real numbers between [0,1] assigned to each subset of this sample space, such that

A.1. $\mathrm{P}(\Omega) = 1$,

A.2. $(\forall A \subseteq \Omega)\ P(A) \geq 0$, and

A.3. $A \subseteq \Omega, \mathrm{B} \subseteq \Omega, \mathrm{A} \cap \mathrm{B} = \emptyset\ \Rightarrow P(A \cup B) = P(A) + P(B)$.

The last property, which in general is imposed upon countably many pairwise disjoint subsets, states that the probabilities of disjoint subsets simply add up. This means that we only need to know the probabilities assigned to individual elements ("atoms") of Ω and the probability of any other subset (=event) can be readily computed. With the additional axiom for conditional probability $P(\mathrm{A} \mid \mathrm{B}) = P(\mathrm{A} \cap \mathrm{B})/P(\mathrm{B})$, one obtains a powerful framework for uncertainty representation that serves as the basis of Bayesian inference. An equivalent expression for *A.3* above is:

A.3a. $A, \mathrm{B} \subseteq \Omega\ \Rightarrow P(A \cup B) = P(A) + P(B) - P(\mathrm{A} \cap \mathrm{B})$.

The weakness of Bayesian probabilities is in the requirement that all elements of the sample space need to be assigned some probability mass (which might take zero value). The Dempster-Shafer theory of belief functions removed this constraint of atomic assignment of probability; instead probability mass (called "basic probability assignment", abbreviated as *bpa*) can be assigned onto any subset of Ω. It was welcomed by researchers because it provided a better way of representing ignorance. The Dempster-Shafer framework reduces to Bayesian framework when basic probability assignment (*bpa*) is atomic, and to fuzzy probability when the *bpa* is supported on a chain of subsets. From *bpa's*, belief functions can be constructed, which are also real-valued functions defined on the powerset 2^{Ω}. The belief function is, nevertheless, non-additive over disjoint events in general, so *A.3* would not hold as long as there is non-atomic probability assignment. See [1] for a comprehensive treatment of Dempster-Shafer theory.

The nature of the event space and the probability/belief functions defined on it is crucial for information fusion. Since there is a natural partial order (i.e., set-inclusion relation) of subsets of the powerset 2^{Ω}, the algebra of the event space can be studied under the Lattice Theory, an abstract theory of partially ordered sets. Reference [2] provides an excellent introduction. From the lattice theoretic point of view, the powerset 2^{Ω} forms a Boolean lattice. It is then natural to relax the structure of the event space to that modeled by non-Boolean lattices, which may provide a more general modeling environment and more flexible tools. Several researchers recently began to consider this possibility by studying both probabilities and belief functions on a lattice [3] [4].

A special kind of non-Boolean lattice that provides richer mathematical structure for modeling the event space is through the concept of topological space, or sets endowed with a topology, where each open set corresponds to an "event" [5]**.** Topology [6] is a system of subsets of Ω that, compared with Borel sets (which is also a system of subsets of Ω), need not be closed under complementation operations. The powerset (which forms a Boolean lattice) is a topological space, corresponding to the discrete topology. In general, any topology corresponds to a complete, distributive lattice (for definition of "complete" and "distributive", see below), where open sets of a topological space form elements of the lattice. The study of probability and belief functions on topological event spaces amounts to the study of probability/belief functions on distributive lattices. In particular, the pseudo-complementation, which forms the crux of intuitionistic logic, is well-defined on finite distributive lattice and is related to the closure operation defined on a topology [7] [8]. Topological event spaces have the advantage of reducing combinatorial explosion of events as modeled by discrete topology and providing a richer semantics such as "neighborhood", "boundary", "limits" for describing the structure of the event space.

Topological spaces admit a relation of partial order by set inclusion. Thus, it can be shown that, for a given set Ω, all of the possible topologies $\mathfrak{L}^{\mathcal{T}} = \{\mathcal{T}: \mathcal{T}\ is\ a\ topology\ over\ \Omega\}$ form a lattice, called the lattice of topologies [9]. Here we use this lattice of topologies to provide a systematic way of organizing multiple topologies (i.e., multiple contexts) and combining the probabilities or belief functions defined across contexts.

This work proposes a hierarchical scheme for fusion of evidence based on the connections between probabilities/belief functions and topological spaces/lattices as described above. Our conceptualization will be illustrated with an example of evidence fusion for sensor networks.

II. Topological Event Spaces

Throughout this paper, we consider finite sets for simplicity. A *topological space*, or just a *topology*, is an ordered pair $(\Omega, \mathcal{T})$, where Ω is a set (where each element is called a *point*), and $\mathcal{T}$ is a collection of subsets of Ω, such that $\emptyset \in \mathcal{T}$, $\Omega \in \mathcal{T}$, and $\mathcal{T}$ is closed under finite intersections and arbitrary unions of its elements [6].

The elements of $\mathcal{T}$ are called *open sets*. Suppose $A \in \mathcal{T}$. The set-theoretical complement of the subset A is $A^c = \Omega \backslash A$; any subset A^c that is the complement of any open set is called a *closed* set. It is possible for a subset of Ω to be open and closed, referred to as *clopen*, or neither open nor closed. Indeed, consider the *discrete topology*, which consists of *all* subsets of Ω, that is the powerset 2^{Ω}. Under discrete topology every subset is clopen. The discrete topology is the largest topology (largest in the sense of containing the largest number of elements, $2^{|\Omega|}$) that can be defined on a set. The opposite side of the spectrum is the *indiscrete topology* consisting of only two elements: $\emptyset$ and Ω; both are clopen. If we consider an arbitrary topology, some of its open sets may be clopen.

Consider an arbitrary subset $B \subseteq \Omega$, with topology $(\Omega, \mathcal{T})$. The set B does not necessarily belong to the topology; however it can be related to the topology using the concepts of *interior* and *closure*. The *interior* of a set $B \subseteq \Omega$ is the largest open set contained in B. Of course, when B is an open set, its interior coincides with the set itself. The *closure* of a set $B \subseteq \Omega$ is the smallest closed set containing B. When B is a closed set, its closure is the set itself. The interior and closure of an arbitrary B are denoted, respectively, as $Int(B)$ and $Cl(B)$. The *boundary* of B is defined as $Cl(B) \backslash Int(B)$.

We are interested in topological spaces because they can be used to model events, which correspond to arbitrary subsets of the ground set. Ref [10] suggested decomposing any event under consideration into its interior, corresponding to all well-defined possibilities, and the boundary, corresponding to possibilities that are not well-defined but are not too far removed from the well-known events [5]. This kind of modeling becomes trivial and not interesting on the discrete topology (Boolean lattice) where every subset coincides with its closure and interior.

III. Lattice Theory

A comprehensive treatment of Lattice Theory can be found in [2], and here we only provide the basic definitions. A *partial order* on a set P is a binary relation, denoted by $\leq$ that is *reflexive*, *antisymmetric*, and *transitive*. Sets with partial order are called *posets*. Given a subset $S \subseteq P$, an element $x \in P$ is an *upper bound* of S if $(\forall s \in S), s \leq x$. The lower bound is defined analogously. The *least upper bound* is an element $x \in P$ such that x is an upper bound of S, and for any other upper bound y of S, $x \leq y$. The *greatest lower bound* is defined analogously.

Given any two elements, $x, y \in P$, we denote their *least upper bound* and dually, *greatest lower bound*, by $x \vee y$ and $x \wedge y$. These operations are respectively called *join* and *meet*.

A partially ordered set P such that join and meet exist for all pairs of its elements is called a *lattice*. The lattice is called *complete* if meet and join exist for any subset of P.

An alternative way of defining a lattice is to start with two binary operations (call them meet and join) on a set P that satisfy the properties of *associativity*, *commutativity*, *idempotency*, and *absorption*. See [2] for details. Then, the relation $\leq$ can be defined (i.e., induced) in terms of these two binary operations and it can be shown that this is precisely the partial order relation. Thus a lattice can be viewed as an *algebraic* structure as well as an *order* structure.

The algebraic point of view allows further classification of lattices by imposing additional identities to be satisfied by various lattice types. In particular, for all lattice elements a, b, c, the modular law $a \geq c \Rightarrow a \cap (b \cup c) = (a \cap b) \cup c$ is satisfied by *modular lattices*, and the distributive law $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$ is satisfied by *distributive lattices*. Distributivity implies modularity, but not vice versa.

A finite lattice always has the least element, or *bottom*, denoted by 0, and the greatest element (*top*), denoted by 1. An important operation on a lattice is that of *complementation*. For any element $a \in P$, its complement is an element $a' \in P$, such that $a' \wedge a = 0$ and $a' \vee a = 1$. Note that for an arbitrary element of P, its complement may not exist, or it may have multiple complements.

Finite lattice is often visualized as the Hasse diagram. Examples of such diagrams can be seen in Fig. 1-5. The nodes of the diagrams correspond to lattice elements and the edges correspond to the "cover" relations between the nodes: the presence of an edge between two elements a and b, such that a is located lower than b on the diagram, means that a is "covered" by b, that is $a \leq b$ and there is no other element c such that $a \leq c \leq b$. Using the Hasse diagram allows to explain the concept of *join-irreducible* elements of a lattice. An element a is join-irreducible if $a = b \vee c \Longrightarrow a = b\ or\ a = c$. On the diagram, the join-irreducible elements are those that have only one neighbor below that it covers. In Fig. 5 for example, the join-irreducible elements are drawn with thick ovals. The set of join-irreducible elements is denoted by J.

Lattice provides a richer modeling language than powerset. *Lattice of sets* is a lattice formed by some subset of the powerset of Ω, $P \subseteq 2^{\Omega}$, with the binary relation corresponding to set inclusion: for all $\mathrm{A}, \mathrm{B} \in \Omega, \mathrm{A} \leq \mathrm{B} \Leftrightarrow \mathrm{A} \subseteq$

B and the meet and join operations correspond to the set theoretical intersection and union. The top element is the set Ω, and the bottom element is the empty set $\emptyset$. Lattice complementation operation is just the set-theoretic complementation: $A' = \Omega \backslash A$. A lattice of sets is typically smaller than the powerset.

IV. Distributive Lattice and Pseudo-Complement

In a distributive lattice, it is known that an element can have at most one complement. A special kind of distributive lattice is a *Boolean lattice*, where each element has a unique complement. Any Boolean lattice is isomorphic to the lattice of sets $(2^{\Omega}, \subseteq)$, the powerset. Finite distributive lattice can be viewed as some (generally non-Boolean) lattice of sets, and this is expressed by Birkhoff's representation theorem [11]. It states that any finite distributive lattice is isomorphic to the lattice of down-sets of the set of its join irreducible elements [2].

As we mentioned, not every element of a distributive lattice has a complement. A weaker notion of a complement, called *pseudo-complement*, can be defined. For a lattice element $a \in P$, its *pseudo-complement* is an element $a^* \in P$, such that $a^* \wedge a = 0$ and a^* is the largest such element to do so (i.e., meet a to 0). It turns out that every element of a finite distributive lattice has a pseudo-complement, though it may not be true that $a^* \vee a = 1$.

There is a natural link between the operations of closure and interior of a set endowed with a topology and the pseudo-complement. Suppose A is an element of a lattice P of sets, $A \in P$, where $P \subseteq 2^{\Omega}$. The corresponding topological space is given by (Ω, P). Then the following is true:

$$A^* = Int\big(Cl(A^c)\big). \qquad (1)$$

This formula provides a topological interpretation to the pseudo-complement operation.

V. Lattice of Topologies

An order relation can be defined between two topologies $\mathcal{T}_1$ and $\mathcal{T}_2$ over a common ground set Ω: $\mathcal{T}_1 \leq \mathcal{T}_2 \Leftrightarrow \mathcal{T}_1 \subseteq \mathcal{T}_2$, where inclusion relationship is with respect to individual open sets in a topology as distinct points of the lattice. This relation induces a lattice of all possible topologies of a given set Ω [9]. The top element of this set is the discrete topology, and the bottom element is the indiscrete topology. This lattice is *complete*, which means that any subset of lattice elements has a join and a meet. This lattice is in general neither distributive nor modular. Individual topologies in this lattice are ordered by refinement, with finer topologies located above coarser ones. Fig. 4 shows the lattice of topologies for a 3-element ground set $\Omega = \{a, b, c\}$.

VI. Belief Functions and Upper-Lower Probability

Dempster-Shafer theory starts with a finite set Ω of possible answers to a question, referred to as the *frame of discernment*, [1]. The events correspond to subsets of Ω, with their uncertainty quantified by the *basic probability assignment*, which is a function $m: 2^{\Omega} \rightarrow [0,1]$ such that

$$m(\emptyset) = 0, \qquad (2)$$

and

$$\sum_{A \subseteq \Omega} m(A) = 1. \qquad (3)$$

The key idea here is that not all subsets of Ω encode inferred information; some are basic events and can be assigned basic (elementary) probability mass. Those subsets that we assign non-zero $m()$ are called *focal elements*. This structure is called a *body of evidence*.

A belief function $Bel: 2^{\Omega} \rightarrow [0,1]$ is computed from the *bpa* as follows:

$$Bel(A) = \sum_{B \subseteq A} m(B), \qquad (4)$$

where $A \subseteq \Omega$. The *bpa* can be recovered from the belief function using the following formula,

$$m(A) = \sum_{B \subseteq A} (-1)^{|A \backslash B|} Bel(B). \qquad (5)$$

Thus there is a one-to-one correspondence between m and Bel. The belief functions satisfy the following properties:

$$Bel(\emptyset) = 0, \qquad Bel(\Omega) = 1,$$

$$Bel\left(\bigcup_{i=1}^{n} A_i\right) \geq \sum_{I \subset \{1,2..n\}, I \neq \emptyset} (-1)^{|I|+1} Bel\left(\bigcap_{i \in I} A_i\right). \qquad (6)$$

The last property means that belief functions are non-additive even in the case of two disjoined subsets A and B: $Bel(A \cup B) \geq Bel(A) + Bel(B)$, in contrast to additivity axiom *A.3* of a probability function.

The plausibility function is defined as follows:

$$Pl(A) = \sum_{\{B | A \cap B \neq \emptyset\}} m(B) = 1 - Bel(A^c). \qquad (7)$$

Given two events A and B, we say that event B *implies* event A when $B \subseteq A$. The belief function of B therefore is the sum of the evidence that exists for all events that imply B. The plausibility function of B is the sum of the evidence for all events that *may* imply B. It is easy to see that $Bel(A) \leq Pl(A)$, for any set A. These two quantities form an interval that characterizes uncertainty based on available evidence (*interval of uncertainty*).

If, and only if, the focal elements of *bpa* are 1-element subsets ("singletons") of Ω, Bel and Pl become equivalent and reduce to a probability function. The belief function theory offers several combination rules interpretable in different ways and each has its own advantages and disadvantages. The original Dempster's rule [1] has been criticized in [12] and later in [13]. This criticism gave rise to new rules avoiding problems caused by unreliable sources, conflicting evidence, and more recently, by overlapping evidence [14], [15].

Given two bodies of evidence with *bpa*'s m_1 and m_2 that need to be combined or "fused", the *conjunctive rule* is given by the following formula:

$$m_{1\cap 2}(A) = \sum_{B\cap C=A} m_1(B)m_2(C). \tag{8}$$

The result of (8) needs to be normalized in order to satisfy (3). Dividing $m_{1\cap 2}$ by K achieves normalization

$$K = \sum_{B\cap C=\emptyset} m_1(B)m_2(C). \tag{9}$$

Formulas (8) and (9) correspond to the Dempster's rule of combination. The un-normalized version (8) of the Dempster's rule is used in Transferable Belief Models (TBM) [16].

The *disjunctive combination rule* provides a different way of combining evidence [17]:

$$m_{1\cup 2}(A) = \sum_{B\cup C=A} m_1(B)m_2(C). \tag{10}$$

This rule can also be used in both normalized and un-normalized way. This rule is justified in situations when we suspect that only one of the sources of evidence is reliable.

The *commonality* function is defined through the *bpa*,

$$Q(A) = \sum_{B\supseteq A} m(B). \tag{11}$$

The conjunctive combination rule is conveniently expressed in terms of Q as follows

$$Q_{1\cap 2}(A) = Q_1(A)Q_2(A). \tag{12}$$

Symmetrically, the disjunctive rule is expressed using Bel,

$$Bel_{1\cup 2}(A) = Bel_1(A)Bel(A). \tag{13}$$

When considering multiple sources of information, the reliability of each source needs to be quantified. This can be accomplished through discounting. If the reliability of source is measured by some number $w \in [0,1]$, the bpa is discounted using the rule

$$m^w = wm + (1-w)m_\Omega, \tag{14}$$

where m_Ω is the vacuous *bpa,* that is the *bpa* with all the mass allocated to Ω [14].

In order to make rational decisions, the belief functions have to be transformed into probability functions [18]. The *pignistic probability* is derived from a bpa function m as follows:

$$Bet_m(a) = \sum_{\{A|a\in A\}} \frac{m(A)}{|A|} \tag{15}$$

The reason for requiring such a transformation is found in the expected utility theory [19].

VII. Belief Functions on a Lattice

In Dempster-Shafer theory, belief functions are defined on Boolean lattices $(2^\Omega, \subseteq)$. It turns out that the belief function framework can be generalized to the lattice setting [3]. In order to show this we need to define the Möbius transform [20].

For any poset $(P, \leq)$, and a function $f: P \to \mathbb{R}$, the Möbius transform [20] is the function $m: P \to \mathbb{R}$ such that

$$f(x) = \sum_{y\leq x} m(y). \tag{16}$$

This function can be computed as follows:

$$m(x) = \sum_{y\leq x} \mu(y,x)f(y), \tag{17}$$

using the Möbius function, defined recursively as follows:

$$\mu(x,y) = \begin{cases} 1 & if\ x = y \\ -\sum_{x\leq t<y} \mu(x,t) & if\ x < y \\ 0 & otherwise \end{cases}. \tag{18}$$

It is easy to show that the transforms (4) and (5) between the belief function and the basic probability assignment are special cases of the Möbius transform on a Boolean lattice.

Given a lattice $(P, \leq)$, a function $Bel: P \to [0,1]$ is called a (generalized) belief function if $Bel(0) = 0$, $Bel(1) = 1$, Bel is strictly monotonic, and its Möbius transform is non-negative [21]. When a lattice admits a belief function, we can define a *bpa* on the lattice as well.

Under what conditions do belief functions and/or probability functions exist on a lattice? This was investigated by [21], and later [3] who showed that belief functions always exist on a lattice. Ref. [4] further showed that on a distributive lattice, a probability function always exists if the non-zero *bpa* is assigned only to the *join-irreducible* elements of the lattice; in this case the belief function reduced to a probability function. These conclusions provide the mathematical grounding of our hierarchical scheme for modeling uncertainty, as discussed next.

VIII. A Hierarchical Scheme for Uncertainty Inference

Bodies of evidence (BOE) appear naturally in the context of sensor networks. Each sensor provides a detection of an event, and the collection of sensors (forming the network) provides a collection of possible evidence (not necessarily spanning the entire event space) by that sensor network. Multiple sensor networks may be deployed and the corresponding bodies of evidence need to be combined. Our scheme consists of the following steps.

1. Each sensor network *i (i*=1,2 in the example*)* is modeled by a topological event space (distributive lattice) $\mathcal{T}_i$. The details of such modeling are beyond the scope of this work but the main ideas follow [10]. Sensor measurements within a network can be modeled as providing a topological basis for generating the topology, while the open sets of the topology correspond to queries that can be answered with the given sensor data. The notions of topological closure and interior play a prominent role in relating the event space (with not less than 2 but as few elements as one may desire) to the combinatorially explosive powerset. The initial *bpa*'s m_i are transformed into the new *bpa*'s defined on the topology (by filling up with zero values). We denote this symbolically:

$$m_i \to m_i^{\mathcal{T}_i} \qquad (19)$$

2. Various topologies, each modeling a different sensor network, are modeled as elements of the lattice $\mathfrak{L}^{\mathcal{T}}$ of topologies. The lattice $\mathfrak{L}^{\mathcal{T}}$ is completed by creating new topologies where the join operation gives the smallest topology that includes the two to-be-combined topologies.

$$\mathcal{T}_{12} = \mathcal{T}_1 \vee \mathcal{T}_2 \qquad (20)$$

3. Each of the constituent topologies comes with a reliability number w_1 and w_2. We generate the *bpa*'s for the elements of the combined topology $\mathcal{T}_{12}$ by weighted average:

$$m_{12} = \frac{w_1 m_1^{\mathcal{T}_1} + w_2 m_2^{\mathcal{T}_2}}{w_1 + w_2} \qquad (21)$$

4. We then transform m_{12} to become probability mass assignments only to join-irreducible elements by "flowing" mass downwards on the lattice until the join-irreducible elements are reached (see Fig. 5):

$$m_{12} \to \widehat{m}_{12}. \qquad (22)$$

Starting at the top element, the mass is redistributed ("split") between its children, and the process is repeated recursively. After this flow-down procedure converges, the resulting $\widehat{m}_{12}$ on join-irreducible elements will generate a probability function P_{12}

$$P_{12}(x) = \sum_{y \le x} \widehat{m}_{12}(y)\,. \qquad (23)$$

We can show (see Appendix) that the value of P_{12} lies within the original uncertainty interval (between the belief function and plausibility function) for all nodes $x \in \mathcal{T}_{12}$.

5. Decisions are made based on P_{12} which is a probability function on the distributive lattice $\mathcal{T}_{12}$ in the sense of *A.1-A.3* and thus can be used directly for utility-based decision making. However, P_{12} is a belief function on the full Boolean lattice corresponding to discrete topology (Boolean algebra) over the same underlying ground set. This highlights the difference between our topological approach and the Dempster-Shafer belief function approach. In the next section, we illustrate our ideas with a simple example.

IX. AN EXAMPLE OF SENSOR FUSION

Consider the ground set with three possible events: $\Omega = \{a, b, c\}$. Suppose that two sensor networks provided the bodies of evidence corresponding to rows 1 and 5 in Table 1. The focal sets of both BOE's are ordered by set inclusion and are represented in Fig. 1 by Hasse diagrams of distributive lattices, which are two topologies over Ω, denoted by $\mathcal{T}_1$ and $\mathcal{T}_2$. In this simple case, each lattice is actually a chain and hence all elements (except ∅) are join-irreducible. We also assume that both sources have equal reliability weights of 0.5 each.

The topological spaces corresponding to the meet and join of $\mathcal{T}_1$ and $\mathcal{T}_2$ are shown in Fig. 2. The meet is the indiscrete topology $\mathcal{T}_0$ and the join is the topology $\mathcal{T}_{12}$ that contains all the elements from $\mathcal{T}_1$ and $\mathcal{T}_2$ and their unions and intersections. The relationship between the four topological spaces can be visualized in a Hasse diagram shown in Fig. 3. These four topologies, when taken together, form a lattice of their own, denoted as $\mathfrak{L}^{\mathcal{T}}$. Fig. 4 illustrates how this lattice of topologies $\mathfrak{L}^{\mathcal{T}}$ is embedded in the complete lattice of topologies of a three-element set, with 29 elements (with discrete topology at the top and indiscrete topology at the bottom).

Our next step is to compute *bpa*'s for the combined topology $\mathcal{T}_{12}$ using (21). Note that in this new topology the computed probability mass is not necessarily restricted to join-irreducible elements. The combined *bpa* is given in row 15. Our proposed flow-down algorithm for the *bpa's* will redistribute probability mass to join-irreducible elements of the lattice/topology (Fig. 5). We split the probability mass *equally* between the child nodes of each join-reducible element and propagate the mass to the children nodes it covers. The transformed *bpa* for $\mathcal{T}_{12}$, denoted by $\widehat{m}_{12}$ is shown in row 18 of Table 1. We show the original uncertainty interval $[Bel_{12}, Pl_{12}]$, before the "flow-down" procedure, in lines 16 and 17, and the resulting probability function P_{12} in line 19, after the flow-down. As expected, $Bel_{12} \le P_{12} \le \mathrm{Pl}_{12}$. It is easily verified that P_{12} is indeed additive, i.e. satisfies A.3a, among elements of $\mathcal{T}_{12}$. We also compute the belief, plausibility, and commonality functions on the original bodies of evidence and their combination using the regular conjunctive and disjunctive rules of the Dempster-Shafer framework. In Table 1, rows 2, 3, and 4 give the belief, plausibility, and commonality for the first BOE and rows 6, 7, and 8 give the corresponding quantities for the second BOE. Row 9 is the *bpa* obtained from the normalized conjunctive combination rule, and rows 10 and 11 have the corresponding belief and plausibility.

Rows 10 and 11 provide the uncertainty interval for each of the possible events. Rows 12, 13, and 14 contain the combined *bpa*, belief and plausibility respectively, using the disjunctive combination rule.

TABLE 1 COMBINATION OF EVIDENCE IN THE TOPOLOGICALLY ORGANIZED BODIES OF EVIDENCE. NOTE THAT EMPTY CELLS CONTAIN ZEROS.

		1	2	3	4	5	6	7	8
		$\emptyset$	$\{a\}$	$\{b\}$	$\{a,b\}$	$\{c\}$	$\{a,c\}$	$\{b,c\}$	$\{a,b,c\}$
1	m_1			0.6	0.2				0.2
2	Bel_1			0.6	0.8			0.6	1
3	Pl_1		0.4	1	1	0.2	0.4	1	1
4	Q_1	1	0.4	1	0.4	0.2	0.2	0.2	0.2
5	m_2						0.2		0.8
6	Bel_2						0.2		1
7	Pl_2		1	0.8	1	1	1	1	1
8	Q_2	1	1	0.8	0.8	1	1	0.8	0.8
9	$m_{1\cap 2}$		0.0455	0.5455	0.1818		0.0455		0.1818
10	$Bel_{1\cap 2}$		0.045	0.5455	0.7727		0.0909	0.5455	1
11	$Pl_{1\cap 2}$		0.4545	0.9091	1	0.2273	0.4545	0.9545	1
12	$m_{1\cup 2}$								1
13	$Bel_{1\cup 2}$								1
14	$Pl_{1\cup 2}$		1	1	1	1	1	1	1
15	m_{12}			0.3	0.1		0.1		0.1
16	Bel_{12}			0.3	0.4		0.1	0.3	1
17	Pl_{12}		0.7	0.9	1	0.6	0.7	1	1
18	$\hat{m}_{12}$		0.1750	0.4750			0.35		
19	P_{12}		0.1750	0.4750	0.65		0.5250	0.4750	1
20	$Bet_{1\cap 2}$		0.2197	0.6970	0.9167	0.0833	0.3030	0.7803	1
21	$Bet_{1\cup 2}$		0.3333	0.333	0.6667	0.3333	0.6667	0.6667	1

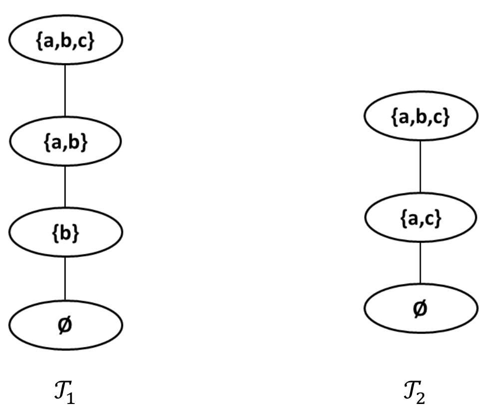

Figure 1 The bodies of evidence and their corresponding distributive lattices (chains in this case).

Notably, in this case the disjunctive evidence combination rule results in vacuous *bpa* indicating we are completely ignorant about the events. The beliefs are transformed to the corresponding pignistic probabilities using (15) and shown in rows 20 and 21 of Table 1.

I. CONCLUSION AND DISCUSSION

This work explored the connections between the theory of belief functions, probability theory, theory of topological spaces, and the lattice theory. The topological event space is a relaxation of Borel sets in the traditional probability theory via removal, in its sigma-algebra, of complementation operation on the event space, in accord with intuitionistic logic.

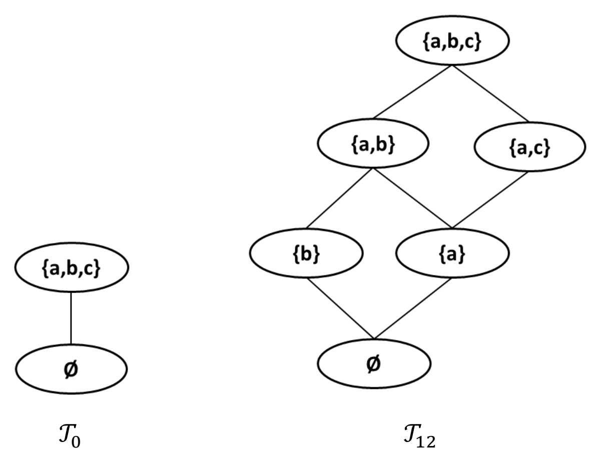

Figure 2 The meet and join of the bodies of evidence $\mathcal{T}_1$ and $\mathcal{T}_2$ from Fig. 1

This was the approach of L. Narens [10]. Here we make use of recent results about belief functions on lattices in general [21], [3] and distributive lattices (which are essentially topologies) in particular [4] to endow basic probability assignments and represent measurement by a sensor network.

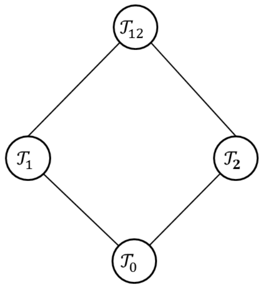

Figure 3 The lattice of topologies corresponding to the bodies of evidence from Fig. 1

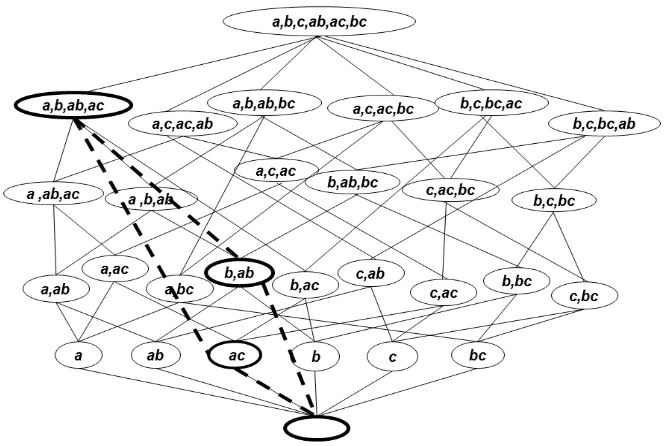

Figure 4 The lattice of topologies of three-element set. Each node shows the elements of the corresponding topology with the empty set and the set Ω omitted for bettet readability. The lattice from Fig. 3 is embedded in the complete lattice and shown by thick lines.

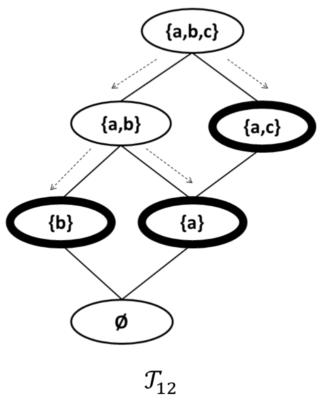

Figure 5 The combined topology and its join-irreducible elements (thick ovals). The dashed arows show the direction of the flow of probability mass towards the join-irreducible elements of the lattice.

We devise a flow-down algorithm for basic probability assignment only on join-irreducible elements of a distributive lattice (topological event space). We further invoke the lattice of topologies for representing different sensor networks, which are treated as different contexts for uncertainty reasoning. Our scheme extends both the Dempster-Shafer scheme by providing i) a principled way for basic probability assignments; and ii) a hierarchical setting for switching and integrating across contexts, and the topological event space approach of Narens by i) considering the upper-lower probabilities and ii) the lattice of topologies that have not so far been used. This forms our main innovative claim. The advantage of our scheme is its ability to model rich sets of structures using different topologies for the same underlying sample space and at the same time to reduce the size of the event space by not having to consider the full Boolean (combinatorial) structure. Our approach yields a probability (modular) function on the topological event space. This is fundamentally different from the Dempster-Shafer theory, which only affords an uncertainty interval on such spaces (note that the same function P, is a belief function and hence super-modular in D-S theory). Future work will involve the investigation of other types of lattices, such as orthocomplemented modular lattice that underlie quantum probability theory, with a goal of achieving a unifying perspective for uncertainty management and integration of bodies of evidence from sensor networks. We will also compare our approach to that of fuzzy rough set theory [22].

ACKNOWLEDGMENT

Support for this research was provided by the Air Force Office of Scientific Research, Department of the Air Force, grant number 14RY04COR. We utilized the Matlab toolbox from Arnaud Martin [23] for belief function computations.

REFERENCES

[1] R. R. Yager and L. Liu, Eds., Classic Works of the Dempster-Shafer Theory of Belief Functions, Springer, 2008.

[2] B. Davey and H. Priestley, Introduction to Lattices and Order, Cambridge: University Press, 2002.

[3] M. Grabisch, "Belief functions on lattice," *ArXiv preprint 0811.3373v1,* 2008.

[4] C. Zhou, "Belief functions on distributive lattices," in *Proceedings of the 26th AAAI Conference on Artificial Intelligence*, Toronto, 2012.

[5] L. Narens, Theories of Probability: An Examination of Logical and Qualitative Foundations, London: World Scientific, 2007.

[6] J. L. Kelley, General Topology, Ishi Press, 2008.

[7] J. McKinsey and A. Tarski, "The algebra of topology," *Ann. of Math.,* no. 45: 141191., 1944.

[8] J. McKinsey and A. Tarski, "On closed elements in closure algebras,," *Ann. of Math,* no. 47: 122162., 1946.

[9] R. E. Larson and S. J. Andima, "The lattice of topologies: a survey," *J. Math.,* vol. 5, no. 2, 1975.

[10] L. Narens, "A foundation for support theory based on a non-Boolean event space," *J. Math. Psychology,* vol. 53, no. 5, pp. 399-407, 2009.

[11] G. Birkhoff, "On the combination of subalgebras," *Proc.*

Cambridge Philos. Soc., no. 29: 441-464., 1933.

[12] L. A. Zadeh, "On the validity of Dempster's rule of combination," Memo M79/24, University of California, Berkeley, USA, 1979.

[13] J. Dezert, P. Wang and A. Tchamova, "On the validity of Dempster-Shafer Theory," in *Information Fusion (FUSION), 2012 15th International Conference on* , 2012.

[14] T. Denoeux, "Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence," *Artificial Inteligence,* no. 172, pp. 234-264, 2008.

[15] E. Blasch, J. Dezert and B. Pannetier, "Overview of Dempster-Shafer and Belief Function Tracking Methods," in *Signal Processing, Sensor Fusion, and Target Recognition XXII, Proc. of SPIE*, 2013.

[16] P. Smets, "The combination of evidence in the transferable belief model," in *Pattern Analysis and Machine Intelligence, IEEE Transactions on* , 1990.

[17] P. Smets, "Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem," *International Journal of approximate reasoning ,* vol. 9, no. 1, pp. 1-35, 1993.

[18] P. Smets, "Constructing the Pignistic Probability Function in a Context of Uncertainty," in *UAI*, 1989.

[19] J. Von Neumann and M. Oskar, Theory of Games and Economic Behavior (60th Anniversary Commemorative Edition)., Princeton university press, 2007.

[20] G. C. Rota, "On the foundations of combinatorial theory I. Theory of Möbius functions.," *Probability theory and related fields,* vol. 2, no. 4, pp. 340-368, 1964.

[21] J. P. Barthélemy, "Monotone functions on finite lattices: an ordinal approach to capacities, belief and necessity functions," in *Preferences and decisions under incomplete knowledge*, Physica-Verlag , 2000, pp. 195-208.

[22] C. Cornelis, J. Medina and N. Verbiest, "Multi-adjoint fuzzy rough sets: Definition, properties and attribute selection," *Int. J. Approx. Reasoning,* vol. 55, no. 1, pp. 412-426, 2014.

[23] A. Martin, "Implementing general belief function framework with a practical codification for low complexity," in *Advances and Applications of DSmT for Information Fusion*, Rehoboth, American Research Press , 2009, pp. 217-273.

APPENDIX

"Flow-down" Procedure. Our algorithm works by passing probability mass $m(x)$ from a join-reducible node x to its children denoted as the set $\mathcal{C}(x)$, one node at a time in top-to-bottom fashion until the probability mass is carried by join-irreducible elements of the lattice only. Note that for each iteration step, x covers each of its child nodes $y \in \mathcal{C}(x)$, by definition of covering, and $|\mathcal{C}(x)| \geq 2$ by the definition of join-reducibility of x. The flow-down will stop upon reaching a join-irreducible element, and the iteration will terminate in finite number of steps.

Lemma. After each iteration of the "flow-down" procedure, yielding a new *bpa* $\widehat{m}$, the resulting belief and plausibility functions $\widehat{\mathrm{Bel}}$ and $\widehat{\mathrm{Pl}}$, satisfy:

$$\mathrm{Bel(x)} \leq \widehat{\mathrm{Bel}}(\mathrm{x}) \leq \widehat{\mathrm{Pl}}(\mathrm{x}) \leq \mathrm{Pl(x)},$$

where the belief and plausibility functions correspond to m and $\widehat{m}$ given by (4) and (7).

Proof. Let $\mathfrak{D}(x) = \{x' | x' \leq x\}$ denote the principal downset of a lattice element x to be "flown-down". Consider an arbitrary lattice element y.

(i) Proof that $Bel(x) \leq \widehat{Bel}(x)$

By definition, the belief function of y is the sum of the mass assigned to elements of $\mathfrak{D}(y)$, where $\mathfrak{D}(y) = \bigcup_{z \in \mathcal{C}(y)} D(z)$.

Case 1: $x \leq y$. So $\{\{x\} \cup \mathcal{C}(x)\} \subseteq \mathfrak{D}(y)$ and $\mathrm{Bel(y)} = \widehat{\mathrm{Bel}}(\mathrm{y})$ since total mass assigned to $\{x\} \cup \mathcal{C}(x)$ does not change.

Case 2: $\exists z \in \mathcal{C}(x), y \leq z$, i.e. y is below one of x's children. $\mathrm{Bel}(y) = \widehat{\mathrm{Bel}}(y)$ since the mass within $\mathfrak{D}(y)$ did not change.

Case 3: $\exists z \in \mathcal{C}(x), z \in \mathfrak{D}(y), x \notin \mathfrak{D}(y)$, i.e. one of the children belongs to $\mathfrak{D}(y)$ and y is not comparable with x. Since $m(z) \leq \widehat{m}(z)$ for such z, $\mathrm{Bel(y)} \leq \widehat{\mathrm{Bel}}(\mathrm{y})$.

Case 4: $\forall z \in \mathcal{C}(x), y \notin \mathfrak{D}(z), z \notin \mathfrak{D}(y)$, which means y is not in the downset of any of the children and no child is in the downset of y. $\mathrm{Bel(y)} = \widehat{\mathrm{Bel}}(\mathrm{y})$, since the mass within $\mathfrak{D}(y)$ does not change.

(ii) Proof that $\widehat{Pl}(x) \leq Pl(x)$

By definition, plausibility of y is the sum of the mass assigned to $S = \{x | x \wedge y \neq \emptyset\} = \{x | \mathfrak{D}(x) \cap \mathfrak{D}(y) \neq \emptyset\}$.

Case 1: $x \in \mathfrak{D}(y)$, this implies $\forall z \in \mathcal{C}(x), z \in \mathfrak{D}(y)$. In this case $\widehat{\mathrm{Pl}}(\mathrm{y}) = \mathrm{Pl(y)}$ because total mass in S remains unchanged.

Case 2: $x \notin \mathfrak{D}(y), \forall z \in \mathcal{C}(x), z \in \mathfrak{D}(y)$, again $\widehat{\mathrm{Pl}}(\mathrm{y}) = \mathrm{Pl(y)}$ since $m(x) = \sum_{z \in \mathcal{C}(x)} \widehat{m}(z)$.

Case 3: $\forall z \in \mathcal{C}(x), \mathfrak{D}(z) \cap \mathfrak{D}(y) = \emptyset$. No change in $\mathrm{Pl(y)}$.

Case 4: $\exists z \in \mathcal{C}(x), \mathfrak{D}(z) \cap \mathfrak{D}(y) \neq \emptyset$. Since $\widehat{m}(z) \geq m(z)$, this implies $\widehat{\mathrm{Pl}}(\mathrm{y}) \leq \mathrm{Pl(y)}$. In words, some mass has been transferred outside of the set S.

(iii) Proof that $\widehat{Bel}(x) \leq \widehat{Pl}(x)$.

Plausibility function is always greater of equal to the corresponding belief function. Hence the claim of the Lemma has been proved. Therefore,

Proposition. The values of the probability function $\widehat{P}$ resulting after the "flow-down" procedure terminates lie within the original interval of uncertainty.