ORNL/TM-9759

ornl

OAK RIDGE
NATIONAL
LABORATORY

MARTIN MARIETTA

O-Theory - A Hybrid Uncertainty Theory

E. M. Oblow

OPERATED BY
MARTIN MARIETTA ENERGY SYSTEMS, INC.
FOR THE UNITED STATES
DEPARTMENT OF ENERGY

Printed in the United States of America. Available from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road, Springfield, Virginia 22161
NTIS price codes—Printed Copy: A03; Microfiche A01

ORNL/TM-9759

Engineering Physics and Mathematics Division

O-THEORY - A HYBRID UNCERTAINTY THEORY

E. M. Oblow

Research sponsored by
U.S. DOE Office of
Basic Energy Sciences

Date Published: October 1985

Prepared by the
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831
operated by
Martin Marietta Energy Systems, Inc.
for the
U.S. DEPARTMENT OF ENERGY
under contract No. DE-AC05-84OR21400

TABLE OF CONTENTS

O-THEORY - A HYBRID UNCERTAINTY THEORY

E. M. OBLOW

ABSTRACT

A hybrid uncertainty theory is developed to bridge the gap between fuzzy set theory and Bayesian inference theory. Its basis is the Dempster-Shafer formalism (a probability-like, set-theoretic approach), which is extended and expanded upon so as to include a complete set of basic operations for manipulating uncertainties in approximate reasoning. The new theory, operator-belief theory (OT), retains the probabilistic flavor of Bayesian inference but includes the potential for defining a wider range of operators like those found in fuzzy set theory.

The basic operations defined for OT in this paper include those for: dominance and order, union, intersection, complement and general mappings. A formal relationship between the membership function in fuzzy set theory and the upper probability function in the Dempster-Shafer formalism is also developed. Several sample problems in logical inference are worked out to illustrate the results derived from this new approach as well as to compare them with the other theories currently being used. A general method of extending the theory using the historical development of fuzzy set theory as an example is suggested.

O-THEORY - A HYBRID UNCERTAINTY THEORY

E. M. OBLOW

1) Introduction

The problem of dealing with uncertainty in inference and reasoning processes is a complex and difficult one. The information available for reasoning is often uncertain, imprecise, and even vague. Approximate means of dealing with the propagation of such data through inference models is crucial to the success of any machine intelligence program. Although no complete solution to this problem is at hand, several different approaches have been pursued. The use of classical Bayesian inference theory (BIT) for instance, is one such approach to this problem which has yielded some success[1,2]. More recently, several attempts derived from set-theoretic formalisms have provided other insights into its solution. These latter methods are represented by Zadeh's fuzzy set theory[3] (FST) and Dempster-Shafer belief theory[4-6] (DST). An excellent unified review of all three of these theories is presented by Prade[7]. Suitable background material for this paper can be found in this latter review article and the extensive list of references cited therein. Preliminary presentations of this paper also suitable for background material appear in Weisbin[8].

In the present article, a different approach to the uncertainty problem will be developed. The motivation behind attempting to develop another approach in this area can better be understood by taking a closer look at the strengths and weaknesses of each of the three uncertainty methodologies just mentioned. For instance, BIT has a strong, well established probability basis but is weak in its applicability to problems which are formulated in set-theoretic language. FST, on the other hand, has a strong and highly developed set theory background but its basic membership function and set operators are less intuitive and physical than those using probability concepts. In between these two extremes lies DST, which has both a probability and set theory basis. Its strongest point is the capability of representing such concepts as noncommital and vacuous belief. On the other hand, however, it lacks the extensive mathematical developments necessary for more general applicability. In addition to these observations, it should be noted that each of these theories gives quantitatively different results in application. Bayes' theorem[1,4], Dempsters' combination rule[4] and fuzzy set rules for union and intersection[3], the basic laws of BIT, DST and FST, respectively, all are quantitatively different ways of combining uncertainty information. Only in certain limiting circumstances do their results all tend to converge[4,7].

In practice then, it should be clear that all three theories (and possibly some new ones) will probably find extensive use in solving the inference-uncertainty problem. Satisfaction with inferential results, computational efficiency, and ease of representation will be the final measures of success of any of these methodologies in any given problem area.

In this paper, an attempt will be made to bridge the gap between FST and BIT using the set-theoretic strengths of the former and the probability basis of the latter. The focal point of this new approach will be the DST, which will be extended so that additional mathematical operators can be used to propagate uncertainties in a wider range of problems. This hopefully will eliminate one of DST's perceived weaknesses compared with FST. The overall approach taken to achieve this synthesis will be to use the available mathematical developments in FST as a model for extending DST while retaining the probabilistic flavor it has in common with BIT. The resulting theory proposed is, therefore, a hybrid-uncertainty theory using the strengths of both FST and BIT on an enhanced DST base.

2) Basis for O-Theory

The basic starting point for the development of operator-belief theory (OT) is DST. A brief outline of the primitive concepts of this theory needed for OT are given here. The reader is referred to Shafer[4] for more details. To begin, use is made of the set of possibilities θ, with elements $x_i \in \theta$, and its power set 2^θ, with elements $x \in 2^\theta$. As in DST, a basic probability mass, $m_\theta(x)$, is assigned to each $x \in 2^\theta$, with the function $m: 2^\theta \rightarrow [0,1]$, which is normalized by

$$\sum_{x \in 2^\theta} m_\theta(x) = 1 \ . \tag{1}$$

Here, x is a set which is a subset of θ (i.e. $x \subset \theta$) and is, therefore, also an element of 2^θ (i.e. $x \in 2^\theta$).

This normalized mass distribution defined on 2^θ is the 'uncertainty' or 'belief' representation of θ, which will be denoted by $\underline{\theta}$ and referred to as the 'belief set of θ'. Any proper subset of θ which will have a mass distribution assigned to it will be denoted by an underlined capital letter (e.g. the set $A \subset \theta$ with mass assignment $m_A(x)$ will be denoted by $\underline{A}$). The set θ will generally be used to represent the largest, finite possibility set under consideration and, therefore, $\underline{\theta}$ will represent the largest belief set.

At this point it should be noted that the normalization given in Eq.(1) represents a departure from DST, in that the assignment of zero mass to the null element of 2^θ (i.e. $m(\emptyset)=0$) is not taken to be such by definition. Mass can be assigned directly to $\emptyset$ or it can be acquired as a result of operations on the set θ. This modification is proposed to allow for the possibility that the original set θ might not have been a complete enumeration of all the possible states of the system under investigation to which mass could be assigned. It also allows $\emptyset$ to be an element of 2^θ into which conflicting information can be gathered to represent incompleteness in θ.

Note here, that the use of a normalized mass distribution for assignment of mass to elements of the possibility set 2^{θ} as given in DST, represents an extension of the concepts of probability theory, and OT, therefore, has this extended probability basis as well. This normalization can also be interpreted in a set theory context in a fashion which ties it to FST. That is, the normalization represents the maximum effective cardinality of the possibility set θ, which is unity. If masses are treated like membership functions, this means that at most only one member (and possibly none if $m_{\theta}(\phi)=1$) can be the true possibility set member.

Two other constructs from DST will also be used in OT. These are the upper and lower probabilities denoted here as P and B, respectively. Their definitions[4], slightly modified, are given for $\forall x,x' \epsilon 2^{\theta}$, as follows:

$$B(x) = \begin{cases} \sum\limits_{x' \subset x \neq \phi} m_{\theta}(x') & x \neq \phi \\ m_{\theta}(\phi) & x = \phi \end{cases}, \tag{2}$$

$$P(x) = 1 - B(\tilde{x}), \tag{3}$$

where $\tilde{x}$ is the complement of x in θ and because of the normalization condition given in Eq.(1), we see that $B(\theta)=1-m(\phi)$, another departure from DST. In this new form we also see that, in addition to its original DST interpretations, $B(\theta)$ can now be used as a measure of the effective cardinality of $\underline{\theta}$.

With these definitions, the basic strengths of DST, in being able to assign mass to any element of the power set of 2^{θ} and the ability to have an amount of belief remain uncommited to any particular element of 2^{θ} (i.e. $P(x)-B(x)\geqslant 0$), are therefore retained in OT.

Further developments of this new theory will now be made by extending this basic framework using analogies derived from the mathematical operators and structures available in FST. In particular a basic set of algebraic operators like the union, intersection, and complement will be proposed first, structural relationships will follow, so that order, dominance and equality can be defined, and finally, a norm will be introduced.

3) Structural Relations and Norm

In order to develop the tools necessary for comparing the uncertainties in various possibility sets, some dominance, order and size relations must first be established. For the order and dominance relations, this is not as easy a task to accomplish as it was for FST. That is, the analogy to set inclusion can not be used in OT, since the probability masses represent a distribution over the power set 2^{θ} and one normalized distribution is not easily included in another. In this case then, the concept of the moment of the distribution was used instead to define an order.

Defining the cardinality of a set x to be $|x|$ (i.e. the number of elements in the set), the dominance of any one member x of a belief set $\underline{\theta}$ over any other member x' can be defined in terms of a cardinality moment as

$$x \succcurlyeq x' , \quad \text{if} \quad m_{\theta}(x)|x| \geq m_{\theta}(x')|x'| , \qquad (4)$$

where $\succ$ represents dominance, $x, x' \in 2^{\theta}$ and the masses of x and x' in 2^{θ} are $m_{\theta}(x)$ and $m_{\theta}(x')$ respectively.

In this same vein, the dominance of one belief set $\underline{A}$ over another $\underline{B}$ where both have the same common power set 2^{θ}, is defined by

$$\underline{A} \succcurlyeq \underline{B} , \quad \text{if} \quad \sum_{x \in 2^{\theta}} m_A(x)|x| \geq \sum_{x \in 2^{\theta}} m_B(x)|x| , \qquad (5)$$

where $m_A(x)$ and $m_B(x)$ are the mass assignments of $\underline{A}$ and $\underline{B}$, respectively, in 2^{Θ}. Equality can also be defined similarly as

$$\underset{\sim}{A} = \underset{\sim}{B}\ , \quad \text{iff} \quad m_A(x) = m_B(x) \quad \text{for} \quad \forall x \in 2^{\Theta}. \qquad (6)$$

As defined, these relationships set up a partial order in $\underset{\sim}{\Theta}$ between various power set mass distributions. The more diffuse the information content of a power set, the more dominant it is in the order. In the case of all the mass being assigned to only a single element of 2^{Θ}, two particularly useful belief sets: $\underset{\sim}{E}$, with $m(\Theta)=1$, and $\underset{\sim}{N}$ with $m(\emptyset)=1$ can be defined. With these two new belief sets, it can be seen that, in an uncertainty context, this order is bounded by them, in that, $\inf(\underset{\sim}{\Theta})=\underset{\sim}{N}$ and $\sup(\underset{\sim}{\Theta})=\underset{\sim}{E}$. Also in this case of all mass being assigned to an individual element of 2^{Θ}, the concept of set inclusion is a limiting case of this order, in that, if $x' \subseteq x$ then $x \succeq x'$.

Finally, the moment sums appearing in Eq.(5) can be normalized to unity by dividing by the cardinality of the possibility set to define the concept of a size or norm, / /, as

$$/\underset{\sim}{\Theta}/ = \sum_{x \in 2^{\Theta}} m_{\Theta}(x)\,|x|/|\Theta| \ . \qquad (7)$$

The limits of this norm are seen to be: $/\underset{\sim}{\Theta}/_{max}=/\underset{\sim}{E}/=1$ and $/\underset{\sim}{\Theta}/_{min}=\underset{\sim}{N}=0$.

4) Intersection and Union Operations

The most important operators needed for any uncertainty algebra are those that allow information from various sources to effectively be combined. In DST, Dempster's rule of combination[4] is the only operator available for pooling uncertainty information. It has very strong intuitive appeal, in that it is based on both a probabilistic and set-theoretic approach. The proportionate distribution of mass between possibilities using mass products, which lies at the heart of this scheme, is a fundamental rule of combination in probability theory[4]. Set theory operations are used, on the other hand, to assign the resulting mass products to each member of the possibility power set. These two strong points make this rule the best choice for the first fundamental operator of OT, that is, the intersection operator $\textcircled{\wedge}$. The definition of the intersection operator for the case $\underline{C} = \underline{A} \textcircled{\wedge} \underline{B}$, where $\underline{A}$, $\underline{B}$, and $\underline{C}$ have power sets 2^{θ} with elements $a,b,c\epsilon 2^{\theta}$ and masses $m_A(a)$, $m_B(b)$ and $m_C(c)$ respectively, is therefore, given as

$$m_C(c) = \sum_{a \cap b = c} m_A(a) m_B(b) \quad . \tag{8}$$

where it is easily shown[4] that $0 \leqslant m_C(c) \leqslant 1$ and unit normalization of masses in $\underline{C}$ is retained.

In this scheme, the mass in any element of $\underline{C}$ (i.e. $m_C(c)$) is given by the sum of all the mass products in which the elements a of $\underline{A}$ and b of $\underline{B}$ intersect in c. This is the essence of Dempster's combination rule except for the fact that mass is allowed to fall into the null set $\emptyset$ if $a \cap b = \emptyset$. In DST these resulting masses are renormalized into all the other elements of $\underline{C}$ such that $m_C(\emptyset) = 0$. The advantage of retaining the masses in $\emptyset$ that result from sets which have no intersection, is to have a measure of the amount of conflict which exists between the two belief sets being combined. Renormalization masks the fact that no common ground can be found to combine such information. In this context also, the mass in $\emptyset$ can also be used as measure of the incompleteness of the original possibility sets.

The choice of a suitable union operator to go along with the intersection rule above is a difficult one and involves many compromises in trying to develop the least restrictive algebra possible. When operator associativity, commutativity, unit normalization, and nonnegativity of masses are considered essential features of this theory, the choice becomes somewhat easier. In this light the mass combination rule most akin to that used in probability and group theory (i.e. $m_1+m_2-m_1*m_2$), must be rejected because it violates either one or the other of the latter two constraints mentioned above in a power set implementation. The final choice was, therefore, made using the MAX and MIN operations in FST as an analogy. That is, an upper and lower bound to the common ground for the pooling of information was formed with the union and intersection rules of set theory. This choice of the union operator $\textcircled{\vee}$, for the case $\underline{C} = \underline{A} \textcircled{\vee} \underline{B}$, was then

$$m_C(c) \quad = \sum_{a \cup b = c} m_A(a) m_B(b) \ . \tag{9}$$

where again because mass products are used, $0 \leqslant m_C(c) \leqslant 1$ and the masses retain unit normalization in $\underline{C}$.

Here, as in Eq.(8), mass products are used to distribute mass between the possibility sets to be combined, but the resulting mass, in this case, is assigned to the union of the subset of the elements being considered and not the intersection.

Use of the two basic operators $\textcircled{\wedge}$ and $\textcircled{\vee}$, together with the identity belief set $\underline{E}$ and the null belief set $\underline{N}$ gives rise to the following relationships for any belief set $\underline{\theta}$:

$$\underline{\theta} \textcircled{\wedge} \underline{E} \quad = \quad \underline{\theta} \ , \tag{10a}$$
$$\underline{\theta} \textcircled{\vee} \underline{E} \quad = \quad \underline{E} \ , \tag{10b}$$

for $\underline{E}$ defined previously as: $m_\theta(x)=0$ for $\forall x \neq \theta$ and $m_\theta(\theta)=1$, and

$$\underline{\theta} \textcircled{\wedge} \underline{N} \quad = \quad \underline{N} \ , \tag{11a}$$
$$\underline{\theta} \textcircled{\vee} \underline{N} \quad = \quad \underline{\theta} \ , \tag{11b}$$

for $\underline{N}$ defined previously as: $m_\theta(x)=0$ for $\forall x \neq \emptyset$ and $m_\theta(\emptyset)=1$.

In summarizing this section, it should be emphasized, that both the union and intersection operators were defined to use the same product rule to combine masses so as to preserve commutativity and unit normalization. It is the use of the set rules $\cap$ and $\cup$ for final placement of these masses in the resulting power set that distinguished them. In this context, the use of $\textcircled{\vee}$ and $\textcircled{\wedge}$ produces upper and lower bounds to the common ground between the two belief sets being combined. This role is similar to that played by the MAX and MIN rules in FST. The OT rules, however, are not distributive in mixed operations and idempotency is also lost. This yields a somewhat less general structural base for future developments but one that appears to be necessary, given the normalization condition which ties this theory to its probability base.

5) Definition of Complementation

The last basic operation needed to complete OT is complementation. This concept can be defined by noting that this theory deals with the power set 2^θ for mass assignments and not just the possibility set θ. In a conventional power set context, every element of the power set has a complement which is also a member of the power set. To preserve mass normalization to unity, then, the complement of a belief set $\underline{\theta}$ is defined to simply shift mass between an element in 2^θ and its complement in 2^θ so that

$$m_{\tilde{\theta}}(x) \quad = \quad m_\theta(\tilde{x}) \quad , \qquad \text{for} \quad \forall x, \tilde{x} \in 2^\theta \ . \qquad\qquad (12)$$

That is, for the belief set $\underline{\theta}$, with mass assignments $m_\theta(x)$ for $\forall x \in 2^\theta$, the complement set representation $\underline{\tilde{\theta}}$, has mass assignments $m_{\tilde{\theta}}(x)$, where $m_{\tilde{\theta}}(x) = m_\theta(\tilde{x})$ for $\forall x \in 2^\theta$.

This definition gives results similar to those derived from FST in the limit of crisp sets, but in its most general form it, like FST, does not preserve the normal set-like rules for a complement. That is,

$$\underline{\theta} \owedge \underline{\tilde{\theta}} \neq \underline{N} \quad , \tag{13}$$

$$\underline{\theta} \ovee \underline{\tilde{\theta}} \neq \underline{E} \quad . \tag{14}$$

In practical application, however, the normal set results are closely approximated, as one would want. In addition, De Morgan's laws and involution, given as follows:

$$\widetilde{\underline{A} \owedge \underline{B}} = \underline{\tilde{A}} \ovee \underline{\tilde{B}} \ , \tag{15}$$

$$\widetilde{\underline{A} \ovee \underline{B}} = \underline{\tilde{A}} \owedge \underline{\tilde{B}} \ , \tag{16}$$

$$\underline{\tilde{\tilde{A}}} = \underline{A} \ , \tag{17}$$

are obeyed in all cases.

The proof of De Morgan's laws can easily be demonstrated using either the $\owedge$ or $\ovee$ operators. In the $\owedge$ case, for example, Eq.(8) can be used to define the operation $\underline{A} \owedge \underline{B} = \underline{C}$ as

$$m_C(c) = \sum_{a \cap b = c} m_A(a) m_B(b) \ , \tag{18}$$

so that using Eq.(12), $\underline{\tilde{C}} = \widetilde{\underline{A} \owedge \underline{B}}$ is then given by

$$m_{\tilde{C}}(c) = m_C(\tilde{c}) = \sum_{a \cap b = \tilde{c}} m_A(a) m_B(b) \ . \tag{19}$$

Noting that the sets a and b are dummy variables in this equation, we can switch to a complement notation to get

$$m_{\tilde{C}}(c) \quad = \sum_{\tilde{a} \cap \tilde{b} = \tilde{c}} m_A(\tilde{a}) m_B(\tilde{b}) . \tag{20}$$

Applying De Morgan's rule to the summation index equation and Eq.(12) to the masses in the summation, yields

$$m_{\tilde{C}}(c) \quad = \sum_{a \cup b = c} m_{\tilde{A}}(a) m_{\tilde{B}}(b) \ , \tag{21}$$

which is identical to $\tilde{\underline{C}} = \tilde{\underline{A}} \circledvee \tilde{\underline{B}}$. Thus De Morgan's law in the form $\widetilde{\underline{A} \circledwedge \underline{B}} = \tilde{\underline{A}} \circledvee \tilde{\underline{B}}$, is proven. Similar manipulations prove Eq.(16).

In summary, a basic set of rules for union, intersection, and complement which can be used to manipulate power sets with mass assignments has been proposed. These rules allow operations to be performed that are not available in DST and should make further developments possible using FST as a model. Despite some similarities to FST, however, the actual results obtained with this approach will be different than those obtained by FST even in very simple cases, as is illustrated below.

6) Examples and Comparison to FST and BIT

To give some idea how the operators defined in the previous sections might be applied in an approximate reasoning problem, a simple example in logic will be worked out. This example was chosen primarily for its simplicity, but it does illustrate clearly some of the differences between OT, DST, BIT and FST. In particular, DST cannot be applied to this problem at all, for reasons to be explained, and FST produces quantitatively different results. It also has some potential applicability in expert systems, where a strict logical interpretation of implication rules with uncertainty might prove useful.

Starting with a simple, single element set $\theta=\{a\}$ and its power set $2^{\theta}=\{\phi,\theta\}$, the definition of the complement can be used to reinterpret this set in terms of the logical constants, T and F (i.e. true and false) if we let $\theta=T$ and $\phi=F$. In DST, we could go no further than this, since the consequent mass assignments would require $m_{\theta}(\phi)=0$ and therefore $m_{\theta}(\theta)=1$. Any further operations with such single element sets would leave these assignments unchanged. An approach to this logic problem in DST would, for example[4], have to start with a minimum of the two element set $\theta=\{T,F\}$. The relaxation of the normalization condition in OT, however, allows mass in ϕ resulting in workable rules for dealing with single element sets. These rules will be illustrated below using the logical interpretations of: $\owedge$ as AND, $\ovee$ as OR, $\tilde{}$ as NOT, and for the two belief sets $\underline{A}$ and $\underline{B}$, the operation $\tilde{\underline{A}} \ovee \underline{B}$ as the implication rule.

Assigning the following masses to the belief sets of $\underline{A}$ and $\underline{B}$ (noting the normalization condition in Eq.(1)):

$$\underline{A} = \{ \overset{1-m_A}{F} , \overset{m_A}{T} \} \text{ and } \underline{B} = \{ \overset{1-m_B}{F} , \overset{m_B}{T} \} , \qquad (22)$$

we can look at the consequences of operating on these sets with the logical constructs, AND, OR, NOT and implication. Note here, that only a single mass assignment is needed to completely characterize the belief sets of $\underline{A}$ or $\underline{B}$ because of the definition of the complement and the normalization condition. The results to follow will, therefore, only deal with the values of m(T) (i.e. the mass of the T power set element). The value of m(F) will always be 1-m(T).

A) AND - Equivalent to $\owedge$

In Boolean logic, the result C in the operation A AND B=C is represented by the following truth table:

		B: F	B: T
A	F	F	F
	T	F	T

Table 1. Boolean truth table for AND operator.

This table represents the limiting behavior of both OT, BIT and FST in this simple case as will be seen below.

For the OT case, using the intersection rule given in Eq.(8) and noting that the masses for T and F add up to unity, we find that the logical AND result, represented by $\underline{A} \oslash \underline{B} = \underline{C}$, is given by

$$\underline{A} \oslash \underline{B} = \underline{C} = \{ \overset{1-m_A m_B}{F} , \overset{m_A m_B}{T} \} . \tag{23}$$

That is,

$$m_C(T) = m_A m_B . \tag{24}$$

This is the same result that would be obtained using BIT when multiplying probabilities p_A for A={T} and p_B for B={T} to get A*B=C. That is,

$$p_C = p_A p_B . \tag{25}$$

The equivalent result in FST is $\underset{\sim}{A} \cap \underset{\sim}{B} = \underset{\sim}{C}$ for the single element fuzzy sets $\underset{\sim}{A}$={T} and $\underset{\sim}{B}$={T}, with membership functions μ_A and μ_B, respectively. That is, the MIN operator in FST is used and

$$\mu_C = MIN(\mu_A, \mu_B) \quad . \qquad (26)$$

Both the OT results for $m_C(T)$, which are also the BIT results for p_C, and the FST results for $\mu_C(T)$ are represented graphically in the form of continuous truth tables in Fig.(1).

As can be seen from this figure, the OT, BIT and FST results are functions representing the Boolean truth table given in Table 1. They all have precisely the same limits as the Boolean results (i.e. as μ, p or m approach 0 or 1), and have qualitatively similar behavior, although the OT and BIT results are smooth functions, in general, and the FST results are only piecewise continuous.

Also evident from this example, is the fact that in OT, with continued application of the $\textcircled{\wedge}$ operator in a sequence of AND operations, the masses in the result will approach a limit in which $m_\theta(\emptyset)=1$ and $m_\theta(\theta)=0$. This result is typical of OT results after application of many $\textcircled{\wedge}$ operations and reflects its probability basis. In BIT (i.e. probability theory), the repeated compounding of probabilities results in monotonically decreasing results and this is precisely what the OT results are duplicating. FST, on the other hand, will always be limited in this compounding effect by the smallest value of the membership function in the series; it acts like a set operator, as opposed to a probability operator. Also evident in this simple example is the almost identical quantitative and qualitative roles played by the mass and membership functions. This gives an indication that these two concepts can be related more rigorously.

B) OR - Equivalent to $\textcircled{\vee}$

The results for the logical OR can be worked out in a similar fashion to those presented above. In Boolean logic the OR operation is represented again by a truth table, which in this case is

ORNL-DWG 85-16199

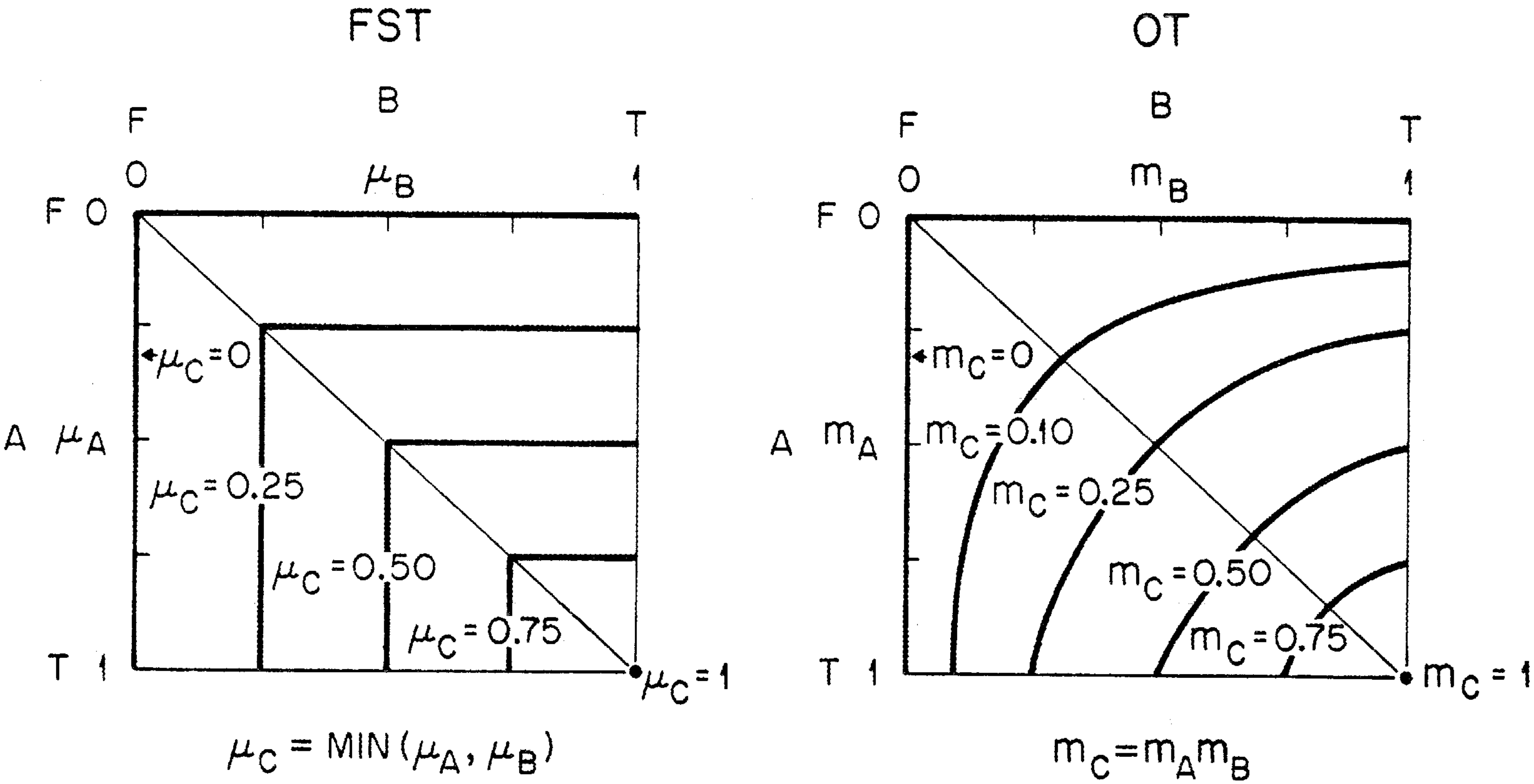

Fig. 1. Truth tables for the AND operator in OT and FST example.

A \ B	F	T
F	F	T
T	T	T

Table 2. Boolean truth table for the OR operator.

In OT, using Eq.(9) for the union rule gives three mass products for $m_C(T)$. This result, rewritten using only the complement of the single mass product result for $m_C(F)$ is

$$\underline{A} \circledvee \underline{B} = \underline{C} = \left\{ \overset{(1-m_A)(1-m_B)}{F} \;,\; \overset{1-(1-m_A)(1-m_B)}{T} \right\} . \qquad (27)$$

That is, the three mass products give

$$m_C(T) = m_A m_B + m_A(1-m_B) + m_B(1-m_A) = 1-(1-m_A)(1-m_B) \ . \qquad (28)$$

In BIT, the equivalent of the logical OR case depends on the addition law for probabilities p_A and p_B. This law preserves normalization to unity, and for this case is A+B=C. The BIT result is then

$$p_C = p_A + p_B - p_A p_B \ , \qquad (29)$$

which can also be seen to be equal to

$$p_C = 1 - (1-p_A)(1-p_B) \ . \qquad (30)$$

The equivalence of the OT and BIT results in this case is rather interesting, in that the OT union combination rule uses only mass products to obtain its result. OT could not accommodate the use of a probability addition rule like Eq.(29) in its general formulation because of normalization constraints (see the discussion in Section 4). The same result, however, is generated by using the set union operation to store three mass products in the power set element T, which then add up to the probability rule result.

By way of comparison again, the equivalent FST results are derived from $\underset{\sim}{A} \cup \underset{\sim}{B} = \underset{\sim}{C}$, using the MAX operator. That is,

$$\mu_C = \mathrm{MAX}(\mu_A, \mu_B) \quad . \tag{31}$$

There is no need to graph the results in this case (or for that matter those for the implication rule which follow), since they are simply rotations or inversions of the same general shapes given in Fig.(1). They all have the same limiting behavior as the Boolean truth tables and the OT (and BIT) results are the smooth analogs of the piecewise continuous FST results.

Again, a typical result of OT in general, is obtained in this case with repeated application of the $\bigcirc\!\!\!\!\vee$ operator. When this is done, the limits are seen to be the reverse of those seen in the intersection case, in that now we get $m_\theta(\phi)=0$ and $m_\theta(\theta)=1$. This is the same limit expected from probability theory for the addition of probabilities, as is evident from the equivalence of BIT and OT in this case. FST produces no such limit due to the nature of the MAX operation.

C) Implication - Equivalent to $\tilde{\underline{A}} \bigcirc\!\!\!\!\vee \underline{B}$

As the last part of this example, the strict logical interpretation of implication as $\tilde{\underline{A}} \bigcirc\!\!\!\!\vee \underline{B}$, will be used to highlight the three theories. Although this is a rather simplistic interpretation of implication (compared to what might be done otherwise in OT and what has already been proposed in FST[7]), it is rather instructive and computationally efficient. Current use of certainty factors in expert systems[1] is certainly on a par with this interpretation as far as computational ease is concerned, although the evidential basis of certainty factors is far more theoretically developed.

As in the last subsection, a Boolean truth table representing the limits of the continuous theories for $\tilde{A} \cup B = C$ is given by

		B: F	B: T
A	F	T	T
	T	F	T

Table 3. Boolean truth table for the implication operator

Application of Eqs.(9) and (12) in OT to the implication rule definition gives the following results:

$$\tilde{\underline{A}} \circledvee \underline{B} = \underline{C} = \{ \overset{m_A(1-m_B)}{F} , \overset{1-m_A(1-m_B)}{T} \} . \qquad (32)$$

The value of $m_C(T)$ is a result again, in this example, of three mass products, which when simplified give

$$m_C(T) = 1 - m_A(1-m_B) . \qquad (33)$$

Noting that the complement of p_A in BIT for this case is $1-p_A$, the BIT results, using the Eq.(29) as the rule for addition of probabilities, are again seen to be equivalent to the OT results. That is, for $\tilde{A}+B=C$,

$$p_C = 1-p_A+p_B-(1-p_A)p_B = 1-p_A(1-p_B) . \qquad (34)$$

In FST, the equivalent result is obtained by taking $\tilde{\underset{\sim}{A}} \cup \underset{\sim}{B} = \underset{\sim}{C}$, where the complement uses a membership function of $1-\mu_A$ for T. This result is simply

$$\mu_C \quad = \quad MAX(1-\mu_A, \mu_B) \quad . \tag{35}$$

A graphical representation of these results are again a rotation of the behavior shown in Fig.(1). All the comments made for the union case apply here as well.

Summarizing the results obtained in this simple example, it should be clear that OT and BIT are strongly connected in a probability context but differ in the way they obtain similar results. In this sense, OT uses both arithmetic and set theory operations while BIT uses arithmetic laws only. The qualitative similarities between OT and FST are also apparent, in that the former is a smooth analog of the latter.

7) General Extensions to OT

A) Connection to FST

In order to make it easier to develop additional operators and concepts in OT, it is useful to make some formal connection between the mass, m, and the membership function, μ. The examples in the last section indicate how these functions are related in a simple case and lead to the belief that a more general relationship can be found. Noting the role played by P, the upper probability in DST (defined in Eq.(3)) and the membership function μ in FST, it was felt that a formal connection could be made between these two concepts.

In DST, the function P(x) can be interpreted as the maximum possible belief in the member x of the power set 2^{θ}. Its range for any element x is always, $0 \leqslant P \leqslant 1$, even though the masses themselves must always sum to unity. The membership function, likewise, represents a possibility (i.e. for membership) and also has the same range of values. The difference between the two, in this regard, is only that μ is defined on θ and P is defined on its power set 2^{θ}. This suggests that a formal connection can be made between μ and P by restricting P, for this discussion, to the elemental members of θ; the defining relationship is then

$$\mu(x_i) = P(x_i) = 1 - B(\tilde{x}_i) = \begin{cases} 1 - \sum\limits_{x' \subset \tilde{x}_i \neq \emptyset} m_\theta(x') & x_i \neq \emptyset \\ 1 - m_\theta(\emptyset) & x_i = \emptyset \end{cases} \qquad (36)$$

for $\forall x_i, \tilde{x}_i \in \theta$ and $\forall x' \in 2^{\theta}$.

Using the definition of B(x) given in Eq.(2), the same restriction (i.e. $x_i \in \theta$) can be made, giving a relationship between $\tilde{\mu}(x_i)$ and $B(\tilde{x}_i)$ which is similarly seen to be

$$\tilde{\mu}(x_i) = 1 - \mu(x_i) = B(\tilde{x}_i) \ . \qquad (37)$$

Although no unique inverse relationship can be postulated between masses m(x) and membership functions $\mu(x)$ using Eq.(36), the above relationships do provide a useful way of comparing theoretical developments and results between OT (and for that matter DST) and FST.

B) Extension to mappings

As an example of the process of extending the range of applicability of OT, the definition of a general mapping rule with uncertainty will be proposed here. As with most of the developments which will be derived from this theory, an analogous path to the extension of FST will be taken. That is, set theory rules will be expanded with the role of membership functions being played by masses.

Looking first at the definition of a general set mapping rule $f: a \rightarrow b$, with $a, b \subset \theta$. We note that, if $a_i \in a$ and $b_j \in b$, then the mapping f gives $b_j \in f\{a_i\}$. In general then, $f\{a_i\}$ is a set with cardinality greater than unity and the mapping is characteristically an element-to-set mapping. If we now define b_a to be the union of the elements in the set $f\{a_i\}$ (the union of f being denoted by $\cup f$), we see that $b_a \epsilon 2^B$. One generalization of this mapping for OT power sets with mass assignments is now suggested (although this choice is certainly not unique).

Thus, define the general belief set mapping $F:\underline{A}\rightarrow\underline{B}$, with $\underline{A}$ and $\underline{B}$ having elements $a\epsilon 2^A$, $b\epsilon 2^B$ and masses $m_A(a)$ and $m_B(b)$, respectively, such that, for each a mapped into b, if $a\epsilon 2^A$ and $b\epsilon 2^B$, then $b\epsilon F\{a\}$. Now, letting the union of the elements of $F\{a\}$ be b_a, as before, we see that $b_a \epsilon 2^B$ and this particular function can be used to obtain the mass assignments for this mapping. That is, in assigning mass for $F:\underline{A}\rightarrow\underline{B}$, use the function $\{b_a\} = \cup F\{a\}$ so that the mass $m_B(b)$ in 2^B can be defined as

$$m_B(b) = \begin{cases} \sum\limits_{b_a = b} m_A(a) & \forall b_a \epsilon 2^B \\ 0 & \text{otherwise} \end{cases}, \qquad (38)$$

with the following mapping and function definitions:

$$F : 2^A \rightarrow 2^B, \quad \{b_a\} = \cup F\{a\} \quad \text{and} \quad a\epsilon 2^A, \; b, b_a \epsilon 2^B \; . \qquad (39)$$

Since b_a is represented by the relationship in Eq.(39), it is also clear that

$$\sum_{b\epsilon 2^B} m_B(b) = \sum_{b\epsilon 2^B} \sum_{b_a = b} m_A(a) = 1 , \qquad (40)$$

and the final mass distribution in 2^B is normalized to unity as it was in 2^A.

In essence, the OT mapping rule replaces the element-to-set nature of the general set mapping rule in Θ with an element-to-element function in 2^Θ. This route was taken so that a normal set mapping rule could be used directly in OT without modification of its definition to distribute set masses appropriately. The rule-of-thumb used here was: collect mass into a set common to all the elements mapped into if no further information was available from the mapping definition to do otherwise. This decision converts mappings to functions, preserves normalization to unity and gives rise to a larger class of mappings in OT which will have unique inverses. The generalized rule should also provide another means of attacking the problem of uncertainty propagation in expert systems, in that the implication rule can alternatively be though of as a general mapping (i.e element-to-set). Future development in OT will be required, however, to bear this conjecture out.

Before concluding, a simple example of an OT mapping will be given to make the concept easier to understand in practice. For the case of $\theta=\{x_1,x_2\}$, the particular mapping example: $F:\underline{A}\text{->}\underline{B}$ with $\underline{A}$ and $\underline{B}$ having elements in 2^θ and respective mass assignments $m_A(x)$ and $m_B(x)$, will be represented in 2^θ by the following figure:

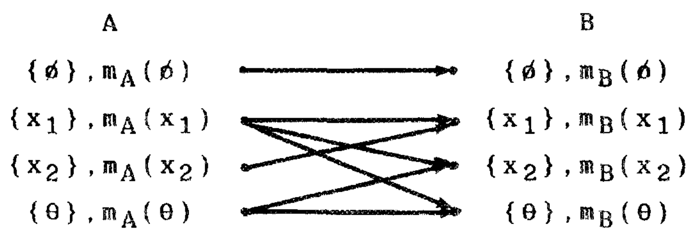

Figure 2. OT representation of the mapping $F:2^A\text{->}2^B$.

and the element mapping rule

$$F\{\emptyset\}=\{\emptyset\},\ F\{x_1\}=\{x_1,x_2,\theta\},\ F\{x_2\}=\{x_1\},\ F\{\theta\}=\{x_2,\theta\} \quad . \qquad (41)$$

Using the OT mapping rule given in Eqs.(38) and (39), this mapping gives rise to the particular function b_a, in which the following relationships hold:

$$
\begin{aligned}
b_{\emptyset} &= \{\emptyset\} = \cup F\{\emptyset\}, \quad b_1 = \{\theta\} = \cup F\{x_1\} \quad , \\
b_2 &= \{x_1\} = \cup F\{x_2\}, \quad b_\theta\{\theta\} = \cup F\{\theta\} \quad .
\end{aligned} \tag{42}
$$

If b_a is used now to distribute masses, the following figure can be used to represent this function and its final results:

$$
\begin{tikzcd}
A & B \\
\{\emptyset\}, m_A(\emptyset) \arrow[r] & \{\emptyset\}, m_B(\emptyset) = m_A(\emptyset) \\
\{x_1\}, m_A(x_1) \arrow[ddr] & \{x_1\}, m_B(x_1) = m_A(x_2) \\
\{x_2\}, m_A(x_2) \arrow[ur] & \{x_2\}, m_B(x_2) = 0 \\
\{\theta\}, m_A(\theta) \arrow[r] & \{\theta\}, m_B(\theta) = m_A(x_1) + m_A(\theta)
\end{tikzcd}
$$

Figure 3. Final OT results for the mapping $F: 2^A \to 2^B$.

Note here that the final form of the mapping is a simple function and the mass distribution in 2^B is normalized to unity. Also, as alluded to previously, this simple case might be useful in expert systems if the set θ were chosen to be $\theta = \{T, F\}$ and the implication rule translated using such a general mapping.

8) Conclusions

A hybrid uncertainty theory has been developed to bridge the gap between fuzzy set theory and Bayesian inference theory. Its basis is the Dempster-Shafer formalism (a probability-like, set-theoretic approach), which has been extended and expanded upon so as to include additional basic operations for manipulating uncertainties in approximate reasoning. The new theory, operator-belief theory (OT), retains the probabilistic flavor of Bayesian inference but includes the potential for defining a wider range of operators like those found in fuzzy set theory.

The basic operations defined for OT in this paper include those for: dominance and order, union, intersection, complement and general mappings. A formal relationship between the membership function in fuzzy set theory and the upper probability function in the Dempster-Shafer formalism was also developed. Several sample problems in logical inference were worked out to illustrate the results derived from this new approach as well as to compare them with the other theories currently being used. A general method of extending the theory using the historical development of fuzzy set theory as an example was suggested.

Future development of OT will concentrate on devising efficient computational algorithms for its implementation in expert or rule-based system applications. The OT union and intersection rules seem to have a natural basis in matrix algebra and are highly suitable for implementation in concurrent algorithmic form on a hypercube computer. Additional work is also needed in defining suitable projection operators for making decisions on the basis of power set mass assignment results. Definitions for suitable direct addition and subtraction operators are also needed so that evidence and belief can be gathered and combined to form an initial power set mass assignment.

The theory will be extensively tested in its current form as part of the Oak Ridge National Laboratory CESAR program in robotics and machine intelligence[9]. It has applicability in this program's planning, sensor fusion, vision and expert system efforts.

ACKNOWLEDGEMENTS

The author sincerely appreciates the support and encouragement of C. R. Weisbin, J. Barhen, F. C. Maienschein, A. Zucker and our DOE sponsor O. P. Manley. This manuscript was prepared solely by the author on his PC, but thanks are always due E. S. Howe and Co. for their efforts in these matters. Special thanks go to E. L. Parsley for creating the perfect environment for producing work like this.

REFERENCES

[1] B. G. Buchanan and E. H. Shortliffe, Rule-Based Expert Systems (Addison-Wesley, California, 1984), sects-11,12.

[2] G. Polya, Mathematics of Plausible Reasoning, Vol II: Patterns of Plausible Inference, 2^{nd} ed. (Princeton University Press, N. J., 1968).

[3] A. Kaufmann, Introduction to the Theory of Fuzzy Subsets, Vol 1 (Academic Press, N. Y., 1975).

[4] G. Shafer, A Mathematical Theory of Evidence (Princeton University Press, N. J., 1976).

[5] J. D. Lowrance and T. D. Garvey, Evidential reasoning: A developing concept, in Proc. IEEE Int. Conf. Cybern. Soc., 1982.

[6] ibid [1], chapter 13.

[7] H. Prade, A computational approach to approximate and plausible reasoning with applications to expert systems, IEEE Trans. Pattern Anal. Mach. Intel., Vol. PAMI-7, May 1983.

[8] C. R. Weisbin, et al., Machine intelligence for robotics applications, in Proc. Conf. Intel. Sys. & Mach., Oakland University, Minn., April 22-24,1985.

[9] CESAR - Center for Engineering Systems Advanced Research, Oak Ridge National Laboratory, Oak Ridge, Tenn. (see [8] for details of program).

INTERNAL DISTRIBUTION

1- 5. L. S. Abbott
6. S. M. Babcock
7-11. J. Barhen
12. G. de Saussure
13. J. R. Einstein
14. M. C. G. Hall
15. W. R. Hamel
16. C. C. Jorgensen
17. D. W. Kammer
18. J. L. Lucius
19-23. F. C. Maienschein
24-28. E. M. Oblow
29. F. G. Perey
30. F. G. Pin
31. V. Protopopescu
32. R. E. Ricks
33-37. C. R. Weisbin
38. A. Zucker
39. P. W. Dickson, Jr. (Consultant)
40. H. Golub (Consultant)
41. D. Steiner (Consultant)
42-43. Central Research Library
44-45. Laboratory Records
46. Laboratory Records, RC
47. ORNL Y-12 Technical, Sec. Document Reference Sec.
48. ORNL Patent Office
49. EPMD Reports Office

EXTERNAL DISTRIBUTION

50. Office of Assistant Manager for Energy Research and Development, DOE-ORO, Oak Ridge, TN 37831

51. David Abraham, Analytic Disciplines, Inc., Second Floor, 1370 Piccard Drive, Rockville, MD 20850

52. Wayne Book, Department of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332

53. B. G. Buchanan, Computer Science Department, Stanford University, Stanford, CA 94305

54. D. Dubois, Université Paul Sabatier, Laboratoire Langages et Systemes Informatiques, 118 route de Narbonne, 31062 TOULOUSE CEDEX - FRANCE

55. K. A. Farry, Flight Dynamics Laboratory, Bldg. 146, AFWAL/FIGC/58683, Wright-Patterson Air Force Base, Ohio 45433

56. R. R. Gawronski, University of West Florida, Systems Sciences Department, Pensacola, FL 32514-003

57. T. D. Garvey, SRI International, 333 Ravenswood Ave., Menlo Park, CA 94025

58. R. Haralick, Machine Vision International, Burlington Center, 325 East Eisenhower, Ann Arbor, MI 48104

59. Ewald Heer, 5329 Crown Avenue, LaCanada, CA 91011

60. F. Holley, Human Engineering Laboratory, Attn: ANXHE-FH (Holley), Aberdeen Proving Grounds, MD 21005

61. Ramesh Jain, The University of Michigan, Department of Electrical Engineering and Computer Science, Ann Arbor, MI 48109-1109

62. M. H. Kalos, AEC Computing & Applied Mathematics Center, Courant Institute of Math Science, New York University, 251 Mercer Street, New York, NY 10012

63. J. D. Lowrance, SRI International, 333 Ravenswood Ave., Menlo Park, CA 94025

64. O. P. Manley, Division of Engineering, Mathematical and Geosciences, Office of Basic Energy Sciences, U.S. Department of Energy - Germantown, Washington, DC 20545

65. Alexander Meystel, Drexel University, ECE, Philadelphia, PA 19104

66. J. F. Palmer, NCUBE Corporation, 915 E. La Vieve Lane, Tempe, AZ 85284

67. J. Pearl, Computer Science Department, University of California at Los Angeles, 405 Hilgard Ave. Los Angeles, CA 90024

68. A. P. Pentland, Artificial Intelligence Center, SRI International, 333 Ravenswood Ave., Menlo Park, CA 94025

69. H. E. Pople, Decision Systems Laboratory, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261

70. H. Prade, Université Paul Sabatier, Laboratoire Langages et Systemes Informatiques, 118 route de Narbonne, 31062 TOULOUSE CEDEX - FRANCE

71. M. Rabins, Mechanical Engineering Department, Wayne State University, Detroit, MI 48202

72. Z. Ras, Computer Science Department, University of Tennessee, 8 Ayres Hall, Knoxville, TN 37946

73. Karl N. Reid, Oklahoma State University, Stillwater, OK 74074

74. E. Sandewall, Department of Computer and Information Science, Linköping University, Linköping, Sweden

75. G. Saridis, Electrical, Computer, and Systems Engineering Department, Rensselaer Polytechnic Institute, 15th Street, Troy, NY 12180

76. G. Shafer, Department of Mathematics, University of Kansas, Lawrence, KS 66045

77. R. Sobek, Laboratoire d'Automatique et d'Analyse des Systemes du C.N.R.S. 7, avenue du Colonel-Roche, F-31077 Toulouse CEDEX, FRANCE

78. T. R. Thompson, Artificial Intelligence Project, System Planning Corporation, 1500 Wilson Boulevard, Arlington, VA 22209

79. P. H. Winston, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, Cambridge, MA, 02139

80. R. Yager, Machine Intelligence Institute, Iona College, New Rochelle, New York 10801

81. L. Zadeh, Computer Science Division, University of California, Berkeley, CA 94720

82. M. Zemankova, Computer Science Department, University of Tennessee, 8 Ayres Hall, Knoxville, TN 37946

83-109. Technical Information Center (TIC)