International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems
Vol. 4, No. 4 (1996) 307–329

AGGREGATION AND COMPLETION OF RANDOM SETS WITH DISTRIBUTIONAL FUZZY MEASURES

CLIFF JOSLYN
Computer and Information Research
Los Alamos National Laboratory

Mail Code B 265, Los Alamos National Laboratory
Los Alamos, NM 87545, USA
E-mail: joslyn@lnl.gov

Received 28 February 1995
Revised 28 August 1996

The two known information theories, probability and possibility theory, are based on t-conorm decomposable fuzzy measures, so that bijective mappings exist between their set-valued measures and their point-valued distributions. Further, their random set (Dempster-Shafer evidence theoretical) interpretations have simple topological structures, with bijective mappings between the subset focal elements and the point singletons. We introduce the concepts of distributional and aggregable random sets and random set completion, and first use them as a model in which to cast probability and possibility measures and distributions. Then, towards the goal of deriving new forms of information theory, general Sugeno conorm decomposable fuzzy measures and ring-like aggregable random sets with set-intersection structural aggregation are examined, but it is shown that in these two cases no new information theories are forthcoming.

Keywords: General information theory; distributional random sets; complete random sets; decomposable fuzzy measures.

1. Introduction

Since the introduction of fuzzy sets [38] and evidence theory [4, 32] in the mid-1960's there has been a proliferation of mathematical methods for the representation of uncertainty and information which generalize beyond probability theory [25]. Recent years have seen increased efforts to synthesize these methods and provide a more coherent, general framework from which they can all be generated [8, 26]. We call this overall field General Information Theory (GIT) [24].

Information theories are essentially a part of the broad field of fuzzy [33] or general measure theory [35], which generalize additive probability measures to a number of different non-additive cases. The ideas in this paper were motivated by the introduction specifically of possibility theory as the first alternative, non-probabilistic form of information theory [23], and thus as a branch of GIT [3, 6]. Possibility theory was originally formulated by Zadeh [39] strictly in relation to fuzzy theory, but researchers are now developing it as an independent field, although still deeply related to both fuzzy and probability theories [2, 17]. For example, possibilistic correlates of histograms, sample statistics, entropy measures, Markov processes, and Monte-Carlo methods are available [12, 14, 15, 19, 36].

Whereas possibility theory is logically independent of probability theory, they are related in GIT: both probability and possibility measures arise in evidence theory as t-conorm decomposable fuzzy measures defined on random sets which always have certain simple topologies; and both probability and possibility distributions are fuzzy sets derived as distributions of those respective fuzzy measures. The natural question which arises is if there are *other* forms of information theory? Are there fuzzy measures which are decomposable for other conorms, which have other simple random set topologies, and yield other distributions which can in turn be used to build new information theoretical structures?

In this paper we will first introduce the concept of a (finite) distributional fuzzy measure as a slight generalization of a t-conorm decomposable fuzzy measure. These are very valuable because while in general the size of the domain of a fuzzy measure $\nu: 2^{\Omega} \mapsto [0,1]$ on a finite universe Ω grows exponentially with $|\Omega|$, distributional fuzzy measures can be constructed from their distributions $q^{\nu}: \Omega \mapsto [0,1]$, whose domains grow only linearly with $|\Omega|$.

Finite random sets are then introduced as set-valued random variables isomorphic to Dempster-Shafer bodies of evidence. Random sets act as a general context from which to generate fuzzy measures and fuzzy sets. They are simple hybrid structures, combining the randomness of a stochastic variable with the nonspecificity of subsets having variable size, shape, and mutual overlap. Moreover, they provide a mechanism to ground both fuzzy and possibility theory on an *objective* basis through the collection of set-valued observations [13, 18, 19].

Aggregable random sets are defined as those where focal elements (those subsets with positive probability) map to specific elements (or singleton subsets) of the universe. Aggregable random sets are sparsely populated, and may have distinctive topologies, so that, similar to distributional measures, knowledge of the structure of the focal elements is obtainable from the singletons, and vice versa. Complete random sets are defined as those where this mapping is a bijection, so that focal elements can be constructed from the singletons uniquely. Probability and possibility theory are then cast in terms of this model, and the topologies and properties of their aggregable and complete random sets identified.

The task is then to search out new forms of information theory by considering both other decomposable fuzzy measures and other aggregable random set topologies. We consider general Sugeno conorm distributional fuzzy measures, and show that for most universes of discourse their random sets are non-aggregable. Then we consider set difference as a new structural aggregation function. This yields random sets with ring-structured topologies, but for many universes of discourse their fuzzy measures are non-distributional.

Assume throughout the paper a finite universe $\Omega := \{\omega_i\}, 1 \leq i \leq n$.

2. Fuzzy Measures and Random Sets

2.1. *Distributional and decomposable fuzzy measures*

Definition 1 (Finite Fuzzy Measure, Normal Fuzzy Measure) [33] If Ω is finite, then $\nu: 2^\Omega \mapsto [0,1]$ is a fuzzy measure if $\nu(\emptyset) = 0$ and $\forall A, B \subseteq \Omega, A \subseteq B \rightarrow \nu(A) \leq \nu(B)$. A finite fuzzy measure ν is normal if $\nu(\Omega) = 1$.

In the sequel assume all fuzzy measures are finite.

Definition 2 (Fuzzy Measure Trace) Given a fuzzy measure ν, then the function $q^\nu: \Omega \mapsto [0,1]$ with $q^\nu(\omega) := \nu(\{\omega\})$ is called the trace of ν. In discrete form denote the vector $\vec{q}^\nu := \langle q_i^\nu \rangle$ where $q_i^\nu := q^\nu(\omega_i), 1 \leq i \leq n$. We will use just q or $\vec{q} := \langle q_i \rangle$ when ν is unambiguous.

Definition 3 (Operator) A function $\oplus: \mathbb{R}^2 \mapsto \mathbb{R}$ is an operator if $\langle \mathbb{R}, \oplus, 0 \rangle$ is an Abelian monoid: $\oplus$ is commutative and associative with identity 0. Further, for $x, y \in \mathbb{R}$ denote $x \oplus y := \oplus(x,y)$; and for $A := \{x_1, x_2, \ldots, x_n\} \subseteq \mathbb{R}$ and $B := \{x_j, x_{j+1}, \ldots, x_k\} \subseteq A$ with $1 \leq j \leq k \leq n$, using the associativity of $\oplus$, define the operator notation

$$\bigoplus_{x_i \in A} x_i = \bigoplus_{i=1}^{n} x_i := x_1 \oplus x_2 \oplus \cdots \oplus x_n,$$

$$\bigoplus_{x_i \in B} x_i = \bigoplus_{i=j}^{k} x_i := x_j \oplus x_{j+1} \oplus \cdots \oplus x_k,$$

with $\bigoplus_{i=j}^{j} x_i := x_j$.

Definition 4 (Distributional Fuzzy Measure) For a fuzzy measure ν and operator $\oplus$, if

$$\forall A \subseteq \Omega, \quad \bigoplus_{\omega_i \in A} q_i^\nu = \nu(A), \tag{5}$$

then ν is called $\oplus$-distributional and q^ν is called the distribution of ν.

Note that for a normal $\oplus$-distributional fuzzy measure,

$$\bigoplus_{i=1}^{n} q_i = 1. \tag{6}$$

Distributionality of a fuzzy measure forces $\oplus$ to be a t-conorm $\sqcup$. Distributionality then becomes equivalent to the traditional property of t-conorm decomposability.

Definition 7 (t-Conorm) [28] An operator $\oplus$ is a t-conorm $\sqcup$ if $\sqcup: [0,1]^2 \mapsto [0,1]$ and $\sqcup$ is weakly monotonic, so that

$$\forall a, b, c, d \in [0,1], \qquad a \leq c, b \leq d \quad \rightarrow \quad a \sqcup b \leq c \sqcup d.$$

Theorem 8 If a finite fuzzy measure ν is $\oplus$-distributional then $\oplus$ is t-conorm $\sqcup$.

Proof: Let ν be $\oplus$-distributional. We need to establish that $\oplus\colon [0,1]^2 \mapsto [0,1]$ and is weakly monotonic. Denote $\bigoplus(A) := \bigoplus_{\omega_i \in A} q_i$.

(i) First, $\mathrm{dom}(\nu) = [0,1] \subseteq \mathbb{R} = \mathrm{dom}(\oplus)$. Recall from Definition (1) that $\forall A \subseteq \Omega, \nu(A) \in [0,1]$. Fix $A \subseteq \Omega$. First let $|A| = 1$, then $A = \{\omega\}$ for some $\omega \in \Omega$, so $\nu(A) = \nu(\{\omega\}) = q(\omega) \in [0,1]$. Then let $2 \le |A| \le n$ and let $B = A - \{\omega\}$ for some $\omega \in A$. Then $\nu(A) = \oplus(A) = [\oplus(B)] \oplus q(\omega) = \nu(B) \oplus q(\omega)$. Since $\nu(A), \nu(B), q(\omega) \in [0,1]$, therefore the range $\mathrm{ran}(\oplus) \subseteq [0,1]$. Proof by induction.

(ii) Assume sets $A, B, C, D \subseteq \Omega$, with $A \subseteq C, B \subseteq D$. Then

$$\bigoplus(A) = \nu(A) \le \nu(C) = \bigoplus(C), \qquad \bigoplus(B) = \nu(B) \le \nu(D) = \bigoplus(D).$$

Now from the associativity and commutivity of $\oplus$, and since $A \cup B \subseteq C \cup D$, therefore

$$\begin{aligned} \left[\bigoplus(A)\right] \oplus \left[\bigoplus(B)\right] &= \bigoplus(A \cup B) = \nu(A \cup B) \\ &\le \nu(C \cup D) = \bigoplus(C \cup D) = \left[\bigoplus(C)\right] \oplus \left[\bigoplus(D)\right], \end{aligned}$$

and so $\oplus$ is monotonic.

■

Definition 9 (t-Conorm Decomposability) [5, 34] Given a fuzzy measure ν and a t-conorm $\sqcup$, if

$$\forall A, B \subseteq \Omega, \qquad A \not\perp B \rightarrow \nu(A \cup B) = \nu(A) \sqcup \nu(B),$$

where $A \perp B := A \cap B = \emptyset$, then ν is $\sqcup$-decomposable.

Note 10 Definition (9) is equivalent to

$$\forall A, B \subseteq \Omega, \qquad \nu(A \cup B) \sqcup \nu(A \cap B) = \nu(A) \sqcup \nu(B), \tag{11}$$

whether $A \perp B$ or not [5].

Corollary 12 A fuzzy measure ν is $\sqcup$-decomposable iff it is $\sqcup$-distributional.

Proof: Case 1: If ν is $\sqcup$-decomposable, then $\forall A \subseteq \Omega$,

$$\nu(A) = \nu\left(\bigcup_{\omega_i \in A} \{\omega_i\}\right) = \bigsqcup_{\omega_i \in A} \nu(\{\omega_i\}) = \bigsqcup_{\omega_i \in A} q_i.$$

Case 2: Let ν be $\sqcup$-distributional and $A, B \subseteq \Omega, A \not\perp B$. Then from associativity and commutativity of $\sqcup$ and the disjointness of A and B,

$$\nu(A \cup B) = \bigsqcup_{\omega_i \in A \cup B} q_i = \left(\bigsqcup_{\omega_i \in A} q_i\right) \sqcup \left(\bigsqcup_{\omega_i \in B} q_i\right) = \nu(A) \sqcup \nu(B).$$

■

Example 13 A classical probability measure Pr is distributional for $\oplus = +$. Then $\vec{p} = \langle p_i \rangle := \vec{q}^{\text{Pr}}$ is its probability distribution, and (11), (5), and (6) assume their traditional additive forms for $A, B \subseteq \Omega$

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cup B), \qquad \Pr(A) = \sum_{\omega_i \in A} p_i, \qquad \sum_{i=1}^{n} p_i = 1.$$

Note that $+$ is not a t-conorm, since $+\colon [0,1]^2 \mapsto [0,2]$. But classical probability also assumes that Pr is normal with $\Pr(\Omega) = 1$. This restricts $+$ to be the bounded sum operator denoted $+_b$ (where $\forall a, b \in [0,1], a +_b b := (a + b) \wedge 1$ and $\wedge$ is the minimum operator), and $+_b$ *is* a conorm. If Pr is supernormal ($\Pr(\Omega) > 1$) then it does not even satisfy Definition (1) for a fuzzy measure. On the other hand, if Pr is subnormal ($\Pr(\Omega) < 1$) then it is a fuzzy measure, but not distributional (see discussion by Klement and Weber [22]).

Example 14 A possibility measure [39] Π is $\vee$-distributional, where $\vee$ is the maximum operator. Then $\vec{\pi} = \langle \pi_i \rangle := \vec{q}^{\Pi}$ is called a possibility distribution, and (11), (5), and (6) take the form

$$\Pi(A \cup B) = \Pr(A) \vee \Pr(B), \tag{15}$$

$$\Pi(A) = \bigvee_{\omega_i \in A} \pi_i, \qquad \bigvee_{i=1}^{n} \pi_i = 1.$$

for $A, B \subseteq \Omega$.

While supernormality ($\Pi(\Omega) > 1$) also precludes Π from being a fuzzy measure, subnormality ($\Pi(\Omega) < 1$) still results in Π being $\vee$-distributional. Also, note that (15) holds whether $A \perp B$ or not, since from Note (10) and the monotonicity of all fuzzy measures we have $\Pi(A \cup B) \vee \Pi(A \cap B) = \Pi(A \cup B)$.

Example 16 A Sugeno measure ν_λ defined by

$$\nu_\lambda(A \cup B) := \nu_\lambda(A) + \nu_\lambda(B) + \lambda \nu_\lambda(A) \nu_\lambda(B), \qquad A \not\perp B, \quad \lambda \in (-1, \infty)$$

is $\oplus_\lambda$-distributional, where $\oplus_\lambda$ is the Sugeno operator

$$a \oplus_\lambda b := a + b + \lambda ab, \qquad a, b \in [0,1], \quad \lambda \in (-1, \infty).$$

By similar reasoning about normalization as with probability, $\oplus_\lambda$ is restricted to be the Sugeno conorm $\sqcup_\lambda$ where $a \sqcup_\lambda b := (a + b + \lambda ab) \wedge 1$ [5].

Example 17 $\sqcup_{-1}$ is the probabilistic sum conorm, where $a \sqcup_{-1} b := a + b - ab$ for $a, b \in [0,1]$. A $\sqcup_{-1}$-distributional fuzzy measure is generally subnormal, unless $\forall A \subseteq \Omega, \nu(A) = 1$.

2.2. *Finite random sets and evidence theory*

Random set theory [21] (and the mathematically isomorphic Dempster-Shafer evidence theory [11]), while somewhat less general than the theory of fuzzy measures, nevertheless provides a broadly satisfying, unifying context within which to develop GIT. They encompass classical information theory and the most important classes of fuzzy measures, and generate fuzzy sets [10].

Random sets were originally developed as a branch of stochastic geometry [21], and involve advanced measure-theoretical concepts of trapping functions and Choquet capacities on Hausdorf spaces. But they can also be seen more simply as random variables taking values on subsets of Ω.

Definition 18 (General Random Set) [1, 9, 30] Given a probability space $\langle X, \Sigma, \Pr\rangle$, then a function $S: X \mapsto 2^\Omega - \{\emptyset\}$, where $-$ is set subtraction, is a random subset of Ω if S is Pr-measurable, so that $\forall A \subseteq \Omega, S^{-1}(A) \in \Sigma$.

Thus a general random set S associates a probability $(\Pr \circ S^{-1})(A)$ to each $A \subseteq \Omega$. When Ω is finite, then following Dubois and Prade [7], a far simpler working definition of random sets is available.

Definition 19 (Evidence Function, Basic Assignment) The function $m: 2^\Omega \mapsto [0,1]$ is an evidence function (basic assignment) when $m(\emptyset) = 0$ and $\sum_{A \subseteq \Omega} m(A) = 1$.

Definition 20 (Finite Random Set, Focal Set) Given a finite universe Ω and evidence function $m: 2^\Omega \mapsto [0,1]$, then let $\mathcal{S} := \{\langle A_j, m_j\rangle : m_j > 0\}$ be a finite random set generated by m, where $\langle\cdot\rangle$ is a vector, $A_j \subseteq \Omega, m_j := m(A_j)$, and $1 \leq j \leq N := |\mathcal{S}| \leq 2^n - 1$. Let $\vec{m} := \langle m_j \rangle$ be an N-vector. Denote the focal set as $\mathcal{F}(\mathcal{S}) := \{A_j : m_j > 0\}$, where each A_j is a focal element. Define the support of $\mathcal{S}$ as $\mathbf{U}(\mathcal{S}) := \bigcup_{A_j \in \mathcal{F}(\mathcal{S})} A_j$.

Proposition 21 [29] Each general random set S on a finite space Ω maps to a unique finite random set $\mathcal{S}$ where $m := \Pr \circ S^{-1}$; conversely, each finite random set $\mathcal{S}$ on Ω maps to a unique general random set S on Ω with measure space $\langle \mathcal{F}(\mathcal{S}), 2^{\mathcal{F}(\mathcal{S})}, \Pr\rangle$ where Pr is obtained by taking m as its distribution function.

In the sequel let "random set" denote "finite random set".

As noted, random sets are isomorphic to Dempster-Shafer bodies of evidence.

Definition 22 (Evidence Measures) [11] Given a random set $\mathcal{S}$, denote two evidence measures called plausibility and belief

$$\mathrm{Pl}(A) := \sum_{A_j \not\subseteq A} m_j, \quad \mathrm{Bel}(A) := \sum_{A_j \subseteq A} m_j, \qquad A \subseteq \Omega \tag{23}$$

respectively.

Pl and Bel are normal and generally non-additive fuzzy measures, and are dual, in that

$$\forall A \subseteq \Omega, \qquad \mathrm{Bel}(A) = 1 - \mathrm{Pl}(\overline{A}), \quad \mathrm{Pl}(A) = 1 - \mathrm{Bel}(\overline{A}). \tag{24}$$

In general only plausibility will be considered below. Bel (and dually Pl) also determine m according to the Möbius transformation

$$m(A) = \sum_{B \subseteq A} (-1)^{|B-A|} \mathrm{Bel}(B), \qquad A \subseteq \Omega. \tag{25}$$

Definition 26 (Plausibility Assignment) The plausibility assignment of a random set $\mathcal{S}$ is $\vec{\text{Pl}} = \langle \text{Pl}_i \rangle := \vec{q}^{\,\text{Pl}}$, where $\text{Pl}_i := q_i^{\text{Pl}} = \text{Pl}(\{\omega_i\})$. Also denote $\text{Pl}(\omega_i) := \text{Pl}_i$.

Note 27 $\vec{\text{Pl}}$ is also known as the one-point coverage function of $\mathcal{S}$ [10]. Also, in general from (23) we have

$$\text{Pl}_i := \sum_{A_j \ni \omega_i} m_j. \tag{28}$$

A linear mapping from the evidence values m_j to the one-point plausibilities Pl_i is available.

Definition 29 (Numerical Aggregator) For a random set $\mathcal{S}$, define the numerical aggregator as the $n \times N$ matrix $\mathbf{H} := [h_{ij}]$, where $h_{ij} := \chi_{A_j}(\omega_i)$ and χ_{A_j} is the characteristic function of A_j, so that

$$\chi_{A_j}(\omega_i) := \begin{cases} 1, & \omega_i \in A_j \\ 0, & \omega_i \notin A_j \end{cases}, \qquad A_j \in \mathcal{F}(\mathcal{S}), \quad \omega_i \in \Omega.$$

The j'th column of $\mathbf{H}$ is the vector representation of the characteristic function of A_j, so that $\mathbf{H}$ is a boolean matrix with no column repeated. Letting $\mathbf{H}(i)$ be the i'th row of $\mathbf{H}$, an N-vector, then it is obvious that

$$\text{Pl}_i = \mathbf{H}(i) \cdot \vec{m}^T, \qquad \vec{\text{Pl}}^T = \mathbf{H}\vec{m}^T. \tag{30}$$

Effectively, the left and right sides of (30) are the vector and matrix representations of (28) and (23) respectively.

The concept of the distributionality of fuzzy measures can be extended to the random sets which generate them.

Definition 31 (Random Set Distribution) For a random set $\mathcal{S}$, if Pl is $\sqcup$-distributional, then $\mathcal{S}$ is called $\sqcup$-distributional and $\vec{\text{Pl}}$ is its distribution.

Note 32 Clearly if $\mathcal{S}$ is $\sqcup$-distributional, then since Pl is normal, $\vec{\text{Pl}}$ is also normal, so that from (6) $\bigsqcup_{\omega_i \in \Omega} \text{Pl}_i = 1$ and from (5)

$$\text{Pl}(A) = \bigsqcup_{\omega_i \in A} \text{Pl}_i, \qquad A \subseteq \Omega. \tag{33}$$

3. Aggregable Random Sets

The existence of a distribution operator $\sqcup$ for a random set $\mathcal{S}$ produces functional relations between the numerical values of the measures m, Pl, and Bel defined on subsets $A \subseteq \Omega$, and the distribution values Pl_i defined on points $\omega_i \in \Omega$. This is the value of distributions, to reduce the complexity of general measures on the subsets to distribution values on the points.

A parallel reduction in complexity can also be identified not in terms of the *numerical* values of these measures, but in terms of the *structural* relation between the focal elements $A_j \subseteq \Omega$ of a random set and the elements $\omega_i \in \Omega$ (or singleton subsets $\{\omega_i\}$). When it is possible to identify a unique universe element ω_i for each focal element A_j, then we call the random set aggregable. Aggregable random sets can have very simple topological structures, and it may be possible to construct the focal sets A_j from the elements ω_i.

Definition 34 (Aggregation Functions) [16] Given a random set $\mathcal{S}$, then if a one to one function $g: \mathcal{F}(\mathcal{S}) \mapsto \Omega$ exists, then g is called a structural aggregation function and $\mathcal{S}$ is called g-aggregable. If $\mathcal{S}$ is g-aggregable, then denote the numerical aggregation function $h: \mathcal{S} \mapsto [0,1]$ with $h(m_j) = \mathrm{Pl}(g(A_j))$.

g maps each focal element A_j to a universe element $g(A_j)$, and h in turn maps that to its plausibility assignment value $h(m_j)$. Note that $h(m_j) = \mathbf{H}(i) \cdot \vec{m}^T$ for that i for which $\omega_i = g(A_j)$.

Also, in general, a random set $\mathcal{S}$ may have multiple g corresponding to the various permutations of the A_j and ω_i. Where possible, we will identify a certain canonical ordering of the A_j and ω_i with respect to each other, and thus a canonical structural aggregation function for each class of random set.

Corollary 35 A random set $\mathcal{S}$ is g-aggregable iff $|\mathcal{S}| = N \leq |\Omega| = n$.

Proof: Case 1: If $\mathcal{S}$ is g-aggregable, then the result follows directly from the definition, since $g: \mathcal{F}(\mathcal{S}) \mapsto \Omega$ and g is one to one.
Case 2: Consider the function

$$g(A_1) := \omega_1, \quad g(A_2) := \omega_2, \quad \cdots, \quad g(A_N) := \omega_N.$$

Clearly g exists only if $N \leq n$, but is then one to one, and thus a structural aggregation function. ■

$1 \leq i \leq n$ has been used to index Ω, while $1 \leq j \leq N \leq 2^n$ has been used to index $\mathcal{F}(\mathcal{S}) \subseteq 2^\Omega$. But when $\mathcal{S}$ is g-aggregable, the points can be coded directly in terms of the focal elements by combining focal element and universe element notation.

Definition 36 (Relabeling) Given a g-aggregable random set $\mathcal{S}$, denote

$$\omega_j := g(A_j), \qquad \mathrm{Pl}_j := \mathrm{Pl}(\{\omega_j\}) = \mathrm{Pl}(\{g(A_j)\}). \tag{37}$$

This relabeling convention establishes a common ordering of the A_j, m_j, ω_i, and Pl_i.

Corollary 35 establishes the general relation between the sizes of random sets and their distributions. But there is also a limiting case, where the sizes of the random set and the distribution become identical.

Definition 38 (Completion) If a random set $\mathcal{S}$ is g-aggregable and $N = |\mathcal{S}| = n = |\Omega|$, then $\mathcal{S}$ is called g-complete. If a g-complete random set $\mathcal{S}$ is also $\sqcup$-distributional, then the distribution $\vec{\mathrm{Pl}}$ is called complete.

In a g-complete random set, the focal elements and universe elements are mutually determining, with each focal element A_j existing as a particular $g^{-1}(\omega_j)$. Therefore $\mathbf{H}$ becomes a square $n \times n = N \times N$ matrix, and by relabeling the indices i and j are identical and can be used interchangeably. Now there is a unique row $\mathbf{H}(j)$, so that $\mathrm{Pl}_j = h(m_j) = \mathbf{H}(j) \cdot \vec{m}^T$ from (30).

Also then g is onto, with inverse $g^{-1}(\omega_j) = A_j$. If $\mathbf{H}$ is invertible, then from (30) $\mathbf{H}^{-1}$ becomes a linear map of the Pl_j back into the m_j, with

$$m_j = \mathbf{H}^{-1}(j) \cdot \vec{\mathrm{Pl}}^T, \qquad \vec{m}^T = \mathbf{H}^{-1}\vec{\mathrm{Pl}}^T.$$

Furthermore, then h^{-1} also exists, so that $m_j = h^{-1}(\mathrm{Pl}_j)$.

Corollary 39 For a g-complete random set $\mathcal{S}$, if $\mathbf{H}^{-1}$ exists then $\mathbf{U}(\mathcal{S}) = \Omega$.

Proof: Assume a g-complete random set $\mathcal{S}$ with $\mathbf{U}(\mathcal{S}) \subset \Omega$ so that $\Omega - \mathbf{U}(\mathcal{S}) \neq \emptyset$. Then $\forall \omega_i \in \Omega - \mathbf{U}(\mathcal{S}), \forall A_j \in \mathcal{F}(\mathcal{S}), \omega_i \notin A_j$, so that $\mathbf{H}(i) = \langle 0, 0, \ldots, 0 \rangle$ is the n-long zero vector, and $\mathbf{H}$ is singular. Proof by contraposition. ■

The relations among the elements of a g-complete, $\sqcup$-distributional random set are diagrammed in Fig. 1, where some of the labels refer to the appropriate equations in the text.

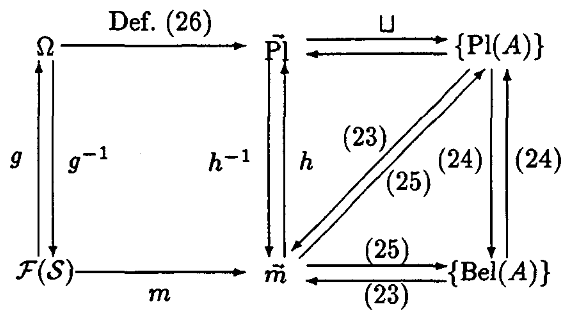

Fig. 1. **Relations in a g-complete, $\sqcup$-distributional random set.**

4. Existing Information Theories

The framework outlined above establishes a general model from within which to consider distributional fuzzy measures defined on random sets. We will now consider two specific cases of existing information theories, casting some known results in terms of the aggregation and completion model, and then introducing some new results.

For the sake of notational convenience, we may sometimes denote the element ω_i and the singleton subset $\{\omega_i\}$ interchangeably. The relations presented here are summarized in Table 1.

4.1. *Probability*

A random set $\mathcal{S}$ is specific if

$$\forall A_j \in \mathcal{F}(\mathcal{S}), \quad |A_j| = 1. \tag{40}$$

Then clearly $\forall A_j \in \mathcal{F}(\mathcal{S}), \exists! \omega_i \in \Omega, A_j = \{\omega_i\}$. Without loss of generality, let the A_j and ω_i be similarly ordered. Then $g_p(A_j) := \omega_i$ such that $A_j = \{\omega_i\}$ is a canonical structural aggregation function, since clearly g_p is one to one. It is then well known [28] that $\Pr(A) := \mathrm{Pl}(A) = \mathrm{Bel}(A)$ is a probability measure with probability distribution $\vec{p} := \vec{\mathrm{Pl}}$.

The specific case is, of course, quite simple, but for completeness we present the following.

Corollary 41 If $\mathcal{S}$ is specific and g_p-complete, then:

(i) The aggregator denoted $\mathbf{H}_p$ is the identity $n \times n$ matrix denoted $\mathbf{I}(n)$, and the numerical aggregation function is $h_p(m_j) = m_j = p_j$;
(ii) $\mathbf{H}_p^{-1} = \mathbf{H}_p = \mathbf{I}(n)$, and $h_p^{-1}(p_j) = p_j = m_j$;
(iii) $\forall \omega_i \in \Omega, p_i > 0$.

Proof:

(i) Follows immediately from (40) and the common ordering forced by the relabeling of (37).
(ii) Follows just as quickly from the inversion of the identity matrix.
(iii) Since $\mathbf{I}(n) \cdot \vec{m}^T = \vec{m}^T = \vec{p}^T$, and $\forall m_j > 0$, therefore $\forall p_j > 0$.

■

Example 42 Consider the example shown in Fig. 2. We have

$$\mathbf{H}_p \vec{m}^T = \mathbf{H}_p^{-1} \vec{p}^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} .2 \\ .3 \\ .5 \end{pmatrix} = \begin{pmatrix} .2 \\ .3 \\ .5 \end{pmatrix} = \vec{p}^T = \vec{m}^T.$$

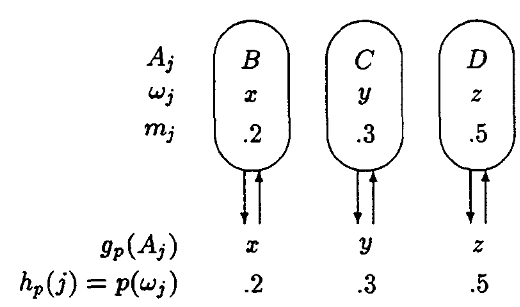

Fig. 2. A complete specific random set.

4.2. *Possibility*

Disjointness in the probabilistic case is one special case of the topological relation between two focal elements. The other extreme is when one is included in the other.

$\mathcal{S}$ is consonant ($\mathcal{F}(\mathcal{S})$ is a nest) when $\forall A_j, A_{j'} \in \mathcal{F}$, either $A_j \subseteq A_{j'}$ or $A_{j'} \subseteq A_j$. Now $\Pi(A) := \text{Pl}(A)$ is a possibility measure, $\eta(A) := \text{Bel}(A)$ is a necessity measure, and $\vec{\pi} := \vec{\text{Pl}}$ is a possibility distribution. Since results for necessity are dual to those of possibility, only possibility will be discussed in the sequel.

Note 43 Without loss of generality for ordering, let $A_0 := \emptyset$ and $A_{j-1} \subseteq A_j$. Further, in the sequel, for any given possibility distribution $\vec{\pi}$ let the π_i be ordered so that

$$\pi_1 \geq \pi_2 \geq \cdots \geq \pi_n. \tag{44}$$

Corollary 45 (Possibilistic Structural Aggregation) If $\mathcal{S}$ is consonant, then any function $g_\pi : \mathcal{F}(\mathcal{S}) \mapsto \Omega$ where $\forall A_j, g_\pi(A_j) \in A_j - A_{j-1}$ is a structural aggregation function.

Proof: In general, $A_{j-1} \subset A_j$, where $1 \leq j \leq N$, and $A_0 := \emptyset$ by convention, so that $A_j = A_{j-1} \cup (A_j - A_{j-1})$. Since the A_j are all distinct, therefore all the $A_j - A_{j-1}$ are also distinct and non-empty. Therefore any function satisfying the premise must be one to one, and thus a structural aggregation function. ■

Let a function g_π constructed according to the above ordering be a canonical possibilistic structural aggregation function. Again, it is necessary to cast some known possibility theoretical results [26] in the context of the aggregation and completion model.

Theorem 46 (Possibilistic Formulae) If $\mathcal{S}$ is consonant and g_π-complete, then (keeping the ordering of (44) and the relabeling convention of (37) in mind):

(i) $g_\pi(A_j) = A_j - A_{j-1} = \{\omega_j\}$.
(ii) $g_\pi^{-1}(\omega_j) = A_j = \{\omega_1, \omega_2, \ldots, \omega_j\}$.
(iii) The aggregator and aggregation function are

$$\mathbf{H}_\pi = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}, \qquad h_\pi(m_j) = \sum_{k=j}^{N} m_k = \pi_j.$$

(iv) The inverse aggregator and aggregation functions exist, and are

$$\mathbf{H}_\pi^{-1} = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 1 & \cdots & 1 & -1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}, \qquad h_\pi^{-1}(\pi_j) = \pi_j - \pi_{j+1} = m_j,$$

where $\pi_{n+1} = 0$ by convention.

Proof:

(i) When $\mathcal{S}$ is complete, then $|\mathcal{F}(\mathcal{S})| = n$. Since all the A_j are distinct, therefore $|A_j - A_{j-1}| = 1$. Since $g_\pi(A_j) \in A_j - A_{j-1}$, therefore $g_\pi(A_j) = \omega_i$ such that $A_j - A_{j-1} = \{\omega_i\}$ unambiguously. By relabeling, $\omega_j := g_\pi(A_j)$.

(ii) Because of the ordering convention of (44), the ω_j are also ordered so that

$$\pi(\omega_1) \geq \pi(\omega_2) \geq \cdots \geq \pi(\omega_n).$$

Furthermore,

$$\omega_1 = g_\pi(A_1) \in A_1 - A_0 = A_1 - \emptyset = A_1,$$

so that $A_1 = \{\omega_1\}$. Similarly,

$$\omega_2 = g_\pi(A_2) \in A_2 - A_1 = A_2 - \{\omega_1\},$$

so that $A_2 = \{\omega_1, \omega_2\}$. Continuing in this way, in general $A_j = \{\omega_1, \omega_2, \ldots, \omega_j\}$.

(iii) $\mathbf{H}_\pi$ and h_π are easily constructed.

(iv) It is easy to verify that $\mathbf{H}_\pi \mathbf{H}_\pi^{-1} = \mathbf{I}$, and then h_π^{-1} is easily constructed from $\mathbf{H}_\pi^{-1}$.

■

In contrast with probabilistic completion of Corollary (41), if π is complete then all the π_i are distinct.

Theorem 47 (Possibilistic Completion) A consonant random set $\mathcal{S}$ is g_π-complete iff

$$1 = \pi_1 > \pi_2 > \cdots > \pi_n > 0.$$

Proof:

(i) Assume a g_π-complete random set $\mathcal{S}$.

- $\pi_1 = 1$ from possibilistic normalization.
- If $\exists j, \pi_j = \pi_{j+1}$ then from the possibilistic formulae of Theorem (46), $\pi_j - \pi_{j+1} = m_j = 0$, which violates Definition (20).
- Finally, if $\exists \pi_i = 0$, then $\forall A_j, \omega_i \notin A_j$, so that $\Omega \notin \mathcal{F}(\mathcal{S})$. But $\Omega \in \mathcal{F}(\mathcal{S})$, because $\mathcal{S}$ is complete and consonant, and otherwise $|\mathcal{S}| < n$. Therefore $\forall \pi_i > 0$.

(ii) Assume a possibility distribution where $1 = \pi_1 > \pi_2 > \cdots > \pi_n > 0$. Let $i = 1$. Then

$$\pi_1 = \sum_{A_j \ni \omega_1} m_j = \sum_{A_j \in \mathcal{F}(\mathcal{S})} m_j > \pi_2 = \sum_{A_j \ni \omega_2} m_j,$$

so that

$$\exists A_{j_1}, A_{j_2}, \quad \omega_1 \in A_{j_1}, A_{j_2}, \quad \omega_2 \in A_{j_1}, \quad \omega_2 \notin A_{j_2}, \qquad m_{j_2} > 0.$$

The same argument holds for general i, therefore $\forall 1 \leq i \leq n, \exists A_j, m_j > 0$, so that $N = n$ and $\mathcal{S}$ is complete.

■

Example 48 Examples of both incomplete and complete consonant random sets are shown in Fig. 3 for $\Omega = \{x, y, z\}$. Note that on the left $g_\pi(C)$ is undetermined and $\pi(y) = \pi(z)$ because $\mathcal{S}$ is incomplete. On the left and right respectively we have for $\mathbf{H}_\pi \vec{m}^T = \vec{\pi}^T$

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} .6 \\ .4 \end{pmatrix} = \begin{pmatrix} 1 \\ .4 \\ .4 \end{pmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} .2 \\ .3 \\ .5 \end{pmatrix} = \begin{pmatrix} 1 \\ .8 \\ .5 \end{pmatrix}.$$

In the inverse case on the right we have

$$\mathbf{H}_\pi^{-1} \vec{\pi}^T = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 1 \\ .8 \\ .5 \end{pmatrix} = \begin{pmatrix} .2 \\ .3 \\ .5 \end{pmatrix} = \vec{m}^T.$$

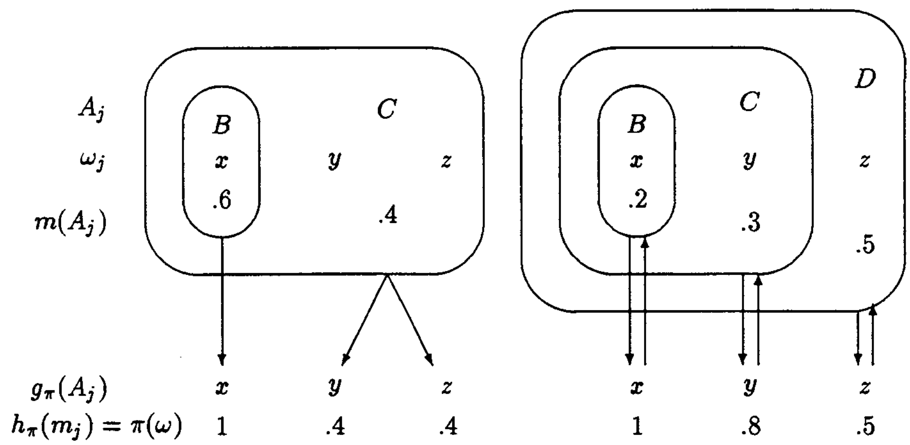

Fig. 3. An incomplete (left) and complete (right) consonant random set.

5. Hypothetical Information Theories

We now search for new special cases which might generate other information theories. In particular, we wish to investigate new distribution operators which might yield aggregable random sets, and new structural aggregation functions which might yield distributional random sets. We consider an obvious example of each, and conclude that their success is somewhat dependent on the cardinality of the universe of discourse, and that in general neither is succesful. This is in strong contrast to both probability and possibility, where all of the relavent mappings exist in all cases.

These results are summarized in Table 2.

Table 1. Summary of the existing information theories.

	General	Probabilistic	Possibilistic
Topology	$2^\Omega - \{\emptyset\}$	Specific	Consonant
Distribution	$q_j = \mathrm{Pl}_j$	$p_j = \Pr(\{\omega_j\})$	$\pi_j = \Pi(\{\omega_j\})$
t-conorm	$\sqcup$	$+$	$\vee$
Focal Element	$A_j = g^{-1}(\omega_j)$	$\{\omega_j\}$	$\{\omega_1, \omega_2, \ldots, \omega_j\}$
Structural Aggregation	$g(A_j) = \omega_j$	A_j	$A_j - A_{j-1}$
Numerical Aggregation	$h(m_j) = \mathrm{Pl}_j$	m_j	$\sum_{k=j}^n m_j$
Inverse	$h^{-1}(\mathrm{Pl}_j) = m_j$	p_j	$\pi_j - \pi_{j+1}$
Completion	$\lvert\mathcal{F}(\mathcal{S})\rvert = \lvert\Omega\rvert$	$p_j > 0$	$\pi_j > \pi_{j+1}$

5.1. *Sugeno Distributional Fuzzy Measures*

First consider the Sugeno conorm $\sqcup_\lambda$ for $\lambda \in (-1, \infty)$, and $\sqcup_\lambda$-distributional random sets. Do they have any special structure? In particular, do they have some canonical structural aggregation function g_λ? We show that they can only for a few universes of discourse, and then their aggregators are not invertible.

Recall the following [5]:

- A Sugeno measure ν_λ is any fuzzy measure which is $\sqcup_\lambda$-decomposable, and thus $\sqcup_\lambda$-distributional from Corollary (12).
- Given $\lambda \in (-1, \infty)$, let $\lambda' := \frac{-\lambda}{1+\lambda} \in (-1, \infty)$. If $\lambda \in [0, \infty)$ then ν_λ is a belief measure, $\lambda' \in (-1, 0]$, and $\nu_{\lambda'}$ is the dual Sugeno plausibility measure of ν_λ; if $\lambda \in (-1, 0]$ then ν_λ is a plausbility measure, $\lambda' \in [0, \infty)$, and $\nu_{\lambda'}$ is the dual Sugeno belief measure of ν_λ; and finally if $\lambda = \lambda' = 0$ then $\nu_\lambda = \nu_{\lambda'}$ is a probability measure.
- For a Sugeno belief measure ν_λ with $\lambda \in (0, \infty)$, the evidence function values are given by
$$m(A) = \lambda^{|A|-1} \prod_{\omega_i \in A} q_i^{\nu_\lambda}, \qquad A \subseteq \Omega. \tag{49}$$

Considering 2^Ω as an n-hypercube, the focal sets of $\sqcup_\lambda$-distributional random sets always fill an $n-k$-sub-hypercube (less $\emptyset$) for some $k \in \{0, 1, \ldots, n-1\}$.

Lemma 50 Let a random set $\mathcal{S}$ be $\sqcup_\lambda$-distributional with $\lambda \neq 0$. Then there exists an $\emptyset \neq \Omega' \subseteq \Omega$ such that $\mathcal{F}(\mathcal{S}) = 2^{\Omega'} - \emptyset$.

Proof: Let $\mathcal{S}$ be $\sqcup_\lambda$-distributional. Then $\mathrm{Pl} = \nu_\lambda$ for some $\lambda \in (-1, 0]$. It is easier to work with with the dual belief measure $\mathrm{Bel} = \nu_{\lambda'}$ where now $\lambda' \in [0, \infty)$. Denote $m^i := m(\{\omega_i\})$. Because $\forall \omega_i \in \Omega, q_i^{\mathrm{Bel}} = \mathrm{Bel}(\{\omega_i\}) = m^i$, therefore (49) becomes

$$\forall A_j \in \mathcal{F}(\mathcal{S}), \qquad m_j = \lambda^{|A_j|-1} \prod_{\omega_i \in A_j} m^i > 0.$$

Therefore $\forall A_j \in \mathcal{F}(\mathcal{S}), \forall \omega_i \in A_j, m^i > 0$, so that $\forall A_j \in \mathcal{F}(\mathcal{S}), \forall \emptyset \neq B \subseteq A_j, m(B) > 0$. Consider $A_j, A_{j'} \in \mathcal{F}(\mathcal{S})$. Then $\forall \omega_i \in A_j \cup A_{j'}, m^i > 0$, so

$m(A_j \cup A_{j'}) > 0$. So $\mathcal{F}(\mathcal{S})$ is closed under unions, and thus there exists a maximal $A_{j^*} \in \mathcal{F}(\mathcal{S})$ with $\forall \emptyset \neq A \subseteq A_{j^*}, A \in \mathcal{F}(\mathcal{S})$. A_{j^*} is just Ω'. ∎

Theorem 51 Assume a $\sqcup_\lambda$-distributional random set $\mathcal{S}$ with $\lambda \neq 0$, and let $K := \lfloor \log_2(n+1) \rfloor$ where $\lfloor x \rfloor \in \mathcal{I}$ is the greatest integer less than or equal to $x \in \mathbb{R}$. Then $\mathcal{S}$ is g_λ-aggregable for some structural aggregation function g_λ iff

$$N \quad \in \quad I := \{1, 3, 7, \ldots, 2^{K-1} - 1, 2^K - 1\}.$$

Proof: Case 1: $K = \lfloor \log_2(n+1) \rfloor \leq \log_2(n+1)$, so

$$N \in I \quad \rightarrow \quad N \leq 2^K - 1 \leq 2^{\log_2(n+1)} - 1 = n,$$

and from Corollary (35) a structural aggregation function g_λ exists.
Case 2: From Lemma (50) $N = |\mathcal{F}(\mathcal{S})| = |2^{\Omega'} - \emptyset| = 2^k - 1$ where $k = |\Omega'| \in \{1, 2, \ldots, n\}$, so that $N \in \{1, 3, 7, \ldots, 2^n - 1\}$. If a structural aggregation function g_λ exists then from Corollary (35) $N = 2^k - 1 \leq n$, so that $k \leq \log_2(n+1)$. Since k must be an integer, therefore $k \leq \lfloor \log_2(n+1) \rfloor$, and so $N \in I$. ∎

Sugeno measures are either degenerate for $\lambda = 0$, and thus already considered above as probability measures, or they produce complete random sets only under special circumstances.

Corollary 52 If a $\sqcup_\lambda$-distributional, g_λ-aggregable random set $\mathcal{S}$ is g_λ-complete, then either $\lambda = 0$ or $\exists k \in \{1, 2, \ldots\}, N = n = 2^k - 1$.

Proof: Case 1: If $\lambda = 0$, then $\sqcup_\lambda = +_b$, so that Pl is a probability measure and $\mathcal{S}$ is specific and g_p-complete from Corollary (41).
Case 2: If $\lambda \neq 0$, then from Theorem (51) and Definition (38), $n = N \in I$, and so the result follows. ∎

Furthermore, even for the restricted set of g_λ-complete random sets, their aggregators are not invertible.

Theorem 53 Given a $\sqcup_\lambda$-distributional, g_λ-complete random set $\mathcal{S}$ with $\lambda \neq 0$, then the aggregator denoted $\mathbf{H}_\lambda$ is not invertible.

Proof: Assume a $\sqcup_\lambda$-distributional, g_λ-complete random set $\mathcal{S}$ with $\lambda \neq 0$. $|\Omega| = 1$ is a degenerate case, so let $n = |\Omega| > 1$. From Corollary (52) $\exists k \in \{1, 2, \ldots\}, N = n = 2^k - 1$. So $k = \log_2(n+1) < n$ because $n > 1$. Therefore $|\Omega'| = k < n = |\Omega|$ and so $\Omega' \subset \Omega$. Since from Lemma (50) $\mathcal{F}(\mathcal{S}) = \{A_j \subseteq \Omega'\} - \{\emptyset\}$, therefore $\mathbf{U}(\mathcal{S}) = \Omega' \subset \Omega$, so that from contraposition and Corollary (39), $\mathbf{H}$ is singular. ∎

Note 54 The converse of Lemma (50) does not hold, so that there are some random sets with hypercube focal sets which are not $\sqcup_\lambda$-distributional. Also, there are some random sets with $|\mathcal{F}(\mathcal{S})| = 2^k - 1$ for $k \in \{1, 2, \ldots, n\}$ which are not hypercubes.

5.2. *Ring-structured random sets*

The purpose of a structural aggregation function g is to provide a mechanism whereby each focal element focuses attention on a specific universe element ω_i or singleton set $\{\omega_i\}$. In turn, the inverse mapping g^{-1} allows the focal elements to be constructed from the points.

Having considered $\sqcup_\lambda$ as an alternate distribution operator function, and having determined that in general it yields no structural aggregation function, let us now consider an alternate structural aggregation function and see if it yields a distribution operator function.

5.2.1. *Set intersection structural aggregation*

We have shown that the canonical structural aggregation functions for probability and possibility are

$$\begin{aligned} g_p(A_j) &= A_j = \{\omega_j\} \\ g_\pi(A_j) &= A_j - A_{j-1} = \{\omega_j\} \end{aligned}$$

respectively. The first (probabilistic) case is degenerate. In the second (possibilistic) case, singletons are recovered by removing one focal element from another by set subtraction.

As another natural way for focal elements to focus on singletons, consider the set *intersection* function, so that

$$\forall \omega \in \Omega, \quad \exists A_j, A_{j'} \in \mathcal{F}(\mathcal{S}), \quad A_j \cap A_{j'} = \{\omega_i\}.$$

When, similarly to the possibilistic case, we assume a linear order among the A_j, but now let $A_0 := A_N$ by convention, then we arrive at the canonical form of a structural aggregation function denoted

$$g_r(A_j) := A_j \cap A_{j-1} = \{\omega_j\}. \tag{55}$$

g_r specifies "ringlike" random sets where focal elements are linked by sharing exactly one point with each of two other focal elements on "either side" of it.

Corollary 56 (Ring Structural Aggregation) For a given g_r-complete random set $\mathcal{S}$

$$\mathcal{F}(\mathcal{S}) = \{\{\omega_1, \omega_2\}, \{\omega_2, \omega_3\}, \ldots, \{\omega_{n-1}, \omega_n\}, \{\omega_n, \omega_1\}\}$$

and $\forall \omega_j \in \Omega, g_r^{-1}(\omega_j) = \{\omega_j, \omega_{j+1}\}$, where $\omega_{n+1} := \omega_1$ by convention.

Proof: Since $\mathcal{S}$ is complete, therefore $N = n$, and we can use the indices j and ignore the indices i. Therefore, by (55) and direct construction we can determine that

$$\begin{array}{rcllllllll} g_r(A_2) &=& A_1 \cap A_2 & & & & &=& \omega_2 & \\ g_r(A_3) &=& & A_2 \cap A_3 & & & &=& \omega_3 & \\ & \vdots & & & \vdots & \vdots & & \vdots & & \\ g_r(A_n) &=& & & & A_{n-1} \cap A_n & &=& \omega_n & \\ g_r(A_1) &=& A_1 \cap & & & & A_n &=& \omega_{n+1} & = \omega_1 \end{array}$$

Only $\omega_2, \omega_3 \in A_2$, for example, so that $A_2 = \{\omega_2, \omega_3\}$. It follows that, in general, $A_j = g^{-1}(\omega_j) = \{\omega_j, \omega_{j+1}\}$. The first result follows immediately. ∎

Example 57 Let $\Omega = \{x, y, z, w\}$, and consider the focal set

$$\mathcal{F}(\mathcal{S}) = \{\{x,y\},\{y,z\},\{z,w\},\{w,x\}\}$$

as shown in Fig. 4.

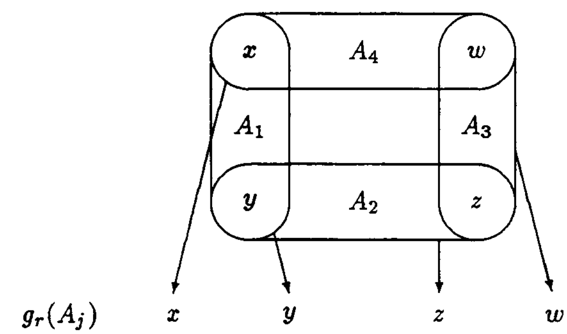

Fig. 4. A complete ring-structured focal set.

5.2.2. *Justification of ring structures*

Mathematical grounds alone are sufficient to justify consideration of g_r as a structural aggregation function: like set subtraction, all simple mechanisms by which focal elements can be constructed from singletons (and vice versa) should be investigated. It is also suggestive that Klir and Ramer have shown that in all three of the probabilistic, possibilistic, and ring-structured cases, when m is uniform and $\mathcal{S}$ complete then the total uncertainty (sum of the nonspecificity and discord) is maximized at $\log_2(n)$ [27].

But there are also important *semantic* criteria to justify consideration of ring-structured random sets, and potential applications. Consider two subsets $A, B \subseteq \Omega$ with $|A| > 0 < |B|$. There are at most three general possible topological relations among them:

$$A \cap B = \emptyset; \qquad A \subseteq B \text{ (or } B \subseteq A); \qquad A \cap B \neq \emptyset, \quad A - B \neq \emptyset, \quad B - A \neq \emptyset.$$

Considering A and B as general focal elements, the first case corresponds to probability and the second to possibility. The third case, where A intersects B "properly", corresponds to ring-structured random sets. Thus fundamentally, where the probabilistic case is degenerate, and the possibilistic case is prototypically linear and ordinal, the ring-structured case is prototypically *cyclic*.

Potential applications include:

- Inferential or evidential situations involving cyclically related evidential claims, for example, in ill-defined inferential systems where linear chains have yet to be completely identified;

- Qualitative modeling of complex dynamical systems whose attractors are regions of cyclic chain dynamics with varying sizes and degrees of stability [20];
- And knowledge-based control systems, using uncertainty modeling, which rely on cyclic feedback relations [31].

5.2.3. *Distributions in ring-structured random sets*

So the questions arise: what is the structure of the aggregator, denoted $\mathbf{H}_r$? is there a distribution operator $\sqcup$ which yields such a distribution r? and does the inverse $\mathbf{H}_r^{-1}$ exist? It will be shown that in general neither r nor $\mathbf{H}_r^{-1}$ exists.

Corollary 58 (Ring Numerical Aggregation) For a g_r-complete random set $\mathcal{S}$, $h_r(m_j) = m_j + m_{j-1}$ and

$$\mathbf{H}_r = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 \end{bmatrix}.$$

Proof: For a general ω_j, from the proof of Corollary (56), $\omega_j \in A_j$ and $\omega_j \in A_{j-1}$. If $\vec{r}$ exists, then

$$h_r(m_j) = \mathrm{Pl}_j = \sum_{A_k \ni \omega_j} m_k = m_j + m_{j-1}.$$

The matrix form follows by recalling that $\omega_{n+1} = \omega_1$. ■

Theorem 59 Let $\mathcal{S}$ be a g_r-complete random set. Then:

(i) If n is odd then

$$\mathbf{H}_r^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 & -1 & 1 & -1 & \cdots & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 & \cdots & -1 & 1 \\ 1 & -1 & 1 & 1 & -1 & \cdots & 1 & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -1 & 1 & -1 & 1 & -1 & \cdots & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & \cdots & -1 & 1 \end{bmatrix}.$$

(ii) If n is even then $\mathbf{H}_r$ is not invertible.

Proof:

(i) Let n be odd. Then $\mathbf{H}_r \mathbf{H}_r^{-1} = \mathbf{I}$ is easily constructed.

(ii) Let n be even. Take $\mathbf{H}_r$ as

$$\begin{matrix} R_1 \\ R_2 \\ R_3 \\ \vdots \\ R_{n-2} \\ R_{n-1} \\ R_n \end{matrix} \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & 1 \\ 1 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & \cdots & 1 & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 & 1 \end{pmatrix}$$

where the row numbers are indicated as $R_1, R_2, \ldots, R_n$. Note that R_{n-k} is even if k is even. Use the standard row reduction method to construct a new equivalent matrix $\mathbf{H}'_r$ in order to attempt to calculate $\mathbf{H}_r^{-1}$ [37]. Retain the first row, so that $R'_1 = R_1$. Replace the second row with the difference of the second row and the first row, so that $R'_2 = R_2 - R'_1$. Continue with the third row, so that $R'_3 = R_3 - R'_2$, through $R'_n = R_n - R'_{n-1}$. In this iterative way, $\mathbf{H}'_r$ is derived as

$$\begin{array}{l} R'_1 = R_1 \\ R'_2 = R_2 - R'_1 \\ R'_3 = R_3 - R'_2 \\ \vdots \\ R'_{n-2} = R_{n-2} - R_{n-3} \\ R'_{n-1} = R_{n-1} - R_{n-2} \\ R'_n = R_n - R'_{n-1} \end{array} \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 1 & \cdots & 0 & 0 & -1 \\ 0 & 0 & \cdots & 0 & 0 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & -1 \\ 0 & 0 & \cdots & 0 & 1 & 1 \\ 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}.$$

Since $R'_n = \langle 0, 0, \ldots, 0 \rangle$, therefore $\mathbf{H}'_r$ is singular.

■

Finally, it can be shown that in general ring-structured random sets do not have a distribution operator $\sqcup$, and thus there is generally no well-defined ring-specific random set distribution r. Thus, unlike specific and consonant random sets, ring-structured random sets cannot serve as bases to construct an alternative information theory.

Theorem 60 A g_r-complete random set $\mathcal{S}$ is not distributional for any distribution operator $\sqcup$.

Proof: Let $\mathcal{S}$ be g_r-complete, and assume $\sqcup$ exists so that $\mathcal{S}$ is $\sqcup$-distributional. Then from (33), (37), and Corollary (58),

$$\mathrm{Pl}(A) = \bigsqcup_{\omega_j \in A} \mathrm{Pl}_j = \bigsqcup_{\omega_j \in A} m_j + m_{j-1}, \qquad A \subseteq \Omega. \tag{61}$$

$n = 1, 2$ are degenerate cases with g_r not well defined. Therefore we must consider two cases.

Case 1: Let $n = 3$, and denote $\Omega = \{x, y, z\}$. Then for all g-complete random sets $\mathcal{S}$, $\mathcal{F}(\mathcal{S}) = \{A_1, A_2, A_3\} = \{\{x,y\},\{y,z\},\{z,x\}\}$. Denote $\vec{m} = \langle m(A_j) \rangle := \langle a, b, c \rangle$ so that $a, b, c \in (0,1], a + b + c = 1$. Therefore $\vec{\mathrm{Pl}} = \langle \mathrm{Pl}(x), \mathrm{Pl}(y), \mathrm{Pl}(z) \rangle = \langle a + c, b + a, c + b \rangle$. Also $\mathrm{Pl}(A_1) = 1 - \mathrm{Bel}(\{z\}) = 1 - m(\{z\}) = 1$, and similarly $\mathrm{Pl}(A_2) = \mathrm{Pl}(A_3) = 1$. Therefore from (61)

$$\begin{aligned} \mathrm{Pl}(A_1) &= (a+b) \sqcup (a+c) = \\ \mathrm{Pl}(A_2) &= (a+b) \sqcup (b+c) = \\ \mathrm{Pl}(A_3) &= (a+c) \sqcup (b+c) = 1, \end{aligned}$$

which can only be the case if $\sqcup$ is the constant function $a \sqcup b \equiv 1, a, b \in (0,1]$, which is not a conorm.

Case 2: Let $n > 3$. Consider an arbitrary subset $A \subseteq \Omega$ with $|A| = 2$, denoted without loss of generality as $A := \{\omega_2, \omega_3\}$, and denote $m_1 := a, m_2 := b, m_3 := c$. Then

$$\mathrm{Pl}(A) = \sum_{A_j \not\perp \{\omega_2, \omega_3\}} m_j = \sum_{A_j \in \{A_1, A_2, A_3\}} m_j = a + b + c.$$

For (61) to hold, we must also have $\mathrm{Pl}(A) = \mathrm{Pl}_2 \sqcup \mathrm{Pl}_3 = (b + a) \sqcup (c + b)$, so that $(a+b) \sqcup (b+c) = a+b+c$. Now consider a new random set $\mathcal{S}'$ on Ω with the same focal set $\mathcal{F}(\mathcal{S}') = \mathcal{F}(\mathcal{S})$, but with new plausibility Pl' and evidence vector

$$\vec{m}' := \langle a - \delta, b + \delta, c - \delta, m'_4, \ldots, m'_N \rangle ,$$

for some $\delta \in (0, 1]$, where the quantity δ has been added to the $m_4, \ldots, m_N$ collectively under some arbitrary division so that $\sum_{j=4}^N m'_j - \sum_{j=4}^N m_j = \delta$, but still $\sum_{j=1}^N m'_j = 1$, and δ is chosen arbitrarily small so that $\forall j \in \{1, \ldots, N\}, m'_j \in [0, 1]$. Then

$$\begin{aligned} \mathrm{Pl}'(A) &= \sum_{A_j \not\perp \{\omega_2, \omega_3\}} m'_j = \sum_{A_j \in \{A_1, A_2, A_3\}} m'_j \\ &= (a - \delta) + (b + \delta) + (c - \delta) = a + b + c + \delta \\ &> a + b + c = \mathrm{Pl}(A). \end{aligned}$$

But also from (61)

$$\begin{aligned} \mathrm{Pl}'(A) &= \mathrm{Pl}'_2 \sqcup \mathrm{Pl}'_3 \\ &= ((b + \delta) + (a - \delta)) \sqcup ((c - \delta) + (b + \delta)) = (a + b) \sqcup (b + c) = \mathrm{Pl}(A). \end{aligned}$$

This is a contradiction, therefore $\sqcup$ cannot exist as a function $\sqcup : [0,1]^2 \mapsto [0, 1]$. ■

Finally, the completeness conditions for ring-structured random sets are similar to those of specific random sets.

Corollary 62 A g_r-aggregable random set $\mathcal{S}$ is g_r-complete iff $\forall \omega_j \in \Omega, \mathrm{Pl}_j > 0$.

Proof: Assume a g_r-aggregable random set $\mathcal{S}$.
Case 1: Assume $\mathcal{S}$ is g_r-complete. Since $\forall \omega_j \in \Omega, A_j \cap A_{j-1} = \{\omega_j\}$ uniquely, therefore $\forall \omega_j \in \Omega, \exists A_j \in \mathcal{F}(\mathcal{S}), \omega_j \in A_j$, and so $\mathrm{Pl}_j > 0$.
Case 2: Assume $\forall \omega_j \in \Omega, \mathrm{Pl}_j > 0$. If $N < n$, then $\exists \omega_{j^*}, \forall A_j \in \mathcal{F}(\mathcal{S}), \omega_{j^*} \notin A_j$, which would force $\mathrm{Pl}_{j^*} = 0$, which is a contradiction, so $N = n$ and $\mathcal{S}$ is g_r-complete. ■

6. Further Questions

Of course, these results are far from satisfactory. A number of open questions remain:

- Ring-structured random sets are attractive, since the three cases of degenerate (probabilistic), linear (possibilistic), and cyclic (ring-structured) structures capture the obvious topological relations among focal elements. Is there a cyclic structural aggregation function other than g_r which might yield a distribution operator $\oplus$?

Table 2. Summary of the special random set cases.

	Sugeno	**Ring**
Topology	Sub-hypercube	Ring
t-conorm	$\sqcup_\lambda$	None
Focal Element	Any $\emptyset \neq A \subseteq \Omega$	$\{\omega_j, \omega_{j+1}\}$
Structural Aggregation	Only for $N \leq 2^{\lfloor \log_2(n+1) \rfloor} - 1$	$A_j \cap A_{j+1}$
Numerical Aggregation	None	$m_j + m_{j-1}$
Inverse	Never	For n even
Completion	$n = 2^k - 1$	$\mathrm{Pl}_j > 0$

- If not, what topological structures other than g_p, g_π, and g_r also focus focal elements on singletons?
- What are the the properties of the probabilistic sum operator $\sqcup_{-1}$-distributional fuzzy measures and random sets?
- In general, is there a yet broader framework than the aggregation and completion model where other classes of fuzzy measures with distributions might be generated?
- If not, might it be that g_p and g_π are the *only* structural aggregation functions which yield invertible, distributional random sets? This would be a very strong result indeed, establishing probability and possibility theory uniquely as the proper domains for information theoretical modeling.

Acknowledgments

I would like to thank Prof. Irwin Goodman for his general support in these investigations, and a reviewer whose challenging comments greatly improved this paper.

1. Artstein, Z and Vitale, RA: (1975) "Strong Law of Large Numbers for Random Compact Sets", *Annals of Probability*, v. **3**, pp. 879-882
2. de Cooman, Gert: (1995) "The Formal Analogy Between Possibility and Probability Theory", in: *Foundations and Applications of Possibility Theory*, ed. G de Cooman et al., pp. 71-87, World Scientific, Singapore
3. de Cooman, Gert; Ruan, D; and Kerre, EE, eds.: (1995) *Foundations and Applications of Possibility Theory*, World Scientific, Singapore
4. Dempster, AP: (1967) "Upper and Lower Probabilities Induced by a Multivalued Mapping", *Annals of Mathematical Statistics*, v. **38**, pp. 325-339
5. Dubois, Didier and Prade, Henri: (1982) "Class of Fuzzy Measures Based on Triangular Norms", *Int. J. of General Systems*, v. **8**, pp. 43-61
6. Dubois, Didier and Prade, Henri: (1988) *Possibility Theory*, Plenum Press, New York
7. Dubois, Didier and Prade, Henri: (1990) "Consonant Approximations of Belief Functions", *Int. J. Approximate Reasoning*, v. **4**, pp. 419-449
8. Dubois, Didier and Prade, Henri: (1993) "Fuzzy Sets and Probability: Misunderstandings, Bridges and Gaps", in: *Proc. FUZZ-IEEE '93*
9. Goodman, IR: (1984) "Some New Results Concerning Random Sets and Fuzzy Sets", *Information Sciences*, v. **34**, pp. 93-113

10. Goodman, IR: (1994) "New Characterization of Fuzzy Logic Operators Producing Homomorphic-Like Relations with One-Point Coverages of Random Sets", in: *Advances in Fuzzy Theory and Technology*, ed. PP Wang, pp. 133-159, Bookwrights Press, Raleigh NC
11. Guan, J and Bell, DA: (1992) *Evidence Theory and its Applications*, v. **1-2**, North Holland
12. Janssen, Hugo J; de Cooman, Gert; and Kerre, Etienne: (1996) "First Results for a Mathematical Theory of Possibilistic Processes", in: *Proc. 1996 European Meeting on Cybernetics and Systems Research*, ed. R. Trappl, pp. 341-346
13. Joslyn, Cliff: (1993) "Possibilistic Semantics and Measurement Methods in Complex Systems", in: *Proc. 2nd Int. Symposium on Uncertainty Modeling and Analysis*, ed. Bilal Ayyub, pp. 208-215, IEEE Computer Society Press
14. Joslyn, Cliff: (1994) "On Possibilistic Automata", in: *Computer Aided Systems Theory—EUROCAST '93*, ed. F. Pichler and R. Moreno-Diáz, pp. 231-242, Springer-Verlag, Berlin
15. Joslyn, Cliff: (1994) *Possibilistic Processes for Complex Systems Modeling*, SUNY Binghamton, PhD dissertation, available from UMI Disseration Services, Ann Arbor MI
16. Joslyn, Cliff: (1994) "Aggregation and Completion in Probability and Possibility", in: *Proc. 1994 Joint Conf. on Information Systems*, ed. PP Wang, pp. 333-336, Pinehurst NC
17. Joslyn, Cliff: (1995) "In Support of an Independent Possibility Theory", in: *Foundations and Applications of Possibility Theory*, ed. G de Cooman et al., pp. 152-164, World Scientific, Singapore
18. Joslyn, Cliff: (1996) "Some New Results on Possibilistic Measurement", in: *Fuzziness: An Introduction to Theory and Applications*, ed. M. Mcallister, SIAM, in press
19. Joslyn, Cliff: (1996) "Measurement of Possibilistic Histograms from Interval Data", *Int. J. General Systems*, in press
20. Kauffman, Stuart A: (1990) "Requirement for Evolvability in Complex Systems: Orderly Dynamics and Frozen Components", in: *Complexity, Entropy and the Physics of Information*, ed. WH Zurek, pp. 151-192, Addison-Wesley, Redwood City
21. Kendall, DG: (1974) "Foundations of a Theory of Random Sets", in: *Stochastic Geometry*, ed. EF Harding and DG Kendall, pp. 322-376, Wiley, New York
22. Klement, Erich P and Weber, Siegfried: (1991) "Generalized Measures", *Fuzzy Sets and Systems*, v. **40**, pp. 375-394
23. Klir, George: (1984) "Possibilistic Information Theory", in: *Cybernetics and Systems Research*, v. **2**, ed. R. Trappl
24. Klir, George: (1991) "Generalized Information Theory", *Fuzzy Sets and Systems*, v. **40**, pp. 127-142
25. Klir, George: (1993) "Developments in Uncertainty Based Information", in: *Advances in Computers*, v. **36**, ed. M. Yovitz, pp. 255-332, Academic Press
26. Klir, George and Folger, Tina: (1987) *Fuzzy Sets, Uncertainty, and Information*, Prentice Hall
27. Klir, George and Ramer, Arthur: (1991) "Uncertainty in Dempter-Shafer Theory: Critical Reexamination", *Int. J. of General Systems*, v. **18**, pp. 155-166
28. Klir, George and Yuan, Bo: (1995) *Fuzzy Sets and Fuzzy Logic*, Prentice-Hall, New York
29. Kõrner, R and Nãther, W: (1995) "On the Specificity of Evidences", *Fuzzy Sets and Systems*, v. **71**, pp. 183-196
30. Mahler, Ronald: (1996) "Representing Rules as Random Sets, I: Statistical Correlations between Rules", *Information Science*, v. **88**, pp. 47-68
31. Richardson, George: (1991) *Feedback Thought*, U Pennsylvania Press, Philadelphia
32. Shafer, Glen: (1976) *Mathematical Theory of Evidence*, Princeton U Press, Princeton

33. Wang, Zhenyuan and Klir, George J: (1992) *Fuzzy Measure Theory*, Plenum Press, New York
34. Weber, S: (1984) "$\perp$-Decomposable Meaures and Integrals for Archimedean t-Conorms $\perp$", *J. Mathematical Analysis and Applications*, v. **101**, pp. 114-138
35. Weber, S: (1991) "Uncertainty Measures, Decomposability and Admissibility", *Fuzzy Sets and Systems*, v. **40**, pp. 395-405
36. Wierman, Mark J: (1996) "Assessing Fuzzy Sets and the Potential of Possibility Theory", in: *Information Sciences*, v. **88**, pp. 247-261
37. Wylie, Ray C: (1975) *Advanced Engineering Mathematics*, McGraw-Hill, New York
38. Zadeh, Lotfi A: (1965) "Fuzzy Sets", *Information and Control*, v. **8**, pp. 338-353
39. Zadeh, Lotfi A: (1978) "Fuzzy Sets as the Basis for a Theory of Possibility", *Fuzzy Sets and Systems*, v. **1**, pp. 3-28