ON FUZZINESS AND LINGUISTIC PROBABILITIES

by

Hung T. Nguyen

Memorandum No. ERL-M595

14 June 1976

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

ON FUZZINESS AND LINGUISTIC PROBABILITIES

Hung T. Nguyen

Department of Electrical Engineering and Computer Sciences
and the Electronics Research Laboratory
University of California, Berkeley, California 94720

1. Introduction

In a recent paper [12], L. A. Zadeh introduced the concept of linguistic probabilities. In this note, we examine this concept in greater detail and explore its relations with fuzziness in measure theory, and the application of the extension principle of Zadeh [12] to operations on the space of probability distribution functions. As a preliminary, we shall review some of the basic concepts in the theory of fuzzy sets which are relevant to our analysis.

2. Fuzzy Sets and the Extension Principle

2.1 Notation

Let U be a set. We denote by $\mathcal{P}(U)$ the collection of all subsets of U. If $A \in \mathcal{P}(U)$, then 1_A is the membership function of A, defined by:

$$1_A : U \to \{0,1\}$$

$$1_A(u) = \begin{cases} 1 & \text{if } u \in A \\ 0 & \text{if } u \notin A \end{cases} \tag{2.11}$$

We denote by $\underset{\sim}{\mathcal{P}}(U)$ the collection of all fuzzy subsets of U. Each $A \in \underset{\sim}{\mathcal{P}}(U)$ is characterized by its membership function $\mu_A : U \to [0,1]$, and

Research supported in part by Army Research Office Grant DAHC04-75-G0056.

we write symbolically

$$A = \int_U \mu_A(u)/u \quad , \tag{2.12}$$

signifying that A is the union of fuzzy singletons $\mu_A(u)/u$.

An order relation in $\underset{\sim}{\mathcal{P}}(U)$ is defined by:

$$A \subseteq B \Leftrightarrow \mu_A \le \mu_B \tag{2.13}$$

If $A, B \in \mathcal{P}(U)$, then we define:

$$A \cup B \quad \Leftrightarrow \quad \mu_{A \cup B} = \mu_A \vee \mu_B \tag{2.14}$$

$$A \cap B \quad \Leftrightarrow \quad \mu_{A \cap B} = \mu_A \wedge \mu_B \tag{2.15}$$

where $\vee$ and $\wedge$ stand for maximum and minimum respectively. With these operations, $\underset{\sim}{\mathcal{P}}(U)$ is a (complete) distributive lattice with minimal element ϕ ($\mu_\phi \equiv 0$), and maximal element U ($\mu_U \equiv 1$).

2.2 Fuzzy Negation

Since the unit interval [0,1] is a distributive lattice, which is not complemented, it follows that $\underset{\sim}{\mathcal{P}}(U)$ is not complemented. If we define the relative pseudo-complement of a fuzzy set A relative to a fuzzy set B, we obtain a Brouwerian lattice, as shown in [2]. But, in fuzzy logic, the complement or, equivalently, the fuzzy negation of a fuzzy set A, denoted by A^c, is defined by:

$$\mu_{A^c} \equiv 1 - \mu_A \ . \tag{2.21}$$

Although obvious, it is important to note that A^c is not a complement or a pseudo-complement of A in the algebraic sense. However, the operation $A \to A^c$ satisfies the following two properties of the complement operation in an arbitrary distributive lattice (with maximal and minimal elements):

$$(A^c)^c = A \text{ (involution)} \tag{2.22}$$

$$A \subseteq B \Rightarrow A^c \supseteq B^c \text{ (order reversing)} \tag{2.23}$$

Thus, the definition of fuzzy negation is motivated by the fact that the unit interval [0,1] is a complete, distributive lattice with order reversing involution: $x \to 1 - x$.

2.3 Reverse and Dual

If A and B are two subsets of the real line $\mathbb{R}$ (or more generally, of $\mathbb{R}^n$), then

$$A \ominus B = \{a - b \mid a \in A,\ b \in B\} \tag{2.31}$$

Now if A and B are fuzzy subsets of $\mathbb{R}$, then it is natural to define $A \ominus B$ as the fuzzy subset of $\mathbb{R}$ characterized by:

$$\mu_{A\ominus B}(x) = \bigvee_{\begin{cases}(u,v) \in \mathbb{R}^2 \\ u - v = x\end{cases}} [\mu_A(u) \wedge \mu_B(v)] \tag{2.32}$$

In particular, if $A = \{t\}$, then

$$\mu_{t\ominus B}(x) = \mu_B(t-x) \tag{2.33}$$

which implies that $t \ominus B$ is the fuzzy symmetrical set of B with respect to the point t/2.

Let U be a bounded invertal in $\mathbb{R}$, say $U = [a,b]$, and $A \in \underset{\sim}{\mathcal{P}}[a,b]$; then by the reverse of A we mean the fuzzy set A* defined by:

$$\mu_{A*}(u) = \mu_A(a+b-u) \quad , \quad \forall u \in [a,b] \tag{2.34}$$

i.e.

$$A^* = (a+b) \ominus A \tag{2.35}$$

Note that in the case of $\mathbb{R}$, we define A* as:

$$\mu_{A*}(u) = \mu_A(-u) \quad , \quad \forall u \in \mathbb{R} \tag{2.36}$$

In the case of U = [0.1] (and, more particularly, in application to linguistic truth-values and linguistic probabilities [12]), we have:

$$\mu_{A*}(u) = \mu_A(1-u) \quad , \quad u \in [0,1] \tag{2.37}$$

In this case, A* is called the dual of A [1]. Note that $A^* \neq A^c$, but the operation $A \to A^*$ is also an ordered reversing involution (with respect to the extended order relation of $\underset{\sim}{\mathcal{P}}[0,1]$, via the extension principle [12], as will be described in greater detail in Sec. 2.5).

2.4 Fuzzy Relations

If $U = U_1 \times \ldots \times U_n$, then an n-ary fuzzy relation in U is a fuzzy subset of U.

If R (resp. S) is a (binary) fuzzy relation in $X \times Y$ (resp. $Y \times Z$), then the (max-min) composition of R and S is the fuzzy relation $R \circ S$ in $X \times Z$ characterized by:

$$\mu_{R\circ S}(x,z) = \bigvee_{y\in Y} [\mu_R(x,y) \wedge \mu_S(y,a)] \tag{2.41}$$

Note that if R and S are nonfuzzy relations, then:

$$R \circ S = \{(x,z) \in X \times Z \mid \exists\, y \in Y \text{ such that } (x,y) \in R \text{ and } (y,z) \in S\} \tag{2.42}$$

Note also that $R \circ S$ can be expressed in terms of fuzzy projections as follows:

Let $\bar{R}$ and $\bar{S}$ be fuzzy subsets of $X \times Y \times Z$ characterized by:

$$\mu_{\bar{R}}(x,y,z) = \mu_R(x,y) \quad (2.43)$$

$$\mu_{\bar{S}}(x,y,z) = \mu_S(y,z) \quad (2.44)$$

The projection T of $\bar{R} \cap \bar{S}$ on $X \times Z$, is defined by:

$$\mu_T(x,z) = \bigvee_{y \in Y} \mu_{\bar{R} \cap \bar{S}}(x,y,z) \quad (2.45)$$

then we have:

$$R \circ S = T = \Pr_{X \times Z}(\bar{R} \cap \bar{S}).$$

2.5 Extension Principle

The theory of fuzzy sets provides a basis for the development of computation techniques for the manipulation of magnitudes which are expressed in linguistic rather than numerical terms. A device that is particularly useful for this purpose is the extension principle [12], which provides a natural way of extending operations defined on U to $\tilde{\mathcal{P}}(U)$.

More specifically, let f be a mapping from U to V, and let $A \in \tilde{\mathcal{P}}(U)$, with A expressed symbolically as:

$$A = \int_U \mu_A(u)/u \quad (2.51)$$

where μ_A is the membership function of A.

Then the image of A, f(A), is defined to be:

$$f(A) = \int_V \mu_A(u)/f(u) \tag{2.52}$$

where $f(u) \in V$.

More generally, if $*$ is a binary operation in $U \times V$, with values in W, and $A \in \underset{\sim}{\mathcal{P}}(U)$, $B \in \underset{\sim}{\mathcal{P}}(V)$, then

$$A * B = \int_W \mu_A(u) \wedge \mu_B(v)/u*v \tag{2.53}$$

For example, if $A, B \in \underset{\sim}{\mathcal{P}}(\mathbb{R})$; then

$$A + B = \int_{\mathbb{R}} \mu_A(x) \wedge \mu_B(y)/x+y \tag{2.54}$$

The extension principle may be used to extend the usual order relation on the real line $\mathbb{R}$ to $\underset{\sim}{\mathcal{P}}(\mathbb{R})$ [12]. This order relation on $\underset{\sim}{\mathcal{P}}(\mathbb{R})$ extends also the usual order relation in interval analysis [5].

A concept which plays a basic role in the applications of the extension principle is that of a fuzzy restriction [12]. Informally, if X is a variable taking values in U, then a fuzzy restriction, $R(X)$, associated with X is a fuzzy relation in U which acts as an elastic constraint on the values that may be assigned to X. Thus, if <u>small</u> is a fuzzy subset of $\mathbb{R}$ characterized by the membership function μ_{small} : $\mathbb{R} \to [0,1]$, then the fuzzy proposition "X is small" translated into

$$R(X) = \text{small}$$

which implies that the proposition in question induces a fuzzy restriction on the values of X which is given by <u>small</u>. In this sense, a variable X is fuzzy if it is associated with a fuzzy restriction.

In the case of fuzzy variables, a concept which is analogous to that of dependence in the case of random variables, is the concept of interaction [12]. More specifically, the fuzzy variables $X_1,\ldots,X_n$ are non-interactive under the restriction $R(X_1,\ldots,X_n)$ iff

$$R(X_1,\ldots,X_n) = R(X_1) \times \ldots \times R(X_n) \tag{2.55}$$

where X denotes the cartesian product of fuzzy sets [12].

As an illustration, consider the propositions: "X_1 is A" and "X_2 is B" where $A \in \tilde{\mathcal{P}}(U)$, $B \in \tilde{\mathcal{P}}(V)$.

These propositions translate into:

$$R[a(X_1)] = A \quad , \quad R[b(X_2)] = B \tag{2.56}$$

where $a(X_1)$ and $b(X_2)$ are implied attributes of X_1 and X_2 respectively. Thus, X_1 and X_2 are non-interactive iff:

$$R(a(X_1),b(X_2)) = A \times B \tag{2.57}$$

i.e.

$$\mu_{R(a(X_1),b(X_2))}(u,v) = \mu_A(u) \wedge \mu_B(v) \tag{2.58}$$

<u>Remark</u>: In general, for each $(u,v) \in U \times V$, the degree of compatibility of (u,v) with $R(a(X_1),b(X_2))$ is a function of $\mu_A(u)$ and $\mu_B(v)$, say $f(\mu_A(u),\mu_B(v))$. Thus the non-interaction of X_1 and X_2 implies that:

$$f(\mu_A(u),\mu_B(v)) = \mu_A(u) \wedge \mu_B(v). \tag{2.59}$$

It is shown in [1] that the non-interaction of X_1 and X_2 corresponds to the property of non-compensation of the function f. More precisely, a function $f : [0,1]^2 \to [0,1]$ is said to have the non-compensation property if:

for all $\alpha \in [0,1]$, there do not exist $(x,y) \in [0,1]^2$ such that:

(i) $x \wedge y < \alpha < x \vee y$

(ii) $f(x,y) = f(\alpha,\alpha)$.

<u>Examples</u>: $f(x,y) = x \wedge y$, $f(x,y) = x \vee y$

Note that the functions $f(x,y) = xy$, $f(x,y) = \frac{x+y}{2}$ do not have this property. As shown in [1], under suitable conditions on f, if f has the non-compensation property, then f is necessarily of the form

$$f(x,y) = x \wedge y.$$

2.6 <u>Operations on the Space of Probability Distribution Functions</u>

Recall that a probabilistic metric space is an ordered triple $(S,\mathcal{F},\tau)$ such that [8]:

$\mathcal{F}: S \times S \to \Delta^+$ (space of probability distribution functions F such that $F(0) = 0$)

(i) $\mathcal{F}(p,q) = \varepsilon_0 \Leftrightarrow p = q$ (where $\varepsilon_0 = 1_{]0,+\infty[}$.)

(ii) $\mathcal{F}(p,q) = \mathcal{F}(q,p)$

(iii) $\mathcal{F}(p,r) \geq \tau(\mathcal{F}(p,q),\mathcal{F}(q,r))$ where τ is a triangle function (a suitable binary operation on Δ^+)

If τ is the convolution operation, then (iii) is Wald's inequality.

If τ is of the form:

$$\tau_T((F,G)(x) = \sup_{u+v=x} F(F(u),G(v)) \tag{2.61}$$

where T is a t-norm, then (iii) is Menger's inequality. It is shown in [9] that the operation $\tau_\wedge$ is derivable from an operation on random variables, namely, addition, using the technique of copulas. A similar result can be obtained from the fuzzy set point of view, using the extension principle as follows.

Given F, G $\in \Delta^+$, formally F and G define two fuzzy subsets A and B of the real line by putting

$$F = \mu_A \quad , \quad G = \mu_B .$$

Using the extension principle, addition of fuzzy sets A and B is defined as:

$$A + B = \int_{\mathbb{R}} \mu_A(x) \wedge \mu_B(y) / x+y \tag{2.62}$$

By the definition of the union of fuzzy sets, we have:

$$\mu_{A+B}(z) = \bigvee_{\begin{cases}(x,y)\in\mathbb{R}^2 \\ x+y=z\end{cases}} [\mu_A(x) \wedge \mu_B(y)]. \tag{2.63}$$

i.e.

$$\tau_\wedge(F,G)(z) = \mu_{A+B}(z) \tag{2.64}$$

Thus, the operation $\tau_\wedge$ corresponds to the addition of fuzzy subsets of the real line.

Remark: If X and Y are real random variables, with distributions F_X, F_Y respectively, then their joint distribution H_{XY} is of the form:

$$H_{XY}(x,y) = C_{XY}(F_X(x), F_Y(y)) \tag{2.65}$$

where C_{XY} is a connecting copula of X and Y [9]. If X and Y are independent, then $C_{XY}(x,y) = x \cdot y$.

If X and Y are fuzzy variables, with fuzzy restrictions R(X), R(Y) respectively, then X and Y are non-interactive iff:

$$f[\mu_{R(X)}(x), \mu_{R(Y)}(y)] = \mu_{R(X,Y)}(x,y) = \mu_{R(X)}(x) \wedge \mu_{R(Y)}(y). \tag{2.66}$$

Let X and Y be two real random variables such that $X = f(Y)$ or $Y = f(X)$, with f strictly increasing. Consider X and Y as fuzzy variables with fuzzy restrictions:

$$\mu_{R(X)} \equiv F_X \quad , \quad \mu_{R(Y)} \equiv F_Y \tag{2.67}$$

Then the fuzzy variables X and Y are non-interactive iff the fuzzy restriction R(X,Y) is characterized by:

$$\mu_{R(X,Y)} \equiv H_{XY} \tag{2.68}$$

because here we have: $C_{XY}(x,y) = x \wedge y$ [9].

3. Fuzziness in Measure Theory

3.1 The concept of a measure may be extended to fuzzy sets in a variety of ways. In particular, in [10] and [13], a fuzzy measure is defined as a mapping from a collection of fuzzy subsets to the real line which has most of the basic properties of an ordinary measure. More specifically, a measure m on a σ-algebra $\mathcal{A}$ of fuzzy subsets of a set Ω is a mapping from $\mathcal{A}$ to $\overline{\mathbb{R}^+}$ such that:

(i) $A,B \in \mathcal{A}$ and $A \subseteq B \Rightarrow m(A) \leq m(B)$

(ii) $m(\phi) = 0$

(iii) $\forall A,B \in \mathcal{A}$, $m(A) + m(B) = m(A \cup B) + m(A \cap B)$

(iv) If A_n is an increasing sequence in $\mathcal{A}$, then:

$$m(\bigcup_n A_n) = \lim_{n \to \infty} m(A_n).$$

It can readily be shown that, for fuzzy measure, monotone continuity is stronger than σ-additivity.

3.2 An alternative definition of fuzzy measure is given in [11] as follows:

Let $(\Omega,\mathcal{A})$ be a measurable space. A fuzzy measure of $(\Omega,\mathcal{A})$ is a mapping $g : \mathcal{A} \to [0,1]$ such that:

(i) $g(\phi) = 0,\ g(\Omega) = 1$

(ii) $A, B \in \mathcal{A}$ and $A \subseteq B \Rightarrow g(A) \leq g(B)$

(iii) $(A_n)_1^{+\infty}$ in $\mathcal{A}$ and monotone, then: $\lim_{n\to\infty} g(A_n) = g(\lim_{n\to\infty} A_n)$

In this definition, the value $g(A)$ may be interpreted as a subjective measure expressing the grade of fuzziness of the set A.

A fuzzy measure g is said to be F-additive if $\forall A,B \in \mathcal{A}$, $g(A \cup B) = g(A) \vee g(B)$.

These fuzzy measures bear a resemblance to a special class of capacities, namely strong capacities [6].

Examples: The Dirac measure at a point $\omega_0 \in \Omega : \delta_{\omega_0}$.
If $f : \Omega \to [0,1]$, the $A \to \sup_A f(\omega)$ is such a measure. Formally, a probability measure is also a fuzzy measure in the above sense.

3.3 Let $(\Omega,\mathcal{A})$ be a measurable space. By a fuzzy event, we mean a fuzzy subset A of Ω such that its membership function μ_A is $\mathcal{A}$-measurable. Probability measure is extended to fuzzy events in [13] as follows:

Let P be a probability measure on $(\Omega,\mathcal{A})$. Denote by $\underset{\sim}{\mathcal{A}}$ the σ-algebra of fuzzy events. Then P is extended to $\underset{\sim}{\mathcal{A}}$ by putting:

$$A \in \underset{\sim}{\mathcal{A}}, \quad \hat{P}(A) = \int_\Omega \mu_A(\omega)\,dP(\omega)$$

Thus: $\hat{P} : \underset{\sim}{\mathcal{A}} \to [0,1]$.

As a simple illustration of this extension, let $\Omega = \{\omega_1, \ldots, \omega_n\}$ be a finite set with the uniform probability measure P. Let $A \in \underset{\sim}{\mathcal{P}}(\Omega)$, then the fuzzy proportion $\frac{|A|}{|\Omega|}$ is defined by

$$\frac{|A|}{|\Omega|} = \frac{\sum_{i=1}^{n} \mu_A(\omega_i)}{n}$$

Remark: The number $|A| = \sum_{i=1}^{n} \mu_A(\omega_i)$ is called the power of A[12]. More generally, if Ω is an arbitrary set, the power of $A \in \tilde{\mathcal{P}}(\Omega)$ may be viewed as the result of an extension of the counting measure on Ω. More precisely, the power of $A \in \tilde{\mathcal{P}}(\Omega)$, is given by $\sum_{\omega \in S_A} \mu_A(\omega)$ is S_A is countable and $+\infty$ if not, where S_A is the support of A, i.e., $S_A = \{\omega : \mu_A(\omega) \neq 0\}$.

3.4 In generalized information theory [4], the information associated with an event A is interpreted as the information provided by a proposition (in the spirit of statistical mechanics) of the form "The state is in the set A." A natural way of extending this concept of generalized information to fuzzy sets is the following [7]:

Recall that an information measure is a mapping J from a lattice X ($\leq, \wedge, \vee$, minimal element 0, maximal element 1) to $\overline{\mathbb{R}^+}$ such that:

(i) $x \leq y \Rightarrow J(x) \geq J(y)$.

(ii) $J(0) = +\infty$, $J(1) = 0$

(iii) There exists $F : \overline{\mathbb{R}^+} \times \overline{\mathbb{R}^+} \to \overline{\mathbb{R}^+}$ such that:

$$J(x \vee y) = F(J(x), J(y)) \text{ if } x \wedge y = 0.$$

Examples:

a) Let $(\Omega, \mathcal{A}, P)$ be a probability space. The Wiener-Shannon information measure on $X = \mathcal{A}$ is defined by:

$$A \in \mathcal{A}, \quad J(A) = c \text{ Log } \frac{1}{P(A)} \quad , \quad c > 0.$$

$$F(x,y) = \mathrm{Sup}\{0, -c \,\mathrm{Log}\,[e^{-\frac{x}{c}} + e^{-\frac{x}{c}}]\}.$$

b) Let m be the counting measure on an infinite set Ω. Define $J : X = \mathcal{P}(\Omega) \to \overline{\mathbb{R}^+}$ by:

$$J(A) = \frac{1}{m(A)}$$

$$F(x,y) = \frac{1}{1/x + 1/y}$$

c) Let D be the Hausdorff dimension on [0,1], and $\theta : [0,1] \to \overline{\mathbb{R}^+}$, continuous, strictly decreasing with $\theta(0) = +\infty$ and $\theta(1) = 0$.

Define J on $X = \mathcal{P}([0,1])$ by

$$J(A) = \theta[D(A)].$$

$F(x,y) = x \wedge y$ (This is so because, for all $A,B \subseteq [0,1]$, $D(A \cup B) = D(A) \vee D(B)$.)

Given a measurable space $(\Omega, \mathcal{A})$, we denote $\underset{\sim}{\mathcal{A}}$ the lattice of fuzzy subsets (of Ω) A such that μ_A is $\mathcal{A}$-measurable. Let J be an information measure on $\mathcal{A}$, then an information measure for fuzzy events, i.e. $\underset{\sim}{\mathcal{A}}$, can be defined as an information measure $\hat{J}$ on the functional lattice $\mathcal{F}(\underset{\sim}{\mathcal{A}}) = \{\mu_A : A \in \underset{\sim}{\mathcal{A}}\}$, such that:

$$\hat{J}(1_A) = J(A) \text{ for all } A \in \mathcal{A}.$$

<u>Example</u>:

Let $\Omega = [0,1]$, and $(f_t)_{t\in[0,1]}$ a family of functions $f_t : [0,1] \to [0,1]$, such that:

$$f_t(t) = 1 \quad , \quad \forall t \in [0,1].$$

For example:

For $t \in \,]0,1[$, $f_t(x) = \begin{cases} \frac{x}{t} & \text{if } 0 \leq x \leq t \\ \frac{1-x}{1-t} & \text{if } t \leq x \leq 1 \end{cases}$

and: $f_0(x) = 1 - x$

$f_1(x) = x.$

The mapping: $A \in \underset{\sim}{\mathcal{P}}[0,1] \to f_A \in [0,1]^{[0,1]}$ defined by:

$$f_A(x) = \sup_{t \in A} f_t(x) \tag{3.41}$$

is such that:

(i) $f_\phi \equiv 0$

(ii) $f_A = 1$ on A.

(iii) $A \subseteq B \Rightarrow f_A \leq f_B$.

Thus this mapping induces a θ-closure operator [3] (left-continuous) on [0,1]:

$$\theta(A,\alpha) = \{x : f_A(x) \geq \alpha\} \tag{3.42}$$

We denote by $\mathcal{B}(\theta)$ the σ-algebra generated by θ (i.e. by $\theta(A,\alpha)$, $A \subseteq [0,1]$, $\alpha \in [0,1]$.). Let J be an information measure on $([0,1],\mathcal{B}(\theta))$, with $F(x,y) = x \wedge y$.

For A in $\underset{\sim}{\mathcal{B}}(\theta)$, define:

$$\hat{J}(\mu_A) = \operatorname*{Inf}_{\alpha \in]0,1]} [\alpha \vee J(A_\alpha)] \tag{3.43}$$

where $A_\alpha = \{x : \mu_A(x) \geq \alpha\} \in \mathcal{B}(\theta)$.

Then $\hat{J}$ is an information measure on $\mathcal{F}(\underset{\sim}{\mathcal{B}}(\theta))$ with $\hat{F}(x,y) = x \wedge y$ and $\hat{J}(1_A) = J(A)$, $\forall A \in \mathcal{B}(\theta)$.

4. The Concept of Linguistic Probabilities

As in the case of the conventional concept of measure, the fuzzy measures defined in the preceding section take real numbers as their values. The notion of linguistic probabilities, on the other hand, represents a substantial departure from this convention. In what follows, we shall examine this notion in great detail and discuss the problem of the computation of the expectation of a linguistic random variable.

4.1 Let $(\Omega,\mathcal{A},P)$ be a probability space. If $A \in \mathcal{A}$ and $P(A)$ is unknown, we consider $P(A)$ as a numerical variable taking values in [0,1]. The linguistic variable associated with A, denoted by $\hat{P}(A)$, has $P(A)$ as its base variable; that is, each value assigned to $\hat{P}(A)$ represents a fuzzy restriction on $P(A)$. Thus, with each event A we associate a parameter taking values in a subclass of $\underset{\sim}{\mathcal{P}}[0,1]$, which implies that $\hat{P}(A)$ takes values in a (countable) set $W \subseteq \underset{\sim}{\mathcal{P}}[0,1]$, where each element of W is a fuzzy subset of [0,1] whose label belongs to a term-set T(P) [12]; $\hat{P}(A) = I$, $I \in W$, signifies that I is a fuzzy restriction on $P(A)$, and I is called a linguistic probability value. In effect, this means that $\hat{P}$ may be viewed as a multi-valued mapping from $\mathcal{A}$ to W.

Let X be a random variable, taking values in a finite set $U = \{u_1,\ldots,u_n\} \subseteq \mathbb{R}$. With each event $(X=u_i)$, we associate a linguistic variable $\hat{P}_i$, with $p_i = P(X=u_i)$ as base variable. Each n-uple $(A_1,\ldots,A_n)$, where A_i is a linguistic value assigned to $\hat{P}_i$, constitutes a linguistic probability assignment list associated with X. A collection of such lists will be referred to as a linguistic probability distribution of X. A random variable X which is associated with a linguistic probability assignment list is called a linguistic random variable. Note also that each list $(A_1,\ldots,A_n)$ may be subject to different constraints [see 4.2]. As an example, the linguistic uniform assignment list of X may be expressed as

$$A_i = \frac{\text{"1"}}{n} \quad , \quad i = 1,\ldots,n$$

where $\frac{\text{"1"}}{n}$ is a fuzzy subset of [0,1] labeled "close to $\frac{1}{n}$."

4.2 Let X be a linguistic random variable taking values in $U = \{u_1,\ldots,u_n\} \subseteq \mathbb{R}$, with linguistic probability assignment list $(A_1,\ldots,A_n)$. By linguistic expectation of X (with respect to this list), we mean the expression:

$$\hat{E}(X) = u_1A_1 + \ldots + u_nA_n \; . \tag{4.21}$$

The meaning of (4.21) may be deduced from the extension principle. Specifically, note first that each fuzzy subset A_i of [0,1] is a fuzzy restriction on a variable p_i. For this reason, the interaction between the A_i is related to the existence of constraints on the base variables p_i, which are expressed by

$$S \subseteq [0,1]^n \quad , \quad S = \{(p_1,\ldots,p_n) \,|\, p_1+\ldots+p_n = 1\} \tag{4.22}$$

If A_i are non-interactive (apart from the constraint (4.22)) then the restriction imposed by $(A_1,\ldots,A_n)$ is characterized by:

$$\mu_{(A_1,\ldots,A_n)}(x_1,\ldots,x_n) = \left[\bigwedge_{i=1}^{n} \mu_{A_i}(x_i)\right] \cap 1_S(x_1,\ldots,x_n) \; . \tag{4.23}$$

Thus, by using the extension principle, we see that $\hat{E}(X)$ is a fuzzy subset of the real line characterized by:

$$\mu_{\hat{E}(X)}(z) = \bigvee_{\begin{cases}(x_1,\ldots,x_n) \in [0,1]^n \\ x_1+\ldots+x_n = 1 \\ u_1x_1+\ldots+u_nx_n = z\end{cases}} \left[\bigwedge_{i=1}^{n} \mu_{A_i}(x_i)\right] \tag{4.24}$$

Remark:

More precisely, the notion of linguistic expectation might be formulated in the setting of measure theory as follows:

Recall that the Minkowski's addition of two subsets A and B of $\mathbb{R}$ (or more generally of $\mathbb{R}^n$) is defined by:

$$A \oplus B = \{a + b \mid a \in A,\ b \in B\} . \tag{4.25}$$

If A and B are fuzzy subsets of $\mathbb{R}$, then we define the sum $A \oplus B$ by:

$$\mu_{A \oplus B}(x) = \bigvee_{\begin{cases}(u,v) \in \mathbb{R}^2 \\ u + v = x\end{cases}} [\mu_A(u) \wedge \mu_B(v)] , \ \forall\, x \in \mathbb{R} . \tag{4.26}$$

On the other hand, for $\ell \in \mathbb{R}$ and $\ell \neq 0$, we define the fuzzy set $\ell.A$ by:

$$M_{\ell A}(x) = \mu_A\left(\frac{x}{\ell}\right) \quad , \quad \forall x \in \mathbb{R} \tag{4.27}$$

and we put $0.A = \{0\}$.

Let $u_i (i=1,\ldots,n)$ be real numbers ($\neq 0$), and A_i $(i=1,\ldots,n)$ be fuzzy subsets of $[0,1]$. We extend each A_i to a fuzzy subset of $\mathbb{R}$, still denoted by A_i, by putting:

$$\mu_{A_i}(x) = 0 \text{ if } x \notin [0,1].$$

Using (4.26) and (4.27), we can then define the fuzzy subset (of $\mathbb{R}$) $\overset{n}{\underset{i=1}{\oplus}} u_i A_i$ by:

$$\forall x \in \mathbb{R} \ , \quad \mu_{\overset{n}{\underset{i=1}{\oplus}} u_i A_i}(x) = \bigvee_{\begin{cases}(x_1,\ldots,x_n) \in \mathbb{R}^n \\ \sum_{i=1}^{n} x_i = x\end{cases}} \left[\bigwedge_{i=1}^{n} \mu_{u_i A_i}(x_i) \right] \tag{4.28}$$

$$= \bigvee_{\begin{cases} (x_1,\dots,x_n)\in\mathbb{R}^n \\ \frac{x_i}{u_i} \in [0,1],\ i=1,\dots,n \\ \sum_{i=1}^{n} x_i = x \end{cases}} \left[\bigwedge_{i=1}^{n} \mu_{A_i}\left(\frac{x_i}{u_i}\right) \right]$$

$$= \bigvee_{\begin{cases} (p_1,\dots,p_n)\in[0,1]^n \\ \sum_{i=1}^{n} u_i p_i = x \end{cases}} \left[\bigwedge_{i=1}^{n} \mu_{A_i}(p_i) \right]$$

In the case of a linguistic probability distribution, we have to add, at least, the constraint $p_1 + \dots + p_n = 1(S)$. Thus:

$$\forall x \in \mathbb{R}\ ,\ \mu_{\overset{n}{\underset{i=1}{\oplus}} u_i A_i}(x) = \bigvee_{\begin{cases} (p_1,\dots,p_n)\in S \\ \sum_{i=1}^{n} u_i p_i = x \end{cases}} \left[\bigwedge_{i=1}^{n} \mu_{A_i}(p_i) \right] \qquad (4.29)$$

4.3 It is pointed out in [12] that the computation of linguistic expectation $\hat{E}(X)$, i.e. the determination of its membership function $\mu_{\hat{E}(X)}$, reduces to the solution of a nonlinear programming problem with linear constraints.

More specifically, let the objective function, defined on $[0,1]^n$, be:

$$f(p_1,\dots,p_n) = \bigwedge_{i=1}^{n} \mu_{A_i}(p_i) \qquad (4.31)$$

then, for each given $x \in \mathbb{R}$, the determination of the value of $\mu_{\hat{E}(x)}(x)$ leads to the problem:

$$\text{Maximize } f(p_1,\ldots,p_n) \tag{4.32}$$

subject to:

$$\begin{cases} \sum_{i=1}^{n} p_i = 1 \\ \sum_{i=1}^{n} u_i p_i = x \end{cases}$$

Apart from the constraint (S) on the base variables, in general there exist (fuzzy) constraints induced by fuzzy relations between the A_i, for example if Q is a fuzzy relation in $[0,1]^2$, and A_i and A_j are such that $A_i = Q \circ A_j$. (Note that this composition reduces to the max-min matrix product when A_j and Q have finite supports.) In such a case, we are faced with the problem of optimization under fuzzy constraints (for this problem, see [14]), for example, consider a mathematical program:

$$\text{Max}\{f(x),\ x \in A\} \tag{4.33}$$

where $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, $A \subseteq D$ and the constraint is interpreted as the condition $x \in A$. We may have

$$A = \{x \in D : g(x) \geq 0\} \tag{4.34}$$

where $g : D \to \mathbb{R}^m$.

If the relation $\geq$ is replaced by a binary fuzzy relation F in $\mathbb{R}^m$, i.e. a fuzzy subset of $\mathbb{R}^{2m}$, then the feasible region A becomes a fuzzy subset of D, defined by:

$$\mu_A(x) = \mu_F(g(x),0) \quad , \quad \forall x \in D \tag{4.35}$$

Thus, in the case where the base variables p_i are constrained by a fuzzy relation, T, we have:

$$\text{Max } f(p_1,\ldots,p_n) \tag{4.36}$$

subject to:

$$\begin{cases} \sum_{i=1}^{n} p_i = 1 \\ \sum_{i=1}^{n} u_i p_i = x \\ T \end{cases}$$

i.e.

$$\text{Max } \mu_T(p_1,\ldots,p_n) \cdot f(p_1,\ldots,p_n) \tag{4.37}$$

Subject to:

$$\begin{cases} \sum_{i=1}^{n} p_i = 1 \\ \sum_{i=1}^{n} u_i p_i = x \end{cases}$$

REFERENCES

[1] R. E. Bellman and L. A. Zadeh, "Local and Fuzzy Logics," Memorandum ERL-M584, University of California, Berkeley (1976).

[2] A. De Luca and S. Termini, "Algebraic properties of fuzzy sets," Jour. Math. Anal. and Appl., 40, pp. 373-386 (1972).

[3] M. J. Frank, "Probabilistic topological spaces," Jour. Math. Anal. and Appl., 34, pp. 67-81 (1971).

[4] J. Kampé de Feŕiet and B. Forte, "Information et probabilité," C. R. Acad. Sc., Paris, 265A, pp. 110-114, pp. 142-146, pp. 350-353 (1967).

[5] R. E. Moore, Interval Analysis, Prentice-Hall (1966).

[6] H. T. Nguyen, "Sur les mesures d'Infomation de type Inf.," Lecture Notes in Math, no. 398, pp. 62-75, Springer-Verlag (1974).

[7] H. T. Nguyen, "Sur l'information et les ensembles flous." Communication at the meeting. Information Theory and Questionnaires, Lyon, France, June 1975 (to appear).

[8] B. Schweizer and A. Sklar, "Statistical metric spaces," Pacific J. Math., vol. 10, no. 1, pp. 313-334 (1960).

[9] B. Schweizer and A. Sklar, "Operations on distribution functions not derivable from operations on random variables," Studia. Math. 52, pp. 43-52 (1974).

[10] R. E. Smith, "Measure Theory on Fuzzy Sets," Thesis, Dept. of Math., Univ. of Saskatchewan, Saskatchewan, Canada (1970).

[11] M. Sugeno, "Theory of Fuzzy Integrals and its Applications," Thesis, Tokyo Inst. of Tech., Tokyo, Japan (1974).

[12] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning," Memorandum ERL, M411, Univ. of Calif., Berkeley (1973). [Published in three parts in Information Sciences

1975: 8, pp. 119-249; 8, pp. 301,357; 9, pp. 43-80.]

[13] L. A. Zadeh, "Probability measures of fuzzy events," Jour. Math. Anal. and Appl., 23, pp. 421-427 (1968).

[14] L. A. Zadeh, "Fuzzy sets and systems," Proc. Symp. on Systems Theory, Poly. Inst. of Brooklyn, pp. 29-39 (1965).