Joslyn, Cliff A: (1993) "Possibilistic Semantics and Measurement Methods in Complex Systems", in: Proc. 2nd Int. Symposium on Uncertainty Modeling and Analysis (ISUMA 1993), ed. Bilal Ayyub, pp. 208-215, IEEE Computer Soc., https://doi.org/10.1109/ISUMA.1993.366766

Possibilistic Semantics and Measurement Methods in Complex Systems

Cliff Joslyn*†
Systems Science
SUNY at Binghamton
Binghamton, New York

Abstract

Possibility theory is a new mathematical theory for the representation of uncertainty. It is related to, but distinct from, probability theory, Dempster-Shafer evidence theory, and fuzzy set theory. It has been applied almost exclusively to knowledge-based engineering systems, with measurements taken from subjective evaluations. Toward the end of developing a strictly possibilistic semantics of natural systems, the following will be considered: the semantics of possibility statements in relation to modal logic, natural language, and mathematical possibility theory; a strong consistency relation for probability and possibility; the basis for the application of possibility theory to complex systems; and physical measurement procedures for possibility.

1 Introduction

Possibility theory, as a new method for the representation of uncertainty, is similar to, and yet distinct from, probability theory. Together with other new mathematical theories of uncertainty (fuzzy sets, fuzzy measures, Dempster-Shafer evidence theory, and random sets), they comprise the field of **Generalized Information Theory** (GIT) [12].

Although possibility and probability are formally independent, they are related within evidence theory in that both are fuzzy measures, and within fuzzy theory in that their distributions are fuzzy sets. But while possibility theory has almost always been related directly to fuzzy set theory, probability theory has been regarded as independent from it. This confusion may have resulted from the fact that possibility theory is a very weak representation of uncertainty, whereas probability makes very strong requirements.

Furthermore, the wedding of possibility theory to fuzzy sets has relegated possibility to interpretation in strict accordance with fuzzy semantics. Since the founding of fuzzy theory by Lotfi Zadeh in the 1960's, fuzziness has been interpreted almost exclusively as a *psychological* form of uncertainty, expressed in natural language (or "linguistic variables"), and measured by the subjective evaluations of human subjects.

It is now necessary to develop a possibilistic semantics of *natural* systems that is independent of *both* fuzzy sets and probability. This will be done within the context of the mathematics of possibility theory (in relation to fuzzy sets and probability) and the meaning of possibility statements (in natural language and probability theory). A strong consistency principle of probability and possibility leads to the consideration of the conceptual basis for the interpretation of possibility in processes and in complex systems.

2 Review of possibility theory

First, the mathematical confluence of evidence theory, probability, fuzzy sets, and possibility theory is reviewed [2, 6, 14]. Given a finite universe $\Omega = \{\omega_i\}, 1 \leq i \leq n$, the set function $m\colon 2^\Omega \mapsto [0,1]$ is an **evidence function** (otherwise known as a **basic probability assignment**) when $m(\emptyset) = 0$ and $\sum_{A \subset \Omega} m(A) = 1$. The **random set** generated by an evidence function is $\mathcal{S} = \{\langle A_j, m_j \rangle : m_j > 0\}$, where $\langle \cdot \rangle$ is a vector, $A_j \subset \Omega, m_j = m(A_j)$, and $1 \leq j \leq |\mathcal{S}| \leq 2^n - 1$. Denote the **focal set** as $\mathcal{F} = \{A_j : m_j > 0\}$ with **core** $\mathrm{C}(\mathcal{F}) = \bigcap_{A_j \in \mathcal{F}} A_j$.

The dual **belief** and **plausibility** measures $\forall A \subset \Omega$ are $\mathrm{Bel}(A) = \sum_{A_j \subset A} m_j$ and $\mathrm{Pl}(A) = \sum_{A_j \cap A \neq \emptyset} m_j$, and are both **fuzzy measures**. The **plausibility assignment** (otherwise known as the **contour func-**

*Mailing address: 327 Spring St. #2, Portland, ME, 04102, USA. Email: cjoslyn@bingsuns.cc.binghamton.edu, joslyn@kong.gsfc.nasa.gov.

†Supported under NASA Grant # NGT 50757.

tion, falling shadow, or **one-point coverage function**) of $\mathcal{S}$ is

$$\vec{\mathrm{Pl}} = \langle \mathrm{Pl}(\{\omega_i\}) \rangle = \langle \mathrm{Pl}_i \rangle, \qquad \mathrm{Pl}_i = \sum_{A_j \ni \omega_i} m_j.$$

When $\forall A_j \in \mathcal{F}, |A_j| = 1$, then $\mathcal{S}$ is **specific**, and $\forall A \subset \Omega, \mathrm{Bel}(A) = \mathrm{Pl}(A) = \Pr(A)$ is a **probability measure** with **probability distribution** $\vec{\mathrm{Pl}} = \vec{p} = \langle p_i \rangle$, additive normalization $\sum_i p_i = 1$, and operator $\Pr(A) = \sum_{\omega_i \in A} p_i$.

$\mathcal{S}$ is **consonant** ($\mathcal{F}$ is a **nest**) when (without loss of generality for ordering, and letting $A_0 = \emptyset$) $A_{j-1} \subset A_j$. Now $\mathrm{Pl}(A) = \Pi(A)$ is a **possibility measure** with dual **necessity measure** $\mathrm{Bel}(A) = \eta(A)$. These are related in that $\forall A \subset \Omega$

$$\Pi(A) = 1 - \eta(\overline{A}), \qquad \eta(A) = 1 - \Pi(\overline{A}) \tag{1}$$

$$\eta(A) > 0 \rightarrow \Pi(A) = 1, \qquad \Pi(A) < 1 \rightarrow \eta(A) = 0. \tag{2}$$

As Pr is additive, so Π is **maximal** in the sense that $\forall A_1, A_2 \in \mathcal{F}, \Pi(A_1 \cup A_2) = \Pi(A_1) \vee \Pi(A_2)$, where $\vee$ is the maximum operator. Denoting $A_i = \{\omega_1, \omega_2, \ldots, \omega_i\}$, and assuming that $\mathcal{F}$ is complete ($\forall \omega_i \in \Omega, \exists A_i \in \mathcal{F}$), then $\vec{\mathrm{Pl}} = \vec{\pi} = \langle \pi_i \rangle$ is a **possibility distribution** with maximal normalization and operator

$$\bigvee_{i=1}^{n} \pi_i = \pi_1 = 1, \qquad \Pi(A) = \bigvee_{\omega_i \in A} \pi_i. \tag{3}$$

The core of $\mathcal{S}$ is then

$$\mathbf{C}(\mathcal{S}) = A_1 \neq \emptyset. \tag{4}$$

A **fuzzy subset** $\widetilde{F}$ of Ω is defined by its **membership function** $\mu_{\widetilde{F}} \colon \Omega \mapsto [0,1]$, which is a generalization to the unit interval from the set $\{0,1\}$ of the characteristic function for crisp sets. Let the cardinality of a fuzzy set be $|\widetilde{F}| = \sum_i \mu_{\widetilde{F}}(\omega_i)$. Kampè de Fériet has shown [11] that for countable Ω, and for some random set $\mathcal{S}$ and fuzzy set $\widetilde{F}$ taken as a vector, that:

- $|\widetilde{F}| \geq 1$ iff $\widetilde{F}$ can be taken as a plausibility assignment $\vec{\mathrm{Pl}}$;
- similarly, $|\widetilde{F}| \leq 1$ iff $\widetilde{F}$ can be taken as a "belief assignment" $\vec{\mathrm{Bel}} = \langle \mathrm{Bel}(\{\omega_i\}) \rangle$; and finally
- $|\widetilde{F}| = 1$ iff $\widetilde{F}$ can be taken as a probability distribution $\vec{p}$.

In the first two cases, only a mapping back to an equivalence class of random sets is guaranteed, whereas in the last case each additive fuzzy set maps to a unique specific (probabilistic) random set. Finally, if $\bigvee_i \mu_{\widetilde{F}}(\omega_i) = 1$ then the first case holds, but $\vec{\mathrm{Pl}} = \vec{\pi}$ is also a possibility distribution mapping back to an equivalence class of consonant (possibilistic) random sets [6].

A number of researchers are developing possibility theory into a complete, alternative information theory. Unique measures of information — analogs of stochastic entropy — have been developed [13], and possibilistic automata have recently been defined [9].

Probability and possibility are formally complementary and independent, rarely even ***defined*** on the same random sets. They refer to distinct forms of uncertainty, each with its own internal logic. Probability represents a highly constrained, indeed a maximally strong, form of uncertainty, whereas possibility is very weak. The specificity, additive normalization, identity of Pl and Bel, and functional mapping of distributions to measures of stochastic random sets all represent maximal constraint. In contrast, the non-specificity, maximal normalization, and maximal divergence of the fuzzy measures of possibilistic random sets all indicate the weak constraints represented by possibilistic uncertainty. The semantic consequences of this will be explored in Section 5 below.

3 Classical semantics of possibility

Historically, the concept of possibility has been based on the view that events can (and should) be classified as either strictly possible or impossible. This is the case in modal logic, natural language, and probability theory. In extending these concepts to include graded possibility, there is no need to avoid, and indeed every need to embrace, compatibility with this "crisp" case. This is exactly the spirit of the crucial **extension principle** of fuzzy theory, which states that everything that holds for crisp sets must hold for fuzzy sets in the limit of crisp membership grades.

3.1 Modal possibility

Possibility and the closely related concepts of necessity, impossibility, and contingency have a long history in philosophy. In modern philosophy these ideas are represented in modal logics [7]. For a proposition p, the possibility statement $M(p)$ and necessity statement $L(p)$ are related in all axiomatizations according to the formula

$$M(p) = \neg L(\neg p), \qquad L(p) = \neg M(\neg p), \qquad (5)$$

where $\neg$ is logical negation. It is comforting to note the equivalence of the dualities of Equation 5 and Equation 1 under the appropriate translation.

Since Aristotle, modal logic has been a crisp theory, with propositions being explicitly possible, necessary, impossible, or contingent. "Quantified" modal operators have been introduced to unite predicate and modal logic, allowing expressions such as $\forall x, L(p(x))$. But no ***multi-valued*** modal logic has been introduced to allow expressions of the form $L(p) = .5$, meaning "p is half-necessary".

Modal logic remains a highly mathematical form of philosophy, with little or no application in science, and no physical measurement or interpretation procedures. It provides a robust mathematical theory of crisp possibility, but little more.

3.2 Natural language possibility

The criteria for possibilistic semantics should also be constrained by the natural language senses of possibility. For example, we might observe a die with six faces, and so let $\Omega = \{1,2,3,4,5,6\}$. Each face is a possible outcome for which there is neither qualification nor quantification: all faces are just *possible.* We may then question whether the die is fair or not, and consider the distribution of the various faces' occurrences. But this is then embarking on a ***probabilistic*** analysis; there is never any question that each face is completely *possible*, no matter how unevenly ***likely*** they may be.

Another factor in the common language usage of crisp possibility is the relation of possibility to occurrence. One definition of "possibility" offered by Webster is "being something that may or may not occur" [21]. Näther [18] observes: "The popular meaning of possibility [is]: events which take place at least one time are possible (but not necessarily probable)". The conclusion is obvious: something that actually ***happens*** must be ***possible.*** This property will be crucial in the following development: if an event $A \subset \Omega$ is observed to occur, then (since this is yet the crisp case) $\Pi(A) = 1$.

But just because an event ***has not*** occurred does not mean that it ***cannot*** occur, that it is ***not*** possible, perhaps even ***completely*** possible. Some event may be possible, but has simply not occurred ***yet.*** The die may be hidden to us before it is rolled, and each roll may produce a new face not previously seen. Just because a five has not yet appeared does not mean that a five is ***not*** possible, only that it ***may or may not*** be possible. But once all six faces have been observed at least once, then they must be given unitary possibility. Therefore $0 < \Pi(A) \leq 1$ means that A is possible, in the sense of "not prohibited", but also not necessarily ***seen.***

3.3 Stochastic possibility

Possibility theory exists as part of GIT, and stands in close relation to probability, both being fuzzy measures with distributions. In moving away from crisp possibility to consider the semantics of graded possibility, it is reasonable to take the semantics of probability into account.

Kosko reminds us that "After the fact 'randomness' looks like fiction [16]," referring to the well-known "paradox" of probability statements: if after flipping the coin a heads is observed, what sense does it make to say that the probability of heads is one half? Some suggest that once an event A has occurred, then it should be realistically said that $\Pr(A) = 1$. Surely then it must also be said that $\Pi(A) = 1$. This is in keeping with the idea that occurrence implies complete possibility.

In their seminal article on possibility theory [5], Gaines and Kohout go further in considering the status of possibility and probability statements over time. They observe that a "likely" event A, here interpreted as $\Pr(A) > 0$, actually has the property of **eventuality:** with increasing time the ***aggregate*** probability of A occurring approaches arbitrarily close to one. Their discussion deserves quotation at length.

> We have been very concerned to embody in our formulation the distinction between possible events that ***may not occur*** and possible events that are ***guaranteed to occur*** sooner or later ... Note that probability theory does not provide an explicatum of the first type of possible event. If for the purposes of analyzing an uncertain system we assign an uncertain event a non-zero probability then we imply that not only may it occur but also, in a sequence of occurrences each of which may be that event, it eventually ***will*** occur with a probability arbitrarily near one. The notional assignment of a definite probability to an event also fails to provide an adequate explicatum of the second type of possible event because it has the stronger implication that the relative frequency of such events in a sequence will tend to converge to

the given probability with increasing length of sequence.

Either or both of these connotations which probability has over possibility may be too strong in practical situations where the concepts of probability theory are being used to express the effects of uncertain behavior. For example, we are often faced with situations where an event E *may* occur, but there is no guarantee that E actually *will* occur, no matter how long we wait. If we ascribe some arbitrary probability to E then we certainly express that it is a possible event. However we are in a position to derive totally unjustified results based on the certainty of some eventual occurrence of E, or meaningless numeric results based on the actual "probability" of occurrence of E.

The possibilistic idea of occurrence should be extended to include the idea that if an event A *must* occur *sometime*, that is if A is *eventual*, then similarly $\Pi(A) = 1$.

4 Probability and possibility consistency

The observations of Section 3 lead directly to the following strong consistency requirement for possibility and probability statements.

4.1 A strong consistency criterion

Definition 1 (GK-consistency) *A function f is Gaines-Kohout-consistent, or just GK-consistent, with a function g on a set $X = \{x\}$ iff $f, g: X \mapsto [0,1]$ and $\forall x \in X, f(x) > 0 \leftrightarrow g(x) = 1$.*

Corollary 1 *If f is GK-consistent with g on X then $\forall x \in X, f(x) = 0 \leftrightarrow g(x) < 1$.*

Proof: Follows trivially from the restrictions $f(x), g(x) \in [0,1]$. ■

Principle 1 (Probability-Possibility Consistency) *Given a probability measure* Pr *and possibility measure* Π *then* Pr *should be GK-consistent with* Π *on* 2^Ω.

Simply stated,

$$\forall A \subset \Omega \qquad \begin{array}{l} \Pr(A) > 0 \leftrightarrow \Pi(A) = 1 \\ \Pr(A) = 0 \leftrightarrow \Pi(A) < 1. \end{array}$$

This states that something having non-zero probability is, following Gaines and Kohout, ***likely***, and therefore ***eventual***, and therefore equivalent to its being ***completely possible***. Conversely, a properly possible event ($\Pi(A) < 1$) must be of probability measure zero, and probability zero may or may not indicate proper possibility.

This is also completely in keeping with the standard probabilistic sense of the term "possible", in which a possible state is one with a non-zero probability. Starke expresses this view in his book on automata.

> We can apply non-deterministic automata to describe the "possibilities" of a given stochastic automaton in that we call "possible" those things which have a positive probability of being turned out. [19, p. 145]

Starke's view of possibility is crisp, and thus the interpretation is that $\Pr(A) > 0 \rightarrow \Pi(A) = 1$.

By Principle 1 an event may have zero probability and still have some degree of necessity, or a positive probability and no degree of necessity.

Corollary 2 *If* Pr *is GK-consistent with* Π *on* 2^Ω*, then* $\eta(A) > 0 \rightarrow \Pr(A) > 0$*, and* $\eta(A) = 0 \rightarrow \Pr(A) = 0$.

Proof: If $\eta(A) > 0$, then by Equation 2, $\Pi(A) = 1$, and so by Principle 1, $\Pr(A) > 0$. If $\eta(A) = 0$, then by Equation 2, $\Pi(A) < 1$, and so by Principle 1, $\Pr(A) = 0$. ■

If a probability and possibility measure are GK-consistent, then so are their (discrete) distributions.

Theorem 1 *If* Pr *is GK-consistent with* Π *on* 2^Ω*, then* $\vec{p}$ *is GK-consistent with* $\vec{\pi}$ *on* Ω.

Proof: $p: \Omega \mapsto [0,1]$ and $\pi: \Omega \mapsto [0,1]$, so the first condition of Definition 1 is satisfied. **Case 1:** Assume $\exists \omega \in \Omega, \exists a = p(\omega) \in (0,1], \exists b = \pi(\omega) \in [0,1)$. Then $\Pr(\{\omega\}) = a > 0, \Pi(\{\omega\}) = b < 1$, which violates the GK-consistency of Pr with Π on 2^Ω. Therefore $\forall \omega \in \Omega, p(\omega) > 0 \rightarrow \pi(\omega) = 1$. **Case 2:** Assume $\exists \omega \in \Omega, \pi(\omega) = 1, p(\omega) = 0$. Then $\Pi(\{\omega\}) = 1, \Pr(\{\omega\}) = 0$, which violates the GK-consistency of Pr with Π on 2^Ω. Therefore $\forall \omega \in \Omega, \pi(\omega) = 1 \rightarrow p(\omega) > 0$. ■

4.2 General consistency requirements

Aside from the subjective evaluation methods for determining possibility values mentioned in Section 1, it is also common to derive a possibility distribution by applying a conversion method to some given probability distribution [15]. These methods have been guided by a general principle of probability-possibility consistency, briefly stated by Delgado and Moral as "the intuitive idea according to which as an event is more probable, then it is more possible [1]," and summarized most generally by the formula

$$\forall A \subset \Omega, \quad \Pr(A) \leq \Pi(A). \tag{6}$$

Zadeh proposed a quantitative measure of this consistency

$$C(\vec{\pi}, \vec{p}) = \sum_i \pi_i p_i \in [0,1]$$

in his initial introduction of possibility theory [22]. $C(\vec{\pi}, \vec{p}) = 0$ indicates complete inconsistency and $C(\vec{\pi}, \vec{p}) = 1$ complete consistency. If $C(\vec{\pi}, \vec{p}) = 1$, then Equation 6 holds.[1]

Principle 1 is a stronger case of Zadeh's consistency: GK-consistency implies that $C = 1$, but not vice versa.

Theorem 2 *If p is GK-consistent with π on Ω, then* $C(\vec{\pi}, \vec{p}) = 1$.

Proof: If $p_i > 0$, then $\pi_i = 1$ and $p_i \pi_i = p_i$. If $p_i = 0$ then $p_i \pi_i = 0$. Therefore $\sum p_i \pi_i = \sum p_i = 1$. ■

4.3 Consequences of strong consistency

It should not be overlooked that Principle 1 is a deliberately strong requirement. But it should also be observed that it is not a *definition:* probability and possibility have been well defined in Section 2. Nor is it a *theorem:* probability and possibility are almost never even *defined* on the same random sets for comparison according to Principle 1.

Instead it is a *principle* which we assert as a *semantic* criterion, intended to relate different aspects of GIT in accordance with this extra-theoretical consideration. Zadeh also emphasized this point when he advanced his consistency measure C.

> It should be understood, of course, that the possibility-probability principle is not a precise law of a relationship that is intrinsic in the concepts of possibility and probability. Rather it is an approximate formalization of the heuristic observation that a lessening of the possibility of an event tends to lessen its probability, but not vice-versa. [22]

Principle 1 effects an interpretation of possibility (resp. probability) statements in the context of a given *specific* (resp. *consonant*) random set, which is, of course, formally inappropriate.

For example, given the following "trapezoidal" possibility distribution on $\Re$ (a typical "fuzzy number")

$$\pi(x) = \begin{cases} x, & x \in [0,1) \\ 1, & x \in [1,2] \\ 3 - x, & x \in (2,3] \\ 0, & \text{elsewhere} \end{cases},$$

what is the status of the expression $p(x)$? Principle 1 allows p to be any member of the class of probability distributions on $\Re$

$$\{p : x \in [1,2] \leftrightarrow p(x) > 0\},$$

effectively restricting the range of p to $\mathbf{C}(\pi) = [1,2]$. It provides no further information to determine p, and so by Insufficient Reason we should choose p to be uniform on $[1,2]$.

Similarly, in our example, there are six possible faces, so that $\forall 1 \leq \omega \leq 6, \Pi(\{\omega\}) = 1, \Pr(\{\omega\}) > 0$, yielding distributions $\vec{p} = \langle p_i \rangle$ with $p_i > 0$ and $\vec{\pi} = \langle 1,1,1,1,1,1 \rangle$. The p_i have yet to be fixed, and will depend on the weightings of the faces. But as long as all six remain *possible*, with *some* positive probability of occurrence, the possibility values π_i must remain at 1.

Conversely, given some Gaussian probability distribution $p(x), x \in \Re$, what is the status of the expression $\pi(x)$? Since $\forall x \in \Re, p(x) > 0$, therefore, by Principle 1, $\forall x \in \Re, \pi(x) = 1$.

In fact, by GK-consistency, any probability distribution with a positive value on $\forall \omega \in \Omega$ yields the possibility distribution $\forall \omega \in \Omega, \pi(\omega) = 1$. This is the **completely uninformed possibility distribution**, which has maximal uncertainty [13]. As Insufficient Reason leads to equiprobability in the stochastic case, this possibility distribution represents complete possibilistic ignorance, with $\forall A \subset \Omega, \Pi(A) = 1$.

For an ω with $p(\omega) = 0$, GK-consistency is not more helpful, saying only that $\pi(\omega) < 1$.

The conclusion is that, by Principle 1, a standard probabilistic analysis yields *essentially no information* of a possibilistic nature. This is in keeping with the logical independence of probability and possibility, and the simultaneous weakness of possibilistic representations of uncertainty. Not *all* possibilistic analyses must yield maximally uninformative distributions, only those with a probabilistic source, and thus with a very *strong* informational structure.

[1] The measure C has been generalized by Delgado and Moral [1].

5 The conceptual basis of possibilistic semantics

The nature of possibilistic categories, processes, and concepts can now be characterized (in a decidedly semi-formal manner). Our thinking about uncertainty and indeterminism has been deeply molded by over three centuries of concepts and methods which have arisen from probability theory and statistics. This is natural since probability is the strongest, simplest, and historically prior representation of uncertainty. In approaching possibility theory, these existing mental models must be modified, if not abandoned.

5.1 Possibilistic mathematics

Possibilistic mathematics provides some indications of how to interpret possibilistic statements.

Nonspecificity: The key feature that distinguishes probability from all other classes of fuzzy measures on random sets is its specificity, that is $\forall A_j \in \mathcal{F}, |A_j| = 1$. Since in probability theory evidence values m_j are attached to singleton sets, and thus essentially to elements ω_i, therefore, assuming no underlying structure in Ω, they can be compared only in *value*.

In general, however, two evidence values m_1, m_2 can also be compared in terms of their relative cardinalities $|A_1|, |A_2|$, and the amount of "overlap" between A_1 and A_2, measured by $|A_1 \cup A_2|, |A_1 \cap A_2|, |A_1 - A_2|$, and $|A_1 \Delta A_2|$ (where Δ is the symmetric difference operator). This hybrid form is captured by the various new uncertainty measures available in possibility theory [13].

Normalization: From Equation 4, in a possibilistic random set $\exists \omega_i \in \mathbf{C}(\mathcal{S})$ for which $\pi_i = 1$, and is thus shared by all evidential claims A_i. It has an invariant core around which the data set or process varies and which is common to all of its components. It should be noted that large cores, and thus large areas which adhere to crisp standards, are not necessarily undesirable.

Consonance: The final essential property that characterizes a fully possibilistic random set is consonance, the nesting of focal elements within each other. The core then becomes the smallest focal element with the smallest cardinality, the innermost box of a nest which spreads out from it.

Linearity: The nested structure of $\mathcal{F}$ imparts a strong linear ordering to a possibilistic random set, including the ordering of the focal elements by inclusion $A_{i-1} \subset A_i$; the ordering of the possibility values $1 = \pi_1 \geq \pi_2 \geq \ldots \geq \pi_n > 0$; and the ordering established on the elements ω_i so that $A_i = \{\omega_1, \omega_2, \ldots, \omega_i\}$ and $\Pi(\{\omega_i\}) = \pi(\omega_i) = \pi_i$.

Thus a discrete possibility distribution consists of two components: the selection of one of the $n!$ permutations of the universe, and the assignment of maximum normalized weights to the ω_i such that $\pi(\omega_1) = \pi_1 = 1$. The ordering specifies a particular path through the universe, whereas the weights represent the "distance" in certainty values between them.

This *ordinal* property of possibilistic information is crucial: knowledge is not just divided among a set of otherwise indistinguishable entities, as in probability, but rather over a linear *structure*.

5.2 Possibilistic processes

Because of nonspecificity, traditional ideas of **randomness** are altered in possibility theory. In probability theory, a given universe is a partition, and the uncertainty is only as to which unit is selected.

> Probability measures apply to precise but differentiated items of information, while possibility measures reflect imprecise but coherent items (i.e., which mutually confirm each other) ... A probabilistic model is suitable for the expression of precise but dispersed information. Once the precision is lacking, one tends to quit the domain of validity of the model. [2, pp. 6, 13]

The size and relative overlap of events in a possibilistic process are constantly changing as it moves not through states $\omega \in \Omega$, but rather through *meta-states* $A \subset \Omega$, in a non-deterministic manner. Thus while probability represents ambiguity (an uncertain choice among distinct alternatives), possibility represents a lack of specificity (an uncertain, but monotonically increasing, distance from a central core).

5.3 Statistical interpretations

In a possibilistic process there is a variation not just of state, but of the *size* of the state: the core remains fixed, and the observed meta-state varies in extent *around* it. The visual image is not of a point traversing a state space, but of a shifting granularity, tolerance, or precision. In statistical terms, there is no shifting sample mean, but rather a shifting *variance* around a *fixed* mean, which is the core. Since the sample variance of a stationary process converges to a fixed value as the sample size increases, possibility theory may therefore be appropriate modeling *non-stationary* processes.

By the law of large numbers, large samples provide sufficient information to construct additive probability distributions. Thus possibility theory may be appropriate to modeling problems with small sample sizes. Here the weakness of a small sample is matched by the weakness of possibilistic information, while the strength of a large sample is matched by the strength of stochastic information. In fact, it might be hoped that with increased sample size, a possibilistic analysis would become *less* useful, just as a stochastic treatment of the same problem would become *more* accurate.

5.4 Locality, extensibility, and mutability

Possibilistic information is highly *local* to specific elements ω_i. Whereas stochastic normalization is necessarily a property of the *whole* distribution $\sum_{i=1}^{n} p_i = 1$, possibilistic normalization can be satisfied by a *single* (perhaps not unique) element of the distribution.

Possibility is also **mutable** and **extensible:** all non-unitary and some unitary elements of the distribution may be modified, added, or removed without other elements being changed, and without any global rescaling, renormalization, or recalculation (although sometimes some global reordering will be required).

Together, these properties may provide important computational efficiencies for possibilistic models.

5.5 Possibility and complexity

The observation about small sample sizes leads to a discussion of **complex systems**. It can be difficult, if not impossible, to re-establish such systems in initial conditions for multiple time trials; once a measurement is made, the system can become perturbed, never to return to its previous state.

Thus complex systems can be relatively impervious to traditional experimental methods, since they generally do not yield the kind of strong time-series data required for stochastic models. But at the same time, their behavior is full of uncertainty. Instead, it seems completely appropriate to attempt possibilistic analyses, which require appropriately weak information, of such systems.

Weaver describes how stochastic methods are especially appropriate for dealing with repeated experiments on simple systems [20]. The simplest systems are indistinguishable, and statistical techniques for handling their uncertainty are very successful, as in statistical physics and thermodynamics.

"Aggregate" systems can become "historically bound", evolving to their future states through a long series of very specific actions. Thus they gain in distinguishability, in the limit of truly complex systems (such as organisms, species, and ecosystems) to actual uniqueness [4]. In these cases, repeated experiments are truly impossible.

Indeed, a hallmark property of complex systems is *exactly* the *possibility*, rather than the *eventuality*, of their states. Such systems are highly "non-ergodic": given a very large state space, only a very small portion of that space could ever be visited. Thus there will be a large number of properly possible states, but a small number of "eventual" states, perhaps even none. Kampis has suggested [10] that this is an appropriate definition of **emergence**, since making a prediction in the state space requires (at least!) an intractable computation. It may be that the properties of possibilistic mutability and extensibility make possibility theory more appropriate to model the "surprises" such systems provide.

Thus there is a link from *simplicity* to *high probability* and then to *crisp possibility*. Conversely, there is a link from *complexity* and increased *difficulty* to *low probability, proper possibility*, and finally *impossibility*.

Models of complex systems risk intractability by "combinatorial explosion". The computational efficiencies of possibility theory may make it attractive for complex systems modeling.

5.6 Capacity vs. frequency concepts

Since possibilistic data are not frequency data, possibility cannot be regarded in the context of concepts such as **likelihood, chance, tendency, propensity**, or **proportion**.

Instead, possibility suggests interpretations in the context of **capacity**. Given a set of "buckets", they can all be completely full, many can be empty, or they may be in some intermediate state, as long as at least one is full (for normalization). The concepts related to possibility include **intensity, degree of fulfillment or satisfaction, ease of fulfillment, distance from optimality, similarity, elasticity**, and **preference**.

These concepts are all *ordinal*, with states measured by their distance from some reference state of maximal capacity (intensity, preference, etc.). Kosko echoes this view, but in the context of general fuzziness, not possibility theory proper.

Fuzzy theory [is] the theory that *all things admit degrees*, but admit them deterministically. Fuzziness ... measures the degree to which an event occurs, not whether it occurs. Randomness describes the uncertainty of *event occurrence.* An event occurs or not, and you can bet on it. [16]

6 Possibilistic measurement

Central to any question of possibilistic semantics is the issue of **measurement:** how is it that possibility values are determined? This has been one of the most vexing questions facing fuzzy set theory over the years, and the lack of objective measurement methods is widely seen as one of its largest weaknesses.

Unfortunately, space does not permit a full discussion of non-subjective possibilistic measurement. These concepts and results are introduced elsewhere [8], and are still under development. They are based on statistics gathered not on elements $\omega_i \in \Omega$, but rather set-statistics on subsets $A_j \subset \Omega$ [3]. Any resulting consonant class yields a possibilistic random set; any globally non-disjoint class yields a non-stochastic, non-possibilistic empirical random set which is possibilistically normal; and possibilistic normalization methods exist for the remaining empirical random sets.

This method is essentially equivalent to the measurement of non-disjoint intervals, which make a probabilistic treatment impossible. Positive overlap amongst all such intervals results in a "possibilistic histogram" similar in form to a fuzzy number. Sources of interval measurements include synchronous measurements from multiple, heterogeneous instruments (similar to Lemmer's "hidden labels" approach [17]); distance of order statistics from a central "possibilistic mean"; and local extrema of time-series data.

References

[1] Delgado, M and Moral, S: (1987) "On the Concept of Possibility-Probability Consistency", *Fuzzy Sets and Systems*, v. **21**, pp. 311-318

[2] Dubois, Didier and Prade, Henri: (1988) *Possibility Theory*, Plenum Press, New York

[3] Dubois, Didier and Prade, Henri: (1989) "Fuzzy Sets, Probability, and Measurement", *Eur. J. of Operations Research*, v. **40**:2, pp. 135-154

[4] Elsasser, WM: (1970) "The Role of Individuality in Biological Theory", in: *Towards a Theoretical Biology*, v. **3**, ed. CH Waddington, pp. 153, Edinburgh U. Press, Edinburgh

[5] Gaines, BR and Kohout, LJ: (1976) "The Logic of Automata", *Int. J. Gen. Sys.*, v. **2**:4, pp. 191-208

[6] Goodman, IR and Nguyen, HT: (1986) *Uncertainty Models for Knowledge-Based Systems*, North-Holland, Amsterdam

[7] Hughes, GE and Creswell, MJ: (1968) *Introduction to Modal Logic*, Methuen, London

[8] Joslyn, Cliff: (1992) "Possibilistic Measurement and Set Statistics", in: *Proc. 1992 NAFIPS Conference*, v. **2**, pp. 458-467

[9] Joslyn, Cliff: (1993) "On Possibilistic Automata", in: *Proc. EuroCAST '93*, to appear

[10] Kampis, George: (1991) *Self-Modifying Systems*, Pergamon, Oxford

[11] Kampè de Fériet, J: (1982) "Interpretation of Membership Functions of Fuzzy Sets in Terms of Plausibility and Belief", in: *Fuzzy Information and Decision Process*, ed. MM Gupta and E Sanchez, pp. 93-98, North-Holland, Amsterdam

[12] Klir, George: (1991) "Generalized Information Theory", *Fuzzy Sets and Systems*, v. **40**, pp. 127-142

[13] Klir, George: (1992) "Probabilistic vs. Possibilistic Conceptualization of Uncertainty", in: *Analysis and Management of Uncertainty*, ed. BM Ayyub et al., pp. 13-25, Elsevier

[14] Klir, George and Folger, Tina: (1987) *Fuzzy Sets, Uncertainty, and Information*, Prentice Hall

[15] Klir, G and Parviz, B: (1992) "Probability-Possibility Transformations: A Comparison", *Int. J. Gen. Sys.*, v. **21**:1

[16] Kosko, B: (1990) "Fuzziness vs. Probability", *Int. J. Gen. Sys.*, v. **17**:2-3, pp. 211-240

[17] Lemmer, JF: (1986) "Confidence Factors, Empiricism and the Dempster-Shafer Theory of Evidence", in: *Uncertainty in Artificial Intelligence*, ed. LN Kanal and JF Lemmer, pp. 117-125, North-Holland, Amsterdam

[18] Näther, Wolfgang: (1990) "On Possibilistic Inference", *Fuzzy Sets and Systems*, v. **36**, pp. 327-337

[19] Starke, PH: (1972) *Abstract Automata*, North Holland, Amsterdam

[20] Weaver, Warren: (1968) "Science and Complexity", *American Scientist*, v. **36**, pp. 536-544

[21] Webster: (1989) *New Merriam Webster Dictionary*, Merriam-Webster, Springfield MA

[22] Zadeh, Lofti A: (1978) "Fuzzy Sets as the Basis for a Theory of Possibility", *Fuzzy Sets and Systems*, v. **1**, pp. 3-28