



Form 4-750 (2004)
OMB No. 1545-0047

[illegible]NSN 7540 01 280 5500

91-25-033

UNCLASSIFIED

NAME OF RESPONSIBLE PERSONNEL

L. R. Goodman

DTIC Accession Number

06190553 4013

DTIC Accession Number

06190553 4013

Accession For	
NTIS CRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Availability for Special
A-1	20



UNCERTAINTY MODELING AND POSSIBILISTIC
APPROACH TO PARAMETER ESTIMATION

by

I.R. GOODMAN *

and

H.T. NGUYEN **

* Surveillance Systems Department, Code 7223
Naval Ocean Systems Center,
San Diego, CALIF. 92152

** Department of Mathematical Sciences,
New Mexico State University,
Las Cruces, N. MEX. 88003

Work was supported in part by the Naval Electronic Systems Command
under Program ELEX NR-021-01 , FY-82.

Uncertainty Modeling and Possibilistic
Approach to Parameter Estimation

Abstract

A general basis is established for modeling uncertainty and concepts of vagueness through a combined possibilistic-probabilistic approach. A survey of the connections between probability theory and fuzzy set theory is undertaken. New results concerning parameter estimation, given both statistical and vague or natural language information, are presented. New discoveries involving possibility measures and linguistic probabilities are also demonstrated. Some specific topics treated include analysis of conjunction and disjunction operators within a general setting, relationships between possibilistic and random set operators, and asymptotic behavior of fuzzy set estimators. Application to the development of a correlation/tracking technique - PACT : Possibilistic Approach to Correlation and Tracking - is presented in outline form.

CONTENTS

Uncertainty Modeling and Possibilistic Approach to Parameter Estimation

by I.R. Goodman and H.T. Nguyen

	page
Abstract	1
1. Introduction	2
2. Motivation	4
3. Basic Framework of Analysis	9
4. Logical or Fuzzy Set Systems: Choice of a Particular System	15
5. Basic Critique and Synopsis of Manes' Generalized Uncertainty Theory..	19
6. Issues Arising in Modeling Uncertainties/ Conclusions	29
7. References	32

Appendix A - Some Connections between Probabilistic and Fuzzy Set Descriptions of Uncertainties

by I.R. Goodman

Appendix B : Part One - Outline of PACT (Possibilistic Approach to Correlation and Tracking) Project
 Part Two- An Approach to the Data Association Problem through Possibility Theory
 Part Three- Micro-Mini Course in Fuzzy Set Theory

by I.R. Goodman

Appendix C - Possibility Measures

by H.T. Nguyen

Appendix D - Linguistic Probabilities

by H.T. Nguyen

1. INTRODUCTION

This report is the final paper during FY-82 for the project "Combined Possibilistic Approach to Parameter Estimation" under the aegis of the Naval Electronic Systems Command, ELEX 6121, Task Area XR-021-01, Program Element 61153 N .

Previously, a quarterly **progress** report was issued by the first named author (see Goodwin [1]). Because of the delay in funds for the project, a reduced work effort for FY-82 has resulted. Consequently, the second, third and last progress reports have been eliminated in favor of this encompassing summary report for FY-82.

In the first quarterly report, the nature and range of the associated NOSC IR/IED PACT Project (Possibilistic Approach to Correlation and Tracking) was outlined. The project is centered about the PACT algorithm. Briefly, the PACT algorithm utilizes probabilistic and fuzzy set or possibilistic concepts in treating statistical and semantic based information for the data association and tracking problem. Basically, the algorithm operates upon a collection of information attribute error distributions, or equivalently, matching tables and a collection of prechosen (relevant to the given scenario) inference rules. The basic output of the algorithm is the posterior distribution (in either the classical statistical or fuzzy set sense) of correlation possibilities between any two given candidate track histories which are suitably updated.

The scope of the task which this report addresses is to define fully the interrelations of possibility and probability theory so that practical quantitative applications can be developed. This obviously relates directly to the PACT Project. Rigorous techniques must be developed for dealing with estimation problems which contain within their formulation further connections between possibility and probability theory. Specifically, this includes, as outlined in a previous communication (see Schuster [2]) :

- (a) Determination of proper possibilistic conjunction and disjunction operators
- (b) Development of relations between fuzzy set operators and ordinary set ones

- (c) Derivation of a sampling theory which parallels random sampling theory for both large and small sample sizes.

It should be noted that the first quarterly progress report addressed in part the last mentioned issue.

This report consists of a main overview text written jointly by both authors, which contains : An introductory section (1) scoping out the paper's purposes; A section (2) concerning motivation for the analysis of fuzzy sets, random sets, and their connections through the presentation of a general example followed by some discussion; A section (3) on the framework of analysis consisting of the topics : internal vs. external modeling, prelinguistic concepts and ideas obtained through natural language, concept of fuzziness as a primitive, development of general logical systems and ambiguities, relationships between fuzzy sets and their operations and probabilistic concepts. A section (4) on logical (or fuzzy set) systems centering about a procedure for choosing a particular system; A section (5) concerning Manes' important generalization of probabilistic and min-max fuzzy set systems, with a new result characterizing Manes' system relative to general fuzzy set systems; and finally, a section (6) concerning open issues and general conclusions.

The first appendix (I.R. Goodman) is a thorough mathematical survey of the various relations between possibility theory and probability theory. In section 10 of this appendix, there is a series of new results concerning parameter estimation, for any given arbitrary combination of probabilistic or possibilistic input descriptions. The second appendix, part one (I.R. Goodman) is an outline - in slide form - of the PACT Project, a direct consequence of earlier and ongoing work in parameter estimation (as mentioned above in the first appendix). The second part of Appendix B (I.R. Goodman) is essentially the same as the latest paper summarizing the theoretical structure of PACT, which has been submitted for inclusion in the Proceedings of the Fifth MIT/ONR Workshop on C³ Systems, 1982. The third part of Appendix B (I.R. Goodman) is a brief summary of basic fuzzy set/random set theory for background purposes. Finally, the last two appendices (Appendix C and D) (H.T. Nguyen) are surveys and analyses of two topics of key interest: possibility measures and linguistic probabilities.

2. MOTIVATION

Consider the following example :

Example 1.

Two target histories of interest are being considered by a sensor operator for possible correlation. He receives information labelled as A_1 , A_2 , A_3 :

- A_1 - two dimensional position observations with associated updated error ellipses of some prescribed confidence level
- A_2 - reports concerning tentative classification of targets, such as Filipino Type Q4 or Liechtensteinian Type R7.9
- A_3 - visual sightings including partial identifications, clues, hull lengths, mast shapes, etc.

Clearly, if information categories A_2 and A_3 are ignored, statistical hypotheses testing theory may be applied to A_1 to establish a standard weighted metric (Mahalanobis distance - see e.g., Rao [3], Chapter 8) for testing for correlation and determining the level of correlation (statistically) between the two track histories of interest. Rough gating procedures could then be added to see if A_2 and A_3 confirm or perhaps contradict the crucial geolocation criterion results for A_1 . (See the Naval Ocean-Surveillance Correlation Handbook [4], [5]) for a listing and descriptions of tracking/correlation systems utilizing the above procedure.) But, how should gates or their softer distributional analogues be systematically established for A_2 and A_3 and integrated with the results for A_1 ? Furthermore, can we use human in-field operator experience to relate in some way information matches, mismatches, and everything in between, occurring for categories A_2 and A_3 and A_1 , as well, to correlation levels? This problem will be solved in detail later in the report. Indeed, this example, is a simplification of the problem that the PACT algorithm addresses. (See Appendix B.)

In the above example, the parameter of interest is the true correlation level.

Other examples of related nature- both military and non-military- may be found in Goodman [6]. The above example motivates us to formulate the following scheme for dealing with parameter estimation problems. We need to:

- (1) Categorize the incoming information into subcategories as $A_1, A_2, A_3, \dots$. These categories should be carefully chosen for nonredundancy. Further analysis of their taxonomy should prove useful.
- (2) Establish a rigorous and systematic framework for quantifying information. This may be identified with the problem of determining the natural domains of attributes $A_1, A_2, A_3, \dots$. For example, in Example 1, A_1 's domain consists of all ordered pairs of 2 by 1 vectors and ellipses, while for A_2 , perhaps simplified labels such as Q4, R7.9, S6, etc. will do.
- (3) Derive matching level tables or equivalently error "distributions" (in a sense possibly extending the classical statistical ones) that can occur between what is reported or observed and what the true values are for each attribute category.
- (4) Determine relative weights of importance between the various attribute categories. For example, how much will we tolerate a total mismatch with respect to A_2 , when a relatively good match occurs relative to A_1 ?
- (5) Establish logical connections - based on either physical considerations or human operator experience - between the various information categories. These connections could be either in the form of inference - modus ponens rules or posterior distributions constraining or delineating the unknown parameter of interest.

(See Goodman [7] - [9] for earlier outlines for approaching parameter estimation when some of the informational input is in linguistic based format. See also [10], section 10, for extensive theoretical results. Note that essentially, Appendix A of this paper is [10], while Appendix B, part two is [9].)

More generally, we might ask: How do we model uncertainty and conclusions concerning this within a general framework?

Obviously, once we remove ourselves from "hard" statistical information, subjectivity and personal interpretive variation play important roles. The logical connections mentioned in item 5 above are usually given in the form of modus ponens inference rules or posterior distributions. In the latter case, at least for the classical statistical situation, Bayes' Theorem is usually invoked with respect to more basic conditional data and prior distributions. For the former case, predicates restricting or describing the unknown parameter of interest and other parameters or values determined through various attribute categories are related through an "if () then ()" structure. Schematically, we have the restrictions on unknown parameter Q, say :

$$\Phi(Q|Z) \longleftrightarrow \text{" If } P(Z) \text{ then } V(Q) \text{ " .}$$

The lefthand side represents the posterior distribution (possibilistic or ordinary - see Appendix for clarification and explanation of possibilistic distributions) of Q given data Z, while the righthand side represents the inference rule " If P(Z) is true then V(Q) must also hold ", where P(Z) is some predicate describing Z and V(Q) is some predicate describing Q. For example, let $Z=(Z^{(i)}, Z^{(j)})$, with $Z^{(i)} = (Z_1^{(i)}, Z_2^{(i)})$ and $Z^{(j)} = (Z_1^{(j)}, Z_2^{(j)})$, the superscript i referring to data for track history i and j for track history j, the subscript 1 referring to attribute category A_1 and 2 referring to attribute A_2 . Then, we may let, for example,

$$P(Z) = \begin{array}{l} \text{" } Z_1^{(i)} \text{ and } Z_1^{(j)} \text{ very mildly (or to a low level) match} \\ \text{and } Z_2^{(i)} \text{ and } Z_2^{(j)} \text{ strongly match " } \end{array}$$

$$V(Q) = \text{" correlation level Q is high at least " .}$$

The words "very mildly", "strongly", "high at least" could, if sufficient information were present, be replaced by more quantitative values such as $\alpha_1, \alpha_2, \alpha_3$, where each α_k is some number between 0 and 1 indicating the intensity of matching level.

(Again, see Appendix B for clarification and elaboration of this idea.)

It should be noted that many posterior distributions and/ or inference rules may be present which delineate the possible values of Q given Z .

With the problem modeled according to the above mentioned scheme, the next basic question concerns the actual mathematization or translation of the problem into symbols which can be manipulated according to some established calculus. How do we translate the problem into a consistent rigorous framework? What means do we use to translate the atomic or fundamental information parts? How are compound informational parts to be exhibited through appropriate choice of operators ? Which parts of the problem are more amenable to ordinary statistical / probabilistic analysis and hence modeling and which parts to possibilistic/ fuzzy set modeling? Certainly, natural language descriptions appear more easily put into fuzzy set structure than classical probabilistic ones . (See [1] for example) On the other hand, even concepts that may appear statistically describable may also be modeled via fuzzy set theory. For example, classification often entails rather overlapping possible values or classes, due to the vector of subcategories in which an object has to lie to qualify to be in a particular class. Indeed, some classes may actually be subsets of others. In addition, it is possible that the exact definitions for the classes themselves are vague and moreover the relations between the classes may not be clear. For example , A_2 could contain C and D , which are defined by knowledge of frequencies, ship size, shape, and number of engine emitting energy sources on-board. Overlaps between C and D may abound. Thus it is not appropriate to consider ordinary probability distributions over A_2 , since the elementary events C and D , among others for example, are not distinct or disjoint. Rather, because of the overlapping flavor of A_2 , either random subsets of A_2 or more generally certain equivalence classes of random subsets of A_2 (equivalence here is in the sense of having the same one-point coverage function - see [12]) or (the same as) fuzzy sets.

Analysis of uncertainty in its most general form requires modeling and

measuring imprecise concepts expressed by natural language. By natural language, we mean that medium through which all human ideas are formed including classical set theory and two-valued logic as well as more ambiguous concepts including "tall" , "happy" , Gaussian distributions , uniform distributions, "within 3 units of" , "close to" , "almost all" , " there exists" , " approximately a subset of to degree 0.6 " , " member of set G " , approximately a member of H to degree 0.4" , etc.

The difficulty with natural language modeling used as ²direct tool for analysis is the disorganization of the field. Despite the heavy influence of pioneer linguists Sapir, Bloomfield , Jespersen, Boas, Whorf (~~and the famous~~ Whorf-Sapir hypothesis on language restricting the thoughts of the native speaker) and the later work of formal linguists Hiz, Harris, and Chomsky , among others (see [13]- [18]) , what is evident is that a unified theory of linguistics entailing both semantics (meaning) and syntactics (operations, form) is needed which is suitable for complete mathematization. One candidate approach is due to Zadeh [19] using fuzzy set theory. (See also the work of Grenander, for a different perspective [20].)

Other approaches to the modeling of uncertainties which do not deal directly with the modeling and emulating of natural language, but do treat the information content content contained therein, include :

1. Probability theory, including random variables, random functions, and more recently, random sets.
2. Multivalued logic/ truth theory and the ensuing set theory developed from it
3. Fuzzy set theory / possibility theory.
4. Flou or multiple set theory. This includes interval and sensitivity theory.
5. Extremal entropy techniques. This area could optionally be treated under probability theory, because of its close relationship.
6. Dempster-Shafer uncertainty theory- belief, doubt, plausibility measures .

3. BASIC FRAMEWORK OF ANALYSIS

The basic analysis of uncertainty modeling revolves around a series of general topics :

1. Internal vs. external modeling.

In the internal approach, explicit analytic relations are sought connecting one approach to uncertainty to another. For example, Negoita and Ralescu, through their Representation Theorem (see [21]) tied up very neatly classical fuzzy set operations with flou set operations. As another example of the internal approach, Goodman, Orlov, Nguyen, and Höhle, among others (see [12], [22]-[24]) demonstrated direct connections between random sets and fuzzy sets and certain of their operations.

In the external approach, unifying generalizations are sought which reduce to various approaches to uncertainty modeling. Here the work of Hirota [25] , Schefe [26] Gaines [27] may be cited for developing structures that simultaneously generalize probability and fuzzy set systems. The most far reaching work in this area is due to Manes [28] who derived a collection of axioms which not only generalize probability theory and min-max fuzzy set theory, but a whole host of other systems, including topological neighborhood theory and credibility theory. (See section 5 for more details on Manes' work, where it is shown certain restrictions must be imposed on fuzzy set systems to satisfy Manes' axioms.)

2. Prelinguistic concepts and ideas obtained through natural language.

This topic concerns itself with the ability of natural language to express ideas accurately and succinctly as well as the formalizing or mathematizing of natural language for dealing with uncertainties. Comments were made previously on the lack of progress in this extremely difficult area. Ironically, we can express within a few words, ideas such as love, happiness, temporal vague concepts, ambiguous descriptions - which are perfectly understandable to another reasonably educated speaker , as well as various combinations and operations on this ideas - yet cannot express this concepts easily within a rigorous framework in terms of all

the component primitive or atomic parts. On the other hand "complicated " mathematical terminology , such as is typically found in category theory or algebraic topology or deductive logic studies, really express concepts far simpler in nature than what language can express. (Of course, we cannot discount the ability of language to represent- albeit how awkward- pure mathematical concepts .)

3. Concept of fuzziness as a primitive

It is the firm conviction of both authors as well as Zadeh and others (see e.g. [29]) that because mathematical analysis has shown that fuzziness is a weak form of randomness , i.e., a looser type of randomness without the constraints of the probability distribution entailed, fuzzy sets and their operations are a natural tool to express linguistic concepts , rather than probability distributions. Thus, the fundamental idea of a point partially belonging to a set with degree specified as some number between 0 and 1 may well be taken as an intuitive concept representing the possibility that the point is in the set, rather than the probability it is in the set (the set now considered as a random set). (See Appendix A for the development of explicit relations between fuzzy sets, random sets, and random variables. Essentially, a fuzzy set -with similar results valid for many fuzzy set operations- is equivalent mathematically to the class of all random sets which have in common the same one-point coverage function, namely the membership function of the given fuzzy set.)

4. Development of general logical (multivalued logic) systems . Ambiguities.

This topic is a basis for further work in developing unified approaches to fuzzy set modeling. Too often in the past (see for example, Dubois and Prade [11]) myriad distinct fuzzy set systems have been proposed for use in modeling uncertainties, without paying attention to the inherent ambiguity of definition present. More specifically , consider the problem of defining an appropriate concept for the intersection of two fuzzy sets. Originally, Zadeh (1965 [30]) proposed that minimum as an operation on the respective membership functions was the most appropriate. Later, Bellman and Giertz

[31] were among the first to justify on a rigorous basis the use of minimum as an intersection operation. (See also the survey of Klement^[32] on rigorous characterizations of various fuzzy set operations, including intersection, union, complementation.) However, the justification required certain constraints (mutual distributivity) which are not realistically required within a general setting. Other definitions for intersection resulted, including the use of product - also justified, with again appropriate restrictions. (Again, see the Dubois and Prade survey [11] .) Clearly, minimum and product, while both extending ordinary intersection (relative to zero-one type membership functions) are considerably different. Which one to choose or not? The answer to this problem may well lie in defining an entire class of fuzzy set operations - not just a single operation- which in the most natural way abstract the ordinary concept of intersection. Such a class has been proposed (see Klement [11] and Goodman [29]), called the class of t-norms (a term borrowed from a branch of probability totally independent of fuzzy set theory, developed by Schweizer and Sklar [33] based on earlier proposals of Menger [34]). These operators are symmetric, associative, usually assumed also continuous, obey certain boundary conditions for compatibility with ordinary intersection, are nondecreasing in their arguments, and numerically are bounded above by (the largest t-norm) the minimum operation. (See, again [33] for properties) See also Haack [35] and Rescher [36] for listings of multivalued logic systems where special cases of t-norm and t-conorm operators are used for "and" and "or".)

Similar, other classes of definitions may be developed for union (t-conorms) complement (involutions) , and in turn these general definitions, à la multivalued logic may be used to develop general compound fuzzy set operations and relations, including implication, the quantifications "for all" and "there exists" , as well as subset and arithmetic (in a fuzzy set sense) relationships. In addition, this leads to the general concept of conditional fuzzy sets (analogous to conditional probability distributions), fuzzy set Bayes' theorem and in turn, a theory of small and large fuzzy set sampling . (See Goodman [37] .)

Even with the general unifying approach as described , problems of ambiguity of definition still arise. For example, the fuzzy set quantifier "for all x in X , () holds" may be definable by the relation , as expressed in English, " $()_{x_1} \& ()_{x_2} \& ()_{x_3} \& \dots \& ()_{x_n}$ " ,

where $()_{x_j}$ represents some relationship such as "if x_j is in A then x_j is in B", x_j varying over the universe of discourse X , and where A and B are fuzzy sets predetermined here.

On the other hand, the same concept could conceivably be expressed directly by a unary operation on $\bigcap_{j=1}^n \phi(x_j)$, which is a fuzzification of "for all", i.e., a monotonically increasing function over the unit interval which rises sharply towards one near domain value one and which otherwise is zero before these values. (Similar monotonic operations are used in fuzzy set theory to define "almost all", "at least most", etc.) It is easy to see that in general, though again the two concepts extend the ordinary meaning in zero-one set theory of "for all", they represent two different approaches to the universal quantifier. Both definitions, it should be noted, depend on the general t-norm definition of intersection (through $\&$), the second depending as well through the "if () then ()" relation on the general definition of union and complement. Which one to choose? In turn, the problem of appropriately modeling a particular fuzzy set with respect to the corresponding original numerical or linguistic concept also arises. What individual variation of response should be allowed? How specific should the universe of discourse be? For example, when considering the fuzzy set "long", do we consider ships, cars, both; is there some grand scale where "long" can be quantitatively established through its membership function, other than the obvious fact that it is some monotone increasing membership function? One answer to this problem is analagous to the problem of modeling an appropriate probability distribution: parameterize and then choose the most appropriate value of the parameter-and hence membership function from the collection through some estimation technique based on empirically obtained evidence.

5. Relationships between fuzzy sets and their operations and probabilistic concepts.

The details of these relationships are spelled out in Appendix A. Recalling the last comments of subsection 3 above, given a fuzzy set, it may be expedicious to choose one particular random set equivalent to it. Which one to choose? How much information loss occurs when one random set is chosen as opposed to the entire equivalence class?

Could some mathematical criterion be used to weed out this random set such as maximal entropy? What about semantic content? For example, consider the simple fuzzy set representing "tall". Clearly, this also is a monotone increasing fuzzy set, i.e., the membership function must be monotonically increasing. However, it can be shown that among the equivalent random sets to any such monotone increasing fuzzy set (see e.g. [29]), two very different ones can be explicitly shown: the so-called S_U -type, which is a random interval with right end point fixed at the maximal universe of discourse element, and the T-type, which is very broken-up in structure and is not any kind of interval. Clearly, the first is more compatible with the concept of tall- if one point is covered randomly by a random set representing "tall", shouldn't all points to the right, i.e., having larger heights? Similarly, there may be a most natural choice of random set representation for a given fuzzy set, when the latter has a membership of some prescribed type, such as unimodal, continuous, discrete, step-form, etc.

In a related vein, we may pose the question as how should semantic based information, and thus fuzzy sets, be combined with independently derived random set or random variable information concerning a common unknown parameter vector?

Finally, it is of importance to ascertain, through the relationships mentioned above, if random set theory could be used to derive results for fuzzy set theory.

A general schematic outlining the various approaches to uncertainty modeling is given in Figure 1.

4. LOGICAL OR FUZZY SET SYSTEMS : CHOICE OF A PARTICULAR SYSTEM

All of fuzzy set theory can be developed in a natural way directly through homomorphic-like evaluations from multiple valued truth theory . In turn, multivalued truth theory can be motivated through natural language analogues consisting of attributes and linguistic operators.

In natural language, there exist many compound attributes and operators built up from more primitive ones. Three connectives play a very significant role in natural language descriptions : "not" , "and" (usually denoted by &), and "or" . It can be shown that all other logical connections can be reduced to these in classical two-valued logic. Indeed, theoretically, all operations can be reduced to one operator , Shefer's stroke operator or alternative denial (see 38 ,sect.1.5). More recently, it has been shown (see [39]) that for multivalued logic systems, a similar reduction is also formally valid. In any case, it is natural to begin with a small collection of primitive logical connectives, add procedures for compounding them, and denote the more important results by particular terms. We will call the triple of operators (not, & , or) a logical (or fuzzy set) system. The interpretation of this triple of operators - with some collection of predicates or sentences subsumed on which these operators act- will be at three conceptually different levels : linguistic , multivalued logic, and fuzzy set (or numerical). The operator not is a unary operator on $\mathcal{P}$ while & and or are binary operators over $\mathcal{P}$. Difficulties arise as to the numerical evaluation of these operators, at the fuzzy set interpretation. For example (as stated earlier in this text) & could yield both product and minimum as interpretations at the fuzzy set end. Which to choose or other operators to consider? As mentioned previously, a basic choice of permissible operators for a logical system is for not to correspond to some decreasing involutive operator, while & corresponds to some t-norm and or to some t-conorm. Other choices for interpreting & and or may lose the flavor of what we really mean by "and" and "or" . For example, it is possible that the weighted sum operator (see [29],e.g.) - which is its own DeMorgan dual (in general, & and or can be interpreted so that they are not in the classical DeMorgan relationship) - could be chosen in place of & and or for a logical system , But note that this averaging type of operator is neither a t-norm nor a t-conorm.

Choice of a specific logical system - i.e., a specific triple of admissible operators for not, &, or - should be based on some combination of practical considerations and theoretical basis. It is shown in Appendix A, Theorems 10-12, that the subfamily of semidistributive t-norms and t-conorms (DeMorgan property may be assumed) provides a theoretical basis of operators to choose from in setting up a logical system, when not is associated with the fuzzy set operator $1-(\cdot)$. This basis essentially allows weak homomorphic identification of arbitrary combinations of the t-norm and t-conorm operators in question, as well as a large class of fuzzy arithmetic and functional transform operators involving these t-norms and t-conorms, with naturally corresponding ordinary intersections and unions relative to random sets (equivalent in the usual one point coverage sense to the original fuzzy sets and determined also by the choice of t-norm and t-conorm, with respect to jointness of distributions).

With the theoretical basis established firmly: we should only in general seek those t-norms and t-conorms which are semidistributive, we must use the structure of a given scenario to further define the appropriate choice for the logical system to be used for the modeling.

In review (see [29]), the semidistributive family of t-norms consists of all discrete weighted sums of the form

$$\psi_{\&} = \sum_{k=1}^{m+1} \alpha_k \cdot F_k + \sum_{k=1}^m \beta_k \cdot G_k ; \quad 0 \leq \alpha_k, \beta_k \leq 1 ; \quad \sum_{k=1}^{m+1} \alpha_k + \sum_{k=1}^m \beta_k = 1 ,$$

where the weights

$$\alpha_k \stackrel{d}{=} a_k - b_{k-1} , \quad \beta_k \stackrel{d}{=} b_k - a_k ; \quad k = 1, 2, \dots, m+1 ;$$

$$0 = b_0 \leq a_1 \leq b_1 \leq a_2 \leq b_2 \leq \dots \leq a_m \leq b_m \leq a_{m+1} = 1 \quad \text{fixed arbitrary constants ;}$$

each F_k is that bivariate probability distribution function corresponding to the random vector

$\begin{pmatrix} U_1 \\ U_2 \end{pmatrix}$, where $U_1 = U_2$ is uniformly distributed over the interval $[b_{k-1}, a_k]$, and thus

$F_k\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) \neq 0$ iff $\begin{pmatrix} x \\ y \end{pmatrix}$ is in the square $[b_{k-1}, a_k] \times [b_{k-1}, a_k]$, in which case,

$$F_k\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \min((x - b_{k-1}) / (a_k - b_{k-1}), (y - b_{k-1}) / (a_k - b_{k-1})) ;$$

each G_k is that bivariate probability distribution function corresponding to the random vector $\begin{pmatrix} U_1 \\ U_2 \end{pmatrix}$, where now U_1 and U_2 are statistically independent identically distributed uniformly over the interval $[a_k, b_k]$, thus $G_k\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) \neq 0$ iff $\begin{pmatrix} x \\ y \end{pmatrix}$ is in the square $[a_k, b_k] \times [a_k, b_k]$, in which case,

$$G_k\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = ((x-a_k)/(b_k-a_k)) \cdot ((y-a_k)/(b_k-a_k)).$$

Note the important special cases $\psi_{\&} = F_1 = \min$ and $\psi_{\&} = G_1 = \text{prod.}$

Suppose now that a given scenario holds so that operator (t-norm) $\psi_{\&}$ is obtainable at least approximately empirically through its use, by a tabulation of its values over the unit square. In general, this empirically obtained function, denoted by $\psi_{\&}^o$ is not a semidistributive t-norm. However, we can approximate $\psi_{\&}^o$ by a semidistributive t-norm, determining the coefficients α_k and β_k in the expansions given above by matching moments to any degree desirable between $\psi_{\&}$ and $\psi_{\&}^o$. Since the moments of F_k and G_k are particularly easy to obtain, we immediately obtain the fundamental relations

$$\begin{aligned} \mu(\gamma_1, \gamma_2) &= (1/(\gamma_1 + \gamma_2 + 1)) \cdot \sum_{k=1}^{m+1} (a_k^{\gamma_1 + \gamma_2 + 1} - b_{k-1}^{\gamma_1 + \gamma_2 + 1}) \\ &+ (1/((\gamma_1 + 1) \cdot (\gamma_2 + 1))) \cdot \sum_{k=1}^m ((b_k^{\gamma_1 + 1} - a_k^{\gamma_1 + 1}) \cdot (b_k^{\gamma_2 + 1} - a_k^{\gamma_2 + 1})) / (b_k - a_k), \end{aligned}$$

where the $(\gamma_1, \gamma_2)^{\text{th}}$ central moment for $\psi_{\&}^o$ is given by

$$\mu(\gamma_1, \gamma_2) \stackrel{d}{=} \int \int_{\substack{\text{over all} \\ x, y \text{ in the} \\ \text{unit square}}} x^{\gamma_1} \cdot y^{\gamma_2} d\psi_{\&}(x, y),$$

obtainable by an approximating discretization, for some predetermined size of m and choice of equalities to hold in the above equation for the a_k 's and b_k 's; γ_1 and γ_2 allowed to run separately over the integers $0, 1, 2, \dots, q$, for some appropriate choice of q .

We then seek to solve the above fundamental relations for the coefficients a_k and b_k

in terms of the empirically known $\mu(\gamma_1, \gamma_2)$'s.

For earlier empirical approaches to the determination of $\psi_{\&}$ and ψ_{or} , see the work of Zimmermann [40] and Hersh and Caramazza [41].

5. BASIC CRITIQUE AND SYNOPSIS OF MANES' GENERALIZED UNCERTAINTY THEORY

Manes [28] has recently written one of the most important papers taking the external approach to the analysis of uncertainty modeling. (See the earlier papers of Gaines [27] and Hirota [25] for much more restricted approaches to the development of systems which extend both probability and fuzzy set concepts.) Because of the paper's great importance and relative difficult format, a synopsis with some critical comments will be presented here.

In summary, Manes derives a general class of "fuzzy" theories which reduce as special cases to finitely discrete probability theory, Zadeh's original (see [30]) fuzzy set system $(1-(\cdot), \min, \max)$, topological neighborhood theory, credibility theory, and other approaches to the modeling of uncertainties. (To avoid confusion with fuzzy set theories as discussed in the overview, Manes' models will be designated as "fuzzy" .) Manes' theories are characterized by a triple of functions satisfying three basic axioms. The first function assigns to any space a corresponding space of generalized distributions over that space (the term "generalized" to be made more specific for each particular system satisfying the axioms) The second function imbeds or identifies any given space with the subclass of mass-point distributions from the associated space of generalized distributions. The third function extends conditional generalized distributions -with index set X , say, and distributions in the associated space $T(Y)$ of space Y - to operators taking initial distribution space $T(X)$ to $T(Y)$. As natural as the first two axioms are (see below) in determining the essence of the above concepts, the third axiom, an associative-like condition imposed on the extension function (i.e., the third function described above) may be too restrictive. This is because of the following characterization we have shown : Let $F = (1-(\cdot), \psi_{\&}, \psi_{or})$ be a given fuzzy set system (recalling that $\psi_{\&}$ is some t-norm and ψ_{or} is some t-conorm), with Θ being membership function binary relation composition relative to F . Then F is also a "fuzzy" theory with the mapping Φ_{Θ} as a function of all binary fuzzy set relation membership functions ϕ_R being the same as an extension function in Manes' sense , iff $\psi_{\&}$ is right distributive over ψ_{or} . As a consequence of this result, a whole host of fuzzy set systems

including the important system $(1-(\cdot), \text{prod}, \text{probsum})$ which plays a key role in the weak-homomorphism theory connecting fuzzy set operations with those ordinary ones on random sets (see Appendix A), do not satisfy the required third axiom of Manes' "fuzzy" theories, although it is easily verified that the other two axioms are indeed satisfied always by any fuzzy set system. Consequently, the elegant results obtained for Manes' theories do not apply to non-right distributive fuzzy set systems. The natural problem that follows is how to change the requirements of Manes' third axiom so as to accomodate a larger class of fuzzy set systems.

Next, a brief treatment with some analysis of Manes' results is presented.

A "fuzzy" theory is a triple $\mathcal{T} = (T, e, \#)$ of functions so that for all well-defined spaces X, Y, Z

- (i) $T(X)$ is the space of generalized distributions associated with X
- (ii) $e(X): X \rightarrow T(X)$ is the special function called the imbedding function of X into the mass-point distributions in $T(X)$
- (iii) $\# : T(Y) \xrightarrow{X} T(Y) \times T(X)$ is the extension or lifting operator, noting that the dependence of $\#$ on X and Y is not denoted here.

The axioms $\mathcal{T}$ must satisfy are :

For any spaces X, Y, Z and mappings $\alpha : X \rightarrow T(Y)$, $\beta : Y \rightarrow T(Z)$

- (1) $(\#(\alpha)) \circ e(X) = \alpha$
- (2) $\#(e(X)) = \text{ident}_{T(X)}$
- (3) $\#(\#(\beta) \circ \alpha) = (\#(\beta)) \circ (\#(\alpha))$; $\circ$ denoting composition of functions.

Zadeh Fuzzy Set Theory Example - $F = (1-(\cdot), \min, \max)$

$T(X) = [0,1]^X$ and $(e(X)(x))(y) = \delta_{x,y}$ (Kronecker delta) ; $x \in X, y \in Y$ arb.

$$((\#(\phi_R))(\phi_A))(y) = \max_{x \in X} (\min(\phi_R(x,y), \phi_A(x))) = \phi_{R \odot A}(y),$$

for any fuzzy binary relation R on $X \times Y$ and any fuzzy subset A of X ; all $y \in Y$.

Finitely Discrete Probability Theory Example

$T(X) = \{p \mid p \text{ is a probability function over } X \text{ with finite support}\}$

$$(e(X)(x))(y) = \delta_{x,y}, \text{ for all } x \in X, y \in Y,$$

$$(\#(q(\cdot | \cdot)))(p))(y) = \sum_{x \in X} (q(y | x) \cdot p(x)) = E_V(q(y | V)),$$

for all conditional probability functions q over Y , conditioned on X , and all $p \in T(X)$, with random variable V distributed according to probability function p .

Theorem A (Goodman-Nguyen, 1982)

Let $F = (1-(\cdot), \psi_{\&}, \psi_{\text{or}})$ be a given fuzzy set system with $\odot$ denoting the fuzzy set binary relation composition relative to F . Thus (see also [29] for motivation via multivalued truth theory) $\odot$ is defined dually as

$$(\phi_R \odot \phi_A)(y) \triangleq \phi_{R \odot A}(y)$$

$$\triangleq \text{proj}_Y (R \cap (A \otimes Y))(y)$$

$$= \psi_{\text{or}} (\psi_{\&} (\phi_R(x,y), \phi_A(x))) ;$$

$$x \in X$$

and let

$$(e(X)(x))(y) \triangleq \delta_{x,y}, \text{ for all } x \in X, y \in Y, X, Y \text{ arbitrary,}$$

with

$$T(X) \triangleq [0,1]^X \text{ for all } X,$$

and finally define

$$(\#(\phi_R))(\phi_A) \triangleq \phi_{R \odot A}, \text{ for all fuzzy binary relations } R \text{ on } X \times Y$$

and all fuzzy subsets A of X .

Then :

- (1) F satisfies Manes' axioms (1) and (2) for "fuzzy" theory
- (2) F satisfies Manes' axiom (3), and hence is also a "fuzzy" theory, iff $\psi_{\&}$ is right distributive over ψ_{or} , i.e.,

$$\psi_{\&} (a, \psi_{\text{or}}(b, c)) = \psi_{\text{or}}(\psi_{\&}(a,b), \psi_{\&}(a,c)) ; \text{ all } a,b,c \in [0,1].$$

Proof of Theorem A:

Result (1) follows from straightforward algebra.

For result (2) :

The lefthand side of Manes' axiom (3) is evaluated as , for any $\alpha = \phi_R$, $\beta = \phi_S$, and ϕ_A , where R is a fuzzy relation on $X \times Y$, S is a fuzzy relation on $Y \times Z$, and A is a fuzzy subset of X, and $z \in Z$ is arbitrary

$$I(R,S,A;z) \stackrel{d}{=} \psi_{\text{or}} \left(\psi_{\&} \left(\phi_A(x), \psi_{\text{or}} \left(\psi_{\&} \left(\phi_R(x,y), \phi_S(y,z) \right) \right) \right) \right) ,$$

$x \in X \qquad y \in Y$

Similarly, the righthand side of axiom (3) becomes

$$II(R,S,A;z) \stackrel{d}{=} \psi_{\text{or}} \left(\psi_{\&} \left(\psi_{\text{or}} \left(\psi_{\&} \left(\phi_A(x), \phi_R(x,y) \right) \right), \phi_S(y,z) \right) \right) ,$$

$y \in Y \qquad x \in X$

If $\psi_{\&}$ is distributive over ψ_{or} , then by simple induction , the defining equation for right distributivity may be extended to arbitrary number of arguments for the ψ_{or} operator. In turn substituting this into the above expressions for I and II yields the relation

$$I(R,S,A;z) = II(R,S,A;z) = \psi_{\text{or}} \left(\psi_{\&} \left(\phi_A(x), \phi_R(x,y), \phi_S(y,z) \right) \right) .$$

$\left(\begin{matrix} x \in X \\ y \in Y \end{matrix} \right)$

For the converse, if axiom (3) is satisfied , then for all R,S,A,z as above, we must have $I(R,S,A;z) = II(R,S,A;z)$. Choose in particular $A=X$, $Y=\{y\}$, with all other variables arbitrary. It follows immediately that right distributivity holds.

□

As a consequence of the above theorem, Zadeh's original fuzzy set system is also a "fuzzy" theory , as is the non-DeMorgan system $(1-(\cdot), \text{prod}, \text{max})$. On the other hand, the (DeMorgan) fuzzy set system $(1-(\cdot), \text{prod}, \text{probsum})$, and extending this, any semi-distributive fuzzy set system (except for the boundary system $(1-(\cdot), \text{min}, \text{max})$) in general will not be right distributive, and hence will not be a "fuzzy" set theory. Since semi-distributive fuzzy set systems are the main link in the weak homomorphism theory connecting fuzzy set theory with random set theory (see Appendix A for details), the concepts of fuzzy and "fuzzy" are not compatible here.

Manes considers the set of truth distributions $\mathcal{L}_0 \triangleq T(\{T_0, F_0\})$, where T_0 (not to be confused with the function T) and F_0 are fixed values representing truth component and falsehood component, respectively. Alternatively, Manes could have considered the more standard $T(\{T_0\})$, or perhaps expanded the truth indices to three or more such as in $T(\{T_0, F_0, M_0\})$, where M_0 is some intermediate truth index such as "maybe".

In addition, Manes, defines "fuzzy" theory mappings, quotients, canonical subtheories, requiring commutivity for all relevant arguments and functions. For example, a "fuzzy" abstract n -ary operator is defined as an n -ary operator (on the Cartesian product of generalized distribution spaces) which commutes with all mappings arising from the extension operation. Because of this restriction, it is shown (Manes' Theorem 3.3) that any such operator may be always identified with the operator $(\#(q))(p)$, for some fixed $p \in T(\{1, \dots, n\})$ as a function of $q = (q_j)_{j=1, \dots, n}$; $q_j \in T(X)$, $j=1, \dots, n$. Similarly, a "fuzzy" homomorphism $\varrho: T(X) \rightarrow T(Y)$ is characterized by its ability to commute with all abstract n -ary operators. This is equivalent (see Manes' Theorem 4.1) to the equation

$$\varrho = \#(\varrho \circ e(X)); \text{ all } X.$$

Closely related to the above, it is shown that $\varrho = \#(\alpha)$ is always a "fuzzy" homomorphism between $T(X)$ and $T(Y)$, when $\alpha: X \rightarrow T(Y)$. Extending this result, it follows that for any $f: X \rightarrow Y$, $T(f) \stackrel{d}{=} \#(e(Y) \circ f)$ is always a homomorphism also, which respects identity and functional composition, as well as one-to-one ness and onto-ness. (This makes $T(f)$ as a function of f a functor.)

Zadeh Fuzzy Set Theory Example

$$T(f)(\phi_A) = \sup_{x \in f^{-1}(\cdot)} (\phi_A(x)) = \phi_{f(A)}; \text{ for any } f: X \rightarrow Y \text{ and } \phi_A \in T(X).$$

Thus $T(f)$ is the basic fuzzification of functions f .

Finitely Discrete Probability Theory Example

$$(T(f))(p) = \sum_{x \in f^{-1}(\cdot)} p(x) = \text{probability function of r.v. } f(V), \text{ where } V \text{ has probability}$$

function p .

Manes extends the idea of the relationship or compatibility of an ordinary set with a generalized distribution, and in turn the relationship between two generalized distributions as follows :

$$dm_X(S, p) \stackrel{d}{=} (T(g_S))(p),$$

where

$$g_S \stackrel{d}{=} \#(e(\{T_0, F_0\}) \circ \chi_S) : T(X) \longrightarrow \mathcal{L}_0,$$

$$\chi_S(x) = \begin{cases} T_0, & \text{iff } x \in S \\ F_0, & \text{iff } x \notin S \end{cases}, \text{ for any } S \subseteq X; p \in T(X).$$

$$eq_X(p, q) \stackrel{d}{=} (\#(dm_X(\{x\}, q))(p)) ; \text{ all } p, q \in T(X),$$

noting the relation

$$eq_X((e(X))(x), q) = dm_X(\{x\}, q) ; \text{ all } q \in T(X), x \in X.$$

These relationships reduce to quite familiar ones for fuzzy set and probability theories:

Zadeh Fuzzy Set Theory Example

$$(eq_X(\phi_A, \phi_B))(T_0) = \sup_{x \in X} \min(\phi_A(x), \phi_B(x)) = \sup_{x \in X} \phi_{A \cap B}(x)$$

$$(eq_X(\phi_A, \phi_B))(F_0) = \sup_{\substack{x \in X \\ y \in Y \\ x \neq y}} \min(\phi_A(x), \phi_B(y))$$

$$(dm_X(S, \phi_A))(T_0) = \sup_{x \in S} \phi_A(x)$$

$$(dm_X(S, \phi_A))(F_0) = \sup_{x \notin S} \phi_A(x) ; \text{ all fuzzy subsets } A \text{ of } X, B \text{ of } Y$$

Finitely Discrete Probability Theory Example

$$eq_X(p, q) = \sum_{x \in X} p(x) \cdot q(x) = \Pr(V=W)$$

$$dm_X(S, \phi_A) = \sum_{x \in S} p(x) = \Pr(V \in S) ; \text{ all probability functions } p, q \text{ over } X$$

which are finitely discrete, where r.v. V corresponds to p and W to q .

Define the imbedding of true and that of false in the distribution space $\mathcal{L}_0$ as

$$T_1 \stackrel{d}{=} (e(\{T_0, F_0\}))(T_0) \quad ; \quad F_1 \stackrel{d}{=} (e(\{T_0, F_0\}))(F_0) .$$

Then define the following properties which "fuzzy" theories may or may not possess:

A theory is anti-reflexive iff $e_{\mathcal{L}_X}(p, p) = T_1$ implies that $p = (e(X))(x)$; all $p \in T(X)$, $x \in X$, X arbitrary

A theory is faithful iff $d_{\mathcal{L}_X}(\cdot, p)$ as a function of p is one-one and onto.

A theory is propositionally complete iff for any $p, q \in T(X)$, $p \neq q$, there exists a separating homomorphism $\#(\alpha)$, for some $\alpha: X \rightarrow \mathcal{L}_0$, i.e., $(\#(\alpha))(p) \neq (\#(\alpha))(q)$.

A theory is consistent iff for any $f, g: X \rightarrow Y$, with $f \neq g$, $T(f) \neq T(g)$, equivalently, $e(X): X \rightarrow T(X)$ is a one-one function.

Summary of theorems and results shown by Manes for the above properties.

- (1) All faithful theories are propositionally complete.
- (2) Except for **two degenerate cases** of no interest, all theories must be consistent.
- (3) Every theory has a largest (canonical) subtheory with crisp points , i.e., $T(\{x\})$ consists of exactly one generalized distribution for any $x \in X$ arbitrary, equivalently,

$$d_{\mathcal{L}_X}(X, p) = T_1 \text{ for all } p \in T(X) \text{ and } x \in X, \text{ where } T \text{ is associated with the subtheory.}$$

For the Zadeh fuzzy set theory example, this condition is equivalent to the class of all fuzzy subsets of X which have supremum of their membership function being 1.

For the Finitely discrete probability theory example , $T(X)$ itself has crisp points.

- (4) Both examples (Zadeh's fuzzy set theory and finitely discrete probability theory) possess all four theory properties defined above.

Generalization of "fuzzy" theory homomorphisms and the operator $\#$ to multivariate settings and relationships to joint and independent generalized distributions.

An n -homomorphism relative to a theory is a mapping $\mathcal{P}: T(X_1) \times \dots \times T(X_n) \rightarrow T(Y)$ such that $(\mathcal{P} \mid T(X_j))$ separately (all other spaces fixed) is a homomorphism for the theory, $j=1, \dots, n$.

The space of all joint generalized distributions with respect to X and Y is $T(XXY)$.

For any $p \in T(X)$ and $q \in T(Y)$,

$$\Gamma_1(p, q) \triangleq \{ \#(\{ \#((e(XXY))(\cdot, y)) \}_{y \in Y}) \} (q)$$

and

$$\Gamma_2(p, q) \triangleq \{ \#(\{ \#((e(XXY))(x, \cdot)) \}_{x \in X}) \} (p)$$

are the only two ways to minimally extend p and q into a joint generalized distribution over X and Y .

Theorem A (Manes, Theorem 5.7 and related material)

(1) $\Gamma_1 = \Gamma_2$, for all $p \in T(X)$ and all $q \in T(Y)$, denoting the common value mapping Γ ,

and is a 2-homomorphism (with respect to the theory) which extends $e(XXY)$ uniquely

iff for $\alpha: X_1 \times \dots \times X_n \longrightarrow T(Y)$, there is a unique extension map

$$\bar{\alpha}: T(X_1) \times \dots \times T(X_n) \longrightarrow T(Y), \text{ noting that for } n=1, \bar{\alpha} = \#(\alpha), \text{ for all possible } \alpha$$

iff the abstract n -ary operators induced by p and q commute with respect to composition, for all p, q .

(2) For any commutative theory, i.e., a theory in which any of the equivalences in (1) are

valid, the operator eq_X is symmetric in its arguments.

(3) For any commutative theory, having crisp points is equivalent to Γ being a one-one mapping

(4) Both Zadeh's fuzzy set theory and finitely discrete probability theory are commutative theories.

The above theorem motivates the designation of Γ as the independent joint distribution forming mapping.

Zadeh Fuzzy Set Theory Example.

$$\Gamma(\phi_A, \phi_B) = \min(\phi_A, \phi_B) \text{ (pointwise)}$$

$$\bar{\alpha}(\phi_A, \phi_B) = \sup_{\substack{x_j \in X_j, \\ j=1,2}} (\min(\phi_A(x_1), \phi_B(x_2), \alpha(x_1, x_2))) \text{ (pointwise)}; \text{ all fuzzy subsets}$$

A of X_1 , B of X_2 , and all $\alpha: X_1 \times X_2 \longrightarrow T(Y)$; X_1, X_2, Y arbitrary spaces.

Finitely Discrete Probability Theory Example

$$\Gamma(p, q) = p \cdot q$$

$$\bar{\alpha}(p, q) = \sum_{\left(\begin{smallmatrix} x_j \in X_j \\ j=1,2 \end{smallmatrix} \right)} \left(p(x_1) \cdot q(x_2) \cdot \alpha(x_1, x_2) \right) \quad (\text{pointwise})$$

$$= E(\alpha(V_1, V_2)),$$

where r.v. V_1 corresponds to p and r.v. V_2 corresponds to q , the two r.v.'s being statistically independent; for all $p \in T(X_1)$, $q \in T(X_2)$, $\alpha: X_1 \times X_2 \rightarrow T(Y)$ arbitrary.

Corellary to Theorem A. Generalized "Fuzzification" Principle.

Suppose the "fuzzy" theory under consideration is a commutative one. Let $f: X_1 \times \dots \times X_n \rightarrow Y$ be arbitrary. Then, using the notation of Theorem A, $\tilde{f} \stackrel{d}{=} (\overline{e(Y) \circ f}): T(X_1) \times \dots \times T(X_n) \rightarrow T(Y)$ is an n -homomorphism relative to the theory which is the unique extension of $e(Y) \circ f$ and is also called, by definition, the n -homomorphic extension of f . Note that if $n=1$, then $\tilde{f} = \#(e(Y) \circ f)$. f is also the "fuzzification" of f . ■

Let a theory be commutative and as a special case of the above corollary, let $f: \{T_0, F_0\} \times \dots \times \{T_0, F_0\} \rightarrow \{T_0, F_0\}$ be an arbitrary n -ary ordinary truth table function or logical operator. Then $\tilde{f}$ is called the "fuzzification" of operator f . Also, define the "fuzzified" Boolean logic associated with the given theory as $(\mathcal{L}_0, H)$, where H is the set of all possible "fuzzifications" of n -ary logical operators.

Zadeh's Fuzzy Set Theory Example.

For any n -ary logical operator f and any $t_1, \dots, t_n \in \mathcal{L}_0$,

$$(f(t_1, \dots, t_n))(\cdot) = \sup_{\left(\begin{smallmatrix} x_1, \dots, x_n \\ \in f^{-1}(\cdot) \end{smallmatrix} \right)} \min(t_1(x_1), \dots, t_n(x_n))$$

In particular:

For the or operator $\text{or} : \mathcal{L}_0 \times \mathcal{L}_0 \rightarrow \mathcal{L}_0$, where $\text{or}(T_0, T_0) \stackrel{d}{=} \text{or}(T_0, F_0) \stackrel{d}{=} \text{or}(F_0, T_0) \stackrel{d}{=} T_0$ and $\text{or}(F_0, F_0) \stackrel{d}{=} F_0$, for all $t_1, t_2 \in \mathcal{L}_0$,

$$\begin{aligned}(\widetilde{\text{or}}(t_1, t_2))(T_0) &= \max(\min(t_1(T_0), t_2(T_0)), \min(t_1(T_0), t_2(F_0)), \min(t_1(F_0), t_2(T_0))), \\(\widetilde{\text{or}}(t_1, t_2))(F_0) &= \min(t_1(F_0), t_2(F_0)).\end{aligned}$$

For the operator $\& : \mathcal{L}_0 \times \mathcal{L}_0 \rightarrow \mathcal{L}_0$, where $\&(T_0, T_0) \stackrel{d}{=} T_0$ and $\&(T_0, F_0) \stackrel{d}{=} \&(F_0, T_0) \stackrel{d}{=} \&(F_0, F_0) \stackrel{d}{=} F_0$, for all $t_1, t_2 \in \mathcal{L}_0$,

$$\begin{aligned}(\widetilde{\&}(t_1, t_2))(T_0) &= \min(t_1(T_0), t_2(T_0)), \\(\widetilde{\&}(t_1, t_2))(F_0) &= \max(\min(t_1(T_0), t_2(F_0)), \min(t_1(F_0), t_2(T_0)), \min(t_1(F_0), t_2(F_0))).\end{aligned}$$

For the negation operator $\text{not} : \mathcal{L}_0 \rightarrow \mathcal{L}_0$, where $\text{not}(T_0) \stackrel{d}{=} F_0$ and $\text{not}(F_0) \stackrel{d}{=} T_0$, for all $t \in \mathcal{L}_0$,

$$(\widetilde{\text{not}}(t))(T_0) = t(F_0) \quad \text{and} \quad (\widetilde{\text{not}}(t))(F_0) = t(T_0).$$

Finitely Discrete Probability Theory Example.

For any n-ary logical operator f and any $t_1, \dots, t_n \in \mathcal{L}_0$,

$$(f(t_1, \dots, t_n))(\cdot) = \sum_{\substack{(x_1, \dots, x_n) \\ \in f^{-1}(\cdot)}} (t_1(x_1) \cdots t_n(x_n))$$

In particular:

For the operator or (as defined earlier), for all $t_1, t_2 \in \mathcal{L}_0$,

$$(\widetilde{\text{or}}(t_1, t_2)) = \text{probsum}(t_1, t_2) \text{ (pointwise)}.$$

For the operator $\&$ (as defined earlier), for all $t_1, t_2 \in \mathcal{L}_0$,

$$(\widetilde{\&}(t_1, t_2)) = t_1 \cdot t_2 \text{ (pointwise)}.$$

For the operator not (as defined earlier), for all $t \in \mathcal{L}_0$,

$$\widetilde{\text{not}}(t) = 1-t \text{ (pointwise)}.$$

6. ISSUES ARISING IN MODELING UNCERTAINTIES / CONCLUSIONS

This section is loosely held together by the commonality of posing questions (with not too many answers at present, unfortunately) concerning the natural questions that arise between natural language/semantics, linguistics, fuzzy set theory and probability theory as approaches to the modeling of uncertainty.

Dubois & Prade (III, 255-264) have devoted a section of their compendium on the developments in fuzzy set theory and its applications to the modeling of fuzzy set membership functions. A number of basic approaches are considered, with some emphasis-as should be-on the empirical aspects of the modeling. In this vein, it should be added that the weak homomorphic theory developed for example in Appendix A shows that the idea of counting percentage of times a fixed value possesses a given attribute - which is interpreted as the evaluation of the corresponding fuzzy set membership function- as is typically done by survey sampling of individuals, may be identified with the empirical one point coverage probability function generated by a random sample of random sets that are identically distributed and are weak equivalent to the attribute or fuzzy set in question. However, the following issue has not been sufficiently emphasized: In classical statistical techniques, modeling of distribution functions is often carried out in two basic steps. First, a parameterized family of distributions is chosen. This may be for reasons of invariance, shape, use of Central Limit Theory, or via trends of earlier empirical evidence. The family is chosen so that it reasonably contains the viable alternatives for the true distribution and its size is adjusted accordingly. Then empirical data or restrictions are imposed-such as unbiasedness, sufficiency, minimal risk with respect to some choice of loss function on errors and estimates of the unknown parameter value or outcome - yielding either a unique value or a reduced set of values where the unknown parameter lies. This leads to the conclusion that the same procedure should be applied to other approaches to uncertainty modeling, where distributions, in now a generalized or different sense from the classical probabilistic one, play a key role in the theory. Examples of this are fuzzy set theory, with its fuzzy set membership functions, the first topic mentioned above, flow set theory with its index functions relative to the individual

sets forming a given flow set, and topological neighborhood theory with its neighborhood filters. (See Manes' unified treatment of these theories as presented in [28] or section 5 of this paper.)

Other analogues can be established between fuzzy set theory as applied to parameter estimation and classical statistical estimation theory. For example, one can assume a linear regression model is valid connecting observations with an unknown parameter vector, with no specification of the relevant distributions involved-at least, at first. Then least squares, or more generally, a least weighted functional defined on the potential errors between observations and possible values of the parameter in question, is derived, yielding a reasonable value of the unknown parameter as a function of the observations, i.e., a statistic, if distributional assumptions were to be made. Then if a fuzzy set modeling approach is taken, the observations could be assumed to be generated from corresponding fuzzy set membership functions, yielding in turn through the standard fuzzification of an ordinary function (see Goodman [29] or Dubois and Prade [11] for further details on the fuzzification or equivalently the obtaining of the membership function of an ordinary function operating over a space on which fuzzy subsets are also defined) the fuzzy set membership function of the "statistic". In turn this leads in a natural way to confidence sets for the unknown parameter vector, by for example considering the level sets associated with the aforementioned membership function. (See Goodman [12], e.g.) for a related technique. A forthcoming publication will consider this problem in more detail.) Asymptotic properties of these (fuzzy set) estimators may also be obtained as the sample sizes are increased. Bayesian techniques may also be developed involving conditional distributions in the fuzzy set sense. (See Goodman [37] for development of these concepts for general fuzzy set systems. See also Dishkant's related work [42].)

Is fuzzy set theory rich enough to model reasonably natural language, or can it be extended or modified- such as by considering, via the weak homomorphic representation theory of random sets (see Appendix A), two point, and in general, multiple point coverage functions?

Can we express all human (or other?) ideas or concepts in terms of natural language? Can these ideas be reduced to primitives and operators involving them? Can a mathematical/logical procedure be developed for describing and analyzing natural language in a unified way, including attributes and operators? What quantitative relationships can be established between prelinguistic ideas and natural language? How sensitive or robust and how subjective are concepts translated into natural language with respect to the particular language chosen (Whorf-Sapir hypothesis is involved, see [43]), individual , and medium used? Is there always inherent ambiguity in modeling a given concept in natural language? Can we make use of the enormous body of literature available which treats formal linguistics and semantics to develop a systematic unified framework directly relatable to multivalued logic theory? In a related manner, we may ask if a unified approach to uncertainty models (à la Manes (see section 5) , for example) and to natural language would be possible? Can the efficiency of the various approaches to modeling uncertainty be meaningfully compared?

7. REFERENCES

1. I. R. GOODMAN, "First Quarterly Report on FY82 Exploratory Development Task Assigned, 'Combined Possibilistic-Probabilistic Approach to Parameter Estimation,'" NOSC Memo Ser 722/39 IRG:ss of 26 May 1982, Naval Ocean Systems Center, San Diego, CA.
2. J. SCHUSTER, "Task Assignment for Combined Possibilistic-Probabilistic Approach to Parameter Estimation," Memo ELEX 612:JS Ser 2480-612 of 25 Feb 1982, Naval Electronic Systems Command, Washington, D.C.
3. C. R. RAO, "Linear Statistical Inference and Its Applications," 2nd Ed., John Wiley & Sons, Inc., New York, 1973.
4. I. R. GOODMAN, H. L. WIENER, and W. W. WILLMAN, "Naval Ocean-Surveillance Correlation Handbook, 1979," NRL Report 8402, Naval Research Lab, Washington, D.C. 1980.
5. H. L. WEINER, W. W. WILLMAN, I. R. GOODMAN, and J. H. KULLBACK, "Naval Ocean-Surveillance Correlation Handbook, 1978," NRL Report 8430, Naval Research Lab., Washington, D.C. 1979.
6. I. R. GOODMAN, "Identification of fuzzy sets with a class of canonically induced random sets and applications," NRL Report 8415, Naval Research Lab, Washington, D.C., 1980.
7. I. R. GOODMAN, Applications of possibility theory to ocean surveillance correlation, Proc. 48th MORS, Naval Postgraduate School (1981), 101-109.
8. I. R. GOODMAN, Applications of a combined probabilistic and fuzzy set technique to the attribute problem in ocean surveillance, 20th IEEE Conf. Decis. & Control, (1981) 1409-1412.
9. I. R. GOODMAN, An approach to the data association problem through possibility theory, Proc. 5th MIT/ONR Workshop on C3, Naval Postgraduate School, 1982.
10. I. R. GOODMAN, Some connections between probabilistic and fuzzy set descriptions of uncertainties, submitted 1982 to J. Math. Analy. & Applic.
11. D. DUBOIS and H. PRADE, "Fuzzy Sets & Systems," Academic Press, New York, 1980.
12. I. R. GOODMAN, Fuzzy sets as equivalence classes of random sets in "Recent Developments in Fuzzy Set and Possibility Theory," (R. R. Yager, Ed.), Pergamon Press, 1982.
13. J. T. GRINDER and S. H. ELGIN, "Guide to Transformational Grammar," Holt, Rinehart & Winston, Inc., New York, 1973.
14. C. S. OSGOOD, G. J. SUCI, and P. M. TANNENBAUM, "The Measurement of Meaning," Univ. of Illinois Press at Urbana, 1971.
15. J. LYONS, "Introduction to Theoretical Linguistics," Cambridge Univ. Press, 1971.

16. G. HARMAN (Ed.), "On Noam Chomsky: Critical Essays," Anchor Press/Doubleday, New York, 1974.
17. E. H. LENNENBERG with appendices by N. Chomsky and O. Marx, "Biological Foundations of Language," John Wiley & Sons, Inc., New York, 1967.
18. J. H. GREENBERG (Ed.), "Universals of Language," 2nd Ed., MIT Press, Cambridge, MA, 1966.
19. L. A. ZADEH, Quantitative fuzzy semantics, Info. Sci 3 (1971), 159-176.
20. U. GRENANDER, "On Mathematical Semantics: A Pattern Theoretic View," Applied Math. Dept. Report in Pattern Analysis, No. 71, Brown Univ., Providence RI, July 1978.
21. C. V. NEGOITA and D. A. RALESCU, Representation theorems for fuzzy concepts, Kybernetes (UK) 4 (1975), 169-174.
22. A. I. ORLOV, Fuzzy and random sets, Appl. Multi. Stat. Analy. J. Acad. Sci. USSR, Moscow (in Russian) (1978), 262-280.
23. H. T. NGUYEN, Some mathematical tools for linguistic probabilities, Fuzzy Sets & Sys. 2 (1979), 53-65.
24. U. HÖHLE, A mathematical theory of uncertainty in "Applied Systems and Cybernetics, Vol VI," (G. Lasker, Ed.), 2728-2733, Pergamon Press, New York, 1981.
25. K. HIROTRA, Concepts of probabilistic sets, Fuzzy Sets & Sys. 5 (1981), 31-46.
26. P. SCHEFFE, On foundations of reasoning with uncertain facts and vague concepts, Int. J. Man-Mach. Stud. 12 (1980), 35-62.
27. B. P. GAINES, Fuzzy and probability uncertainty logics, Info. & Control 38 (1978), 154-169.
28. E. G. MANES, A class of fuzzy theories, J. Math. Anal. & Applic. 85 (1982), 409-451.
29. I. R. GOODMAN, Some fuzzy set operations which induce homomorphic random set operations, Proc. 26 Conf. Gen. Sys. Res. (AAAS) (1982), 417-426.
30. L. A. ZADEH, Fuzzy sets, Info. & Control 8 (1965), 338-353.
31. R. E. BELLMAN and M. GIERTZ, On the analytic formalism of the theory of fuzzy sets, Info. Sci. 5 (1973), 149-157.
32. E. P. KLEMENT, Operations on fuzzy sets and fuzzy numbers related to triangular norms, Proc. 11 Int. Symp. Multi Logic (1981), 218-225.
33. B. SCHWEITZER and A. SKLAR, "Probabilistic Metric Spaces," book to be published 1983.

34. K. MENGER, Statistical metrics, Proc. Nat. Acad. Sci. USA, 28 (1942), 535-537.
35. S. HAACK, "Philosophy of Logic," Cambridge University Press, Cambridge (GB), 1978.
36. N. RESCHER, "Many-Valued Logic," McGraw-Hill Book Co., New York, 1969.
37. I. R. GOODMAN, Some asymptotic properties of fuzzy set systems, Proc. 2nd World Conf. Math. Serv. Man., Univ Pol. de las Palmas, Can. Is., 1982, 312-317.
38. H. B. ENDERTON, "A Mathematical Introduction to Logic," Academic Press, New York, 1973.
39. Z. MANNA, "Mathematical Theory of Computation," McGraw-Hill, New York, 1974.
40. H.-J. ZIMMERMANN, Results of empirical studies in fuzzy set theory in "Applied General Systems Research," (G. J. Klir, Ed.), pp. 303-312, Plenum Press, New York, 1978.
41. H. M. HERSH and A. CARAMAZZA, A fuzzy set approach to modifiers and vagueness in natural language, J. Exper. Psychol.: Gen. 105, No. 3 (1976), 254-276.
42. H. DISHKANT, About membership function estimation, Fuzzy Sets & Sys. 5 (1981), 141-147.
43. B. L. WHORF, "Language, Thought, and Reality," (J. B. Carroll, Ed.), John Wiley & Sons, Inc., New York, 1956.