A Reflection on what is a Membership Function

E. Trillas\(^1\) and C. Alsina\(^2\)
\(^1\) Dept. Inteligencia Artificial.
\(^2\) Secció de Matemàtiques i Informàtica
Univ. Politècnica de Catalunya Barcelona. Spain.

Abstract

This paper is just a first approach to the idea that the membership function \(\mu_P\) of a fuzzy set labelled \(P\) is, basically, a measure on the set of linguistic expressions “\(x\) is \(P\)” for each \(x\) in the corresponding universe of discourse \(X\). Estimating that the meaning of \(P\) (relatively to \(X\)) is nothing else than the use of \(P\) on \(X\), these “measures” seem to be reached by generalizing to a preordered set the concept of Fuzzy Measure, introduced by M. Sugeno, when the preorder translates the primary use of the predicate \(P\), that is a basic relationship like “\(x\) is less \(P\) than \(y\)”. The paper only deals with predicates whose use is made by means of numerical characteristics, but those cases on which the characteristics (if they exist) are not of a numerical nature are not considered. By generalizing De Luca-Termini’s sharpened order some typical membership functions are studied as measures.

**Keywords:** Use of a predicate. General concept of a measure. General measures on an interval of \(R\). General measures and membership functions.

1 Introduction

A fuzzy set \(P\) on a universe of discourse \(X\) is only known once a membership function \(\mu_P : X \to [0,1]\) is fixed (see [1]), and provided that this function reflects the meaning of the statements “\(x\) is \(P\)” for all \(x\) in \(X\). This is the same as to consider that \(\mu_P\) translates the use of \(P\) on \(X\), and in this sense \(\mu_P\) is a model of this use. Nevertheless, there is certainly a lack on the comprehension of which is the link between the degree of “\(x\) is \(P\)” and the use of \(P\) on \(X\), even though in the applications of Fuzzy Set Theory this link is, in each case, clear enough.

Of course, the use of \(P\) on \(X\) not only depends on both \(X\) and \(P\) but also on the way \(P\) is related to \(X\). For example, if \(P = \text{High}\) is referred either to body temperature or to the temperature of an engine, the corresponding uses of \(P\) will differ either in the range of values of the variable “temperature” or in the degrees on which “\(x\) is \(P\)” The corresponding functions \(\mu_P\) will show different plottings with some shape similarity.
To know the actual use of $P$ on $X$ is to know the adequacy of naming as $P$ each $x$ in $X$, and, usually, such adequacy is addressed by means of a qualification of the statements "$x$ is $P$". This qualification can be obtained by an identification like:

\[ \text{Meaning of } "x \text{ is } P" : = \text{A value on some range } L, \]

a value that is to be interpreted as the degree up to which $x$ is $P$. Of course, the range $L$ should be related to the specific use of $P$. In this sense, $L$ is to be endowed with a convenient structure representing the primary use of $P$ on $X$, that is, with a binary relation translating the comparative statements "$x$ is less $P$ than $y$".

At this point, it is not to be forget that three different sets are under consideration. The ground set $X$, the set $X(P)$ of the statements "$x$ is $P$" and the range $L$. Graphically:

\[ \text{Figure 1} \]

The binary relation $\preceq$ on $L$ should reflect that it is $a \preceq b$ if and only if $a =$Degree of "$x$ is $P"$, $b =$Degree of "$y$ is $P"$ and "$x$ is less $P$ than $y"$. That is, the structure $(L, \preceq)$ should translate the primary use of $P$ on $X$: the comparison of the statements in $X(P)$. By the other side, the function $\mu_P : X \rightarrow L$ should translate to what extent each $x$ in $X$ is $P$; it should give a "measure" of such extent.

Then, a mathematical model of the use of $P$ on $X$ is given by a set $L$ endowed with a binary relation $\preceq$ and a mapping $\mu_P : X \rightarrow L$ fixed in such a way that $\mu_P(x) \preceq \mu_P(y)$ means "$x$ is less $P$ than $y". Only then the triple $(L, \preceq, \mu_P)$ can be considered as a model of the use of $P$ on $X$. But, again, it is required that $\mu_P(x) =$ Degree up to which $x$ is $P$, has the meaning of to what extent it is the case that $x$ fits $P$. The use of $P$ on $X$ is totally known once this (say) secondary use, is clear. A reasonable generalization of the concept of fuzzy measure introduced by M. Sugeno can help to clarify this problem.

Remark It should be noticed that this paper has, in some sense, an antecedent in references [2] and [3]. Notwithstanding, there are at least two main differences between both approaches. First, no measures of sets are here considered, in general.
Second, this paper refers to the use of a predicate on a given universe of discourse; it
tries to reflect upon a descriptive theory of a predicate. Of course, both approaches
are attempts to a better understanding of membership functions, and, consequently,
of fuzzy sets.

2 On a general concept of measure

What is measured when measuring are not the elements of a set $S$, but some
characteristic of these elements. For example, if $S$ consists of cubes we can measure
their surface, their volume or their weight. To measure a characteristic $k$ shown by
the elements of $S$ we need, first of all, to know a comparative relation like “$x$ shows
the characteristic $k$ less than $y$ shows it”, for any $x, y$ in $S$. Let’s write $x \preceq_k y$ to
denote that relation and suppose that $\preceq_k$ is a preorder on $S$, that is, a reflexive
and transitive relation.

Provided we can preorder $S$ in such a way, a function $m : S \to [0, 1]$ is a
$\preceq_k$-measure for $S$ whenever:

1. $m(x_o) = 0$ if $x_o \in S$ is minimal for $\preceq_k$
2. $m(x_1) = 1$ if $x_1 \in S$ is maximal for $\preceq_k$
3. If $x \preceq_k y$, then $m(x) \leq m(y)$.

Example 1 This definition generalizes the concept of Fuzzy Measure given by M.
Sugeno (see [4]). In fact, if $S$ is a lattice (with operations $\land$ and $\lor$, lowest element $0$
and greatest element $1$), it has the natural ordering given by “$x \preceq y$ if and only
if $x \land y = x$” and, then, with this partial ordering (an antisymmetric preorder), a
$\preceq$-measure is any function $m : S \to [0, 1]$ such that:

(a) $m(0) = 0$
(b) $m(1) = 1$
(c) If $x \preceq y$, then $m(x) \leq m(y)$.

Hence, Fuzzy Measures are $\preceq$-measures on the Boolean Algebras $\mathcal{P}(X)$.

Sugeno’s $\lambda$-measures (see [4]) can be defined on any Boolean Algebra (see [5])
$S$ as functions $g_\lambda : S \to [0, 1]$, with $-1 < \lambda$, such that $g_\lambda(1) = 1$ and $g_\lambda(x \lor y) =
g_\lambda(x) + g_\lambda(y) + \lambda g_\lambda(x) g_\lambda(y)$ if $x \land y = 0$. These functions are $\preceq$-measures on $S$ as:

- From $g_\lambda(x \lor x') = g_\lambda(1) = 1$ it follows $g_\lambda(x') = \frac{1 - g_\lambda(x)}{1 + \lambda g_\lambda(x)}$ and $g_\lambda(0) = g_\lambda(1') = 0$.

- If $x \preceq y$, from $x \lor y = y = x \lor (x' \land y)$ it follows

$$
g_\lambda(y) = g_\lambda(x) \left[ 1 + \lambda g_\lambda(x' \land y) + g_\lambda(x' \land y) \right]$$

$$
\geq g_\lambda(x) \left[ 1 + \lambda g_\lambda(x' \land y) \right] \geq g_\lambda(x),$$
because $-1 < \lambda$.

As any probability on a Boolean Algebra $S$ is a $\lambda$-measure with $\lambda = 0$, also probabilities are $\leq$-measures on Boolean Algebras (see [5]).

Obviously, if $m$ is a $\leq$-measure on a Complemented Lattice, its dual function, defined by $m^*(x) = 1 - m(x')$, is also a $\leq$-measure.

Example 2 If $S = F(X)$ is the set of all fuzzy sets in $X$, and $\prec$ is the sharpened order defined by:

\[ \mu \prec \sigma \text{ if and only if } 0 \leq \mu(x) \leq \sigma(x) \leq 1/2 \text{ or } 1/2 \leq \sigma(x) \leq \mu(x) \leq 1, \]

a $\leq$-measure on $F(X)$ is the well known concept of Fuzzy Entropy introduced by DeLuca and Termini (see [6]). In this case there is a greatest element (the function constantly equal to $1/2$), but there are many minimal elements (all crisp subsets of $X$).

Example 3 If $S = F(X)$ and $\leq = \leq$ is the pointwise ordering between functions, there is only one $x_0$ (namely, the membership function of $\varnothing$) and only one $x_1$ (namely, the membership function of $X$). By defining $(\mu \cup \sigma)(x) = \max(\mu(x), \sigma(x))$ for each $x$ in $X$, any function $m : F(X) \to [0,1]$ verifying $m(\varnothing) = 0, m(X) = 1$ and $m(x \lor y) = \max(m(x), m(y))$ is a $\leq$-measure on $F(X)$ as $\mu \leq \sigma$ means $\mu \cup \sigma = \sigma$, and then $m(\sigma) = m(\mu \cup \sigma) = \max(m(\mu), m(\sigma))$ implies $m(\mu) \leq m(\sigma)$. Hence, any Possibility Measure is a $\leq$-measure on $F(X)$ (see [7]).

Also Necessity Measures on $F(X)$ are $\leq$-measures on this set as they are duals of Possibility Measures.

Example 4 If $S = N = \{1, 2, \ldots \}$ is the set of natural numbers, let $\mu \in F(N)$ be the membership function of the fuzzy set labelled “Approximately even”, defined by

\[ \mu(n) = \begin{cases} 
1 & \text{if } n = 2^k \text{ for some } k = 0, 1, 2, \ldots \\
0 & \text{if } n = 2^k + 1 \text{ for some } k = 1, 2, \ldots \\
1 - \frac{1}{2^k} & \text{if } n = 2^k (2p + 1) \text{ for some } k \geq 1, p \geq 1 
\end{cases} \]

That is, $\mu$ vanishes in odd numbers, gives value 1 to all powers of 2 and, since all other numbers can be written in the form $2^k(2p + 1)$, the degree up to which these numbers are approximately even is $1 - 2^{-k}$. Obviously, $\mu$ is not a $\leq$-measure on $N$ if $\leq$ is the usual total ordering of $N$.

Let’s consider the well known Sardovskii’s ordering $\prec$ of $N$, which plays a crucial role in the study of dynamic systems (see [8]):

\[ 3 \leq 5 \leq 7 \leq \ldots \leq 23 \leq 25 \leq \ldots \leq 2^3 \cdot 3 \leq 2^3 \cdot 5 \leq \ldots \leq 2^3 \cdot 3 \leq 2^3 \cdot 5 \leq \ldots \leq 2^3 \leq 2^4 \leq 2 \leq 1, \]

that is, one lists first odd numbers except 1, followed by 2 times the odds, $2^2$ times the odds, etc., and finally one orders the decreasing powers of 2 ending in 1. In this total ordering of $N$, the minimum is 3, the maximum is 1 and function $\mu$ defined above is a $\leq$-measure. Of course, Sardovskii’s order is not the only way of ordering
In to make $\mu$ a $\approx$-measure and, among such orderings, the most interesting will be one translating "$n$ is less approximately even than $m"$, a translation that is not perfectly fitted by Sierpinski's order as, for example, $\mu(7) = 0$ and 7 is not a $\approx$-minimal.

**Remarks**

1. The concept of General Measure is related to various concepts of Measurement Theory (see [15]), where one deals with relations $W$ defined on $S$ by means of utility functions $u : S \rightarrow \mathbb{R}$ such that $xWy \iff u(x) > u(y)$.

2. It should be pointed out that a first classification of measures can be obtained by looking at how they verify axiom 3. That is, at how a measure grows. For example, as it is well known, measures of length, surface, volume or weight grow additively because all of them verify axiom 3 as a consequence of the additivity axiom:

$$\text{If } x \land y = 0, \text{ then } m(x \lor y) = m(x) + m(y).$$

This is also the case of probabilities (Sugeno's $g_0$-measures in a Boolean Algebra), but if $-1 < \lambda < 0$ $\lambda$-measures grow sub-additively as if $x \land y = 0$, it is $g_\lambda(x \lor y) \leq g_\lambda(x) + g_\lambda(y)$. If $\lambda > 0$, $\lambda$-measures grow super-additively as then $g_\lambda(x \lor y) \geq g_\lambda(x) + g_\lambda(y)$ provided $x \land y = 0$.

Possibility Measures are also sub-additive measures, since it is always $\text{Max}(r, s) \leq r + s$ if $r$ and $s$ are positive numbers.

3 **Some measures on closed non-degenerated real intervals**

Let $[a, b]$, with $a \neq b$, be a closed interval of $\mathbb{R}$, endowed with the usual total order $\leq$. From this basic totally ordered structure on $[a, b]$, some other partially ordered structures can be defined on $[a, b]$. This is the case, for example, of the sharpened order on $[0, 1]$, defined by:

$$x \preceq y \iff 0 \leq x \leq y \leq 1/2 \text{ or } 1/2 < y \leq x \leq 1.$$ 

This is obviously a partial order (an antisymmetric preorder) with greatest element $1/2$ and two minimal elements, 0 and 1.

In the same way, for any $r \in (a, b)$ we can define on $[a, b]$ the partial order:

$$x \preceq y \iff a \leq x \leq y \leq r \text{ or } r < y \leq x \leq b,$$

showing the greatest element $r$ and the two minimal elements $a$ and $b$.

Obviously, also the reverse order $\preceq^{-1}$ can be defined on $[a, b]$ by means of

$$x \preceq^{-1} y \iff y \preceq x \iff a \leq y \leq x \leq r \text{ or } r < x \leq y \leq b.$$
With this ordering, \([a, b]\) shows two maximal elements, \(a\) and \(b\), and just one minimal element \(r\).

It should be pointed out that a measure on \([a, b]\), endowed with this last order \(\preceq\), is any function \(m : [a, b] \to [0, 1]\) such that \(m(a) = m(b) = 0, m(r) = 1\) and that is non-decreasing between \(a\) and \(r\) and decreasing between \(r\) and \(b\). Graphically:

![Graph 1](image1.png)

**Figure 2**

In the same vein, if \(a < r_1 < r_2 < \ldots < r_n < b\), the \(n\)-sharpened order \(\preceq_n\) can be defined by means of:

\[
x \preceq_n y \text{ if } a \leq x \leq y \leq r_1 \text{ or } r_1 < x \leq y \leq r_2, \text{ or } r_2 < x \leq y \leq r_3, \text{ etc.}
\]

Notice that the 0-sharpened order is just the total order \(\leq\) of \([a, b]\). Graphically, with \(n = 3\):

![Graph 3](image2.png)

**Figure 3**

In the case \(n = 3\), for example, the partial order \(\preceq_3\) shows two maximal elements, \(r_1, r_3\) and three minimal elements \(a, r_2, b\). The graph of the reverse order \(\preceq_3^{-1}\) is the following:

![Graph 4](image3.png)

**Figure 4**

that shows three maximal elements, \(a, r_2, b\) and two minimal elements \(r_1, r_3\).
A measure \( m \) on \((a, b), \preceq_3\) is any function \( m : [a, b] \to [0, 1] \) such that \( m(a) = m(r_3) = m(b) = 0 \), \( m(r_1) = m(r_2) = 1 \), non-decreasing on \([a, r_1]\) and on \([r_2, r_3]\), and decreasing on \((r_1, r_2)\) and on \((r_3, b]\). Graphically:

![Graphical representation of a membership function](image1)

From now on let us suppose that \([a, b]\) is endowed either with a \( n \)-sharpened order \( \preceq_n \) or with its reverse order. Let \( m : [a, b] \to [0, 1] \) be a function such that:

1. \( m(x_1) = 1 \) if \( x_1 \) is any maximal element for the corresponding order.
2. \( m(x_0) = 0 \) if \( x_0 \) is any minimal element for the corresponding order.
3. If \( x \preceq_n y \) then \( m(x) \leq m(y) \).
3'. If \( x \preceq_n^{-1} y \) then \( m(x) \leq m(y) \).

If \( m \) verifies 1, 2 and 3, it is a measure for \(([a, b], \preceq_n)\). If \( m \) verifies 1, 2 and 3', it is a measure for \(([a, b], \preceq_n^{-1})\).

Of course, it is not necessary that the pieces of \( m \) on the sub-intervals \([a, r_1], [r_1, r_2], \ldots, [r_n, b]\) be either strictly non-decreasing or strictly decreasing. For example, on \([0, 10]\) and for \( \preceq_2 \) with \( r_1 = 2 \) and \( r_2 = 7 \), the following picture represents a measure on \([0, 10]\):

![Graphical representation of a membership function](image2)

Analogously, next picture represents a measure on \(([0, 10], \preceq_2^{-1})\):
Remark By a well-known result of Possibility Theory (see [7]), as all measures are normalized functions, they are distributions for measures of possibility \( \Pi_m \) that, restricted to the singletons of \([a, b] \), give the measure of the corresponding point: 
\( \Pi_m(\{x\}) = m(x) \) for all \( x \in [a, b] \). But there is a main difference on the ideas behind \( \Pi_m \) and \( m : \Pi_m \) is obtained thanks to both the total order of \( R \) and to the t-conorm \( \text{Max} \) (as formula \( P(\mu \cup \sigma) = \text{Max}(P(\mu), P(\sigma)) \) refers to the union of fuzzy sets given by \( (\mu \cup \sigma)(x) = \text{Max}(\mu(x), \sigma(x)) \)) and this is not the case for \( m \), which is only linked to the partial orders \( \preceq_n, \preceq_n^{-1} \) and not to any union on \( F([a, b]) \).

4 Predicates whose use on a set is \( R \)-measurable

Let \( P \) be a predicate, proper name of a property, common name or linguistic label, whose primary use on a universe of discourse \( X \) is known. We will say that the use of \( P \) on \( X \) is \( R \)-measurable if there exists a non-empty subset \( S \) of \( R \) endowed with a preorder \( \preceq \), a mapping \( \phi : X \to S \) and a \( \preceq \)-measure on \( S \) such that:

1. \( \phi(x) \leq \phi(y) \) if and only if \( x \) is less \( P \) than \( y \) for \( x, y \) in \( X \).

2. Degree up to which \( x \) is \( P = m(\phi(x)) \) for each \( x \) in \( X \).

Function \( \phi \) corresponds to a numerical characteristic of how \( P \) is used on \( X \) and, of course, function \( \mu_P : X \to [0, 1] \), given by \( \mu_P(x) := m(\phi(x)) \) for each \( x \) in \( X \), can be defined as the Compatibility Function of \( P \) on \( X \) and, consequently, as the Membership Function of the fuzzy set labelled \( P \). This function reflects the use of \( P \) on \( X \); it is a mathematical model of the use of \( P \) on \( X \).

Obviously, the set \( S \) depends both on \( P \) and \( X \). For example, let \( X \) be a set of material bodies, \( P = \text{High Temperature} \), and \( S \) the interval \([a, b] \) on which temperature of such bodies varies; in the case of human beings, \([a, b] = [33.5, 42.5] \) on the Celsius scale. Concerning the preordering \( \preceq \), it should be experimentally obtained from the use of \( P \) on \( X \). For example, if there are both a greatest value \( r_1 \) from which temperature is definitely high and a lowest value \( r_2 \) from which temperature is definitely not-high, the preorder \( \preceq \) can be taken as the \( \preceq_2 \)-sharpened order with
Finally, as measure $m$ restricted to $[a, r_1]$ is 1, and 0 when restricted to $[r_2, b]$, it only remains to find out its value on $[r_1, r_2]$ knowing that $m(r_1) = 1$ and $m(r_2) = 0$. To this end, a model of $m$ on $[r_1, r_2]$ should be hypothesized and, once determined, experimentally tested before considered as an enough good model of the use of $P$ on $X$.

After such an experimental process, a mathematical model for the use of $P$ on $X$ is obtained. Of course the model can be reached thanks to the fact that $P$ is used on $X$ by means of a numerical characteristic given by $\phi$.

Theorem. Let $X$ be a universe of discourse. The use of any crisp predicate $P$ on $X$ is $\mathcal{R}$-measurable.

Proof. As “$P$ is crisp on $X$” just means that the classical set $P = \{x \in X : \text{"$x$ is } P\text{" is true}\}$ exists, the universe $X$ is perfectly classified by the partition $\{P, P^C\}$. Then we get the function $\phi = \varphi_P$, the characteristic function of $P$, and the partial order of $[0, 1]$ given by:

$$a \leq b \text{ if and only if } (a = 0, b = 0) \text{ or } (a = 0, b = 1) \text{ or } (a = b = 1),$$

for which any $a, b$ in $[0, 1]$ such that $0 < a < 1$ and $0 < b < 1$, or $a = 1, b = 0$, are not comparable. With this (degenerated) ordering the set $[0, 1]$ shows the greatest element 1 and the lowest element 0, and the function $m = id_{[0,1]}$ is a measure for $([0,1], \leq)$ since $m(0) = 0, m(1) = 1$ and if $a \leq b$ then $m(a) = a \leq b = m(b)$. From $m(\phi(x)) = \phi(x) = \varphi_P(x) \Rightarrow$Degree up to which $x$ is $P$, it follows that $P$ is measurable. It should be pointed out that $\phi$ perfectly describes the use of $P$ on $X$ because “$x$ is $P$” is only true (degree 1) or false (degree 0).

Example 5 If $X$ is a big population on which there are people of all ages, the common name $Y = \text{Young}$ has a meaning that is known whenever it is known how $Y$ applies to any individual of $X$; that is, whenever the degree up to which “$x$ is $Y$” is known. In this case, following Zadeh’s ideas, there is in $X$ the fuzzy set $Y$ of young people that is determined by a membership function $\mu_Y$. In a strict interpretation of its use, the predicate $Y$ depends only on the age of each individual $x$ in $X$. That is, the use of $Y$ is made by means of the numerical characteristic “Age of $x$”. Hence, let $[a, b] = [0, 110]$ and let $\phi : X \to [0, 110]$ be the (non-injective) function $\phi(x) = \text{Age of } x$ (this number given with the necessary accuracy).

In $[0, 110]$ there are numbers $r_1$ and $r_2$ ($r_1 < r_2$) such that if $\phi(x) \leq r_1$ the individual $x$ is definitely young, and if $r_2 \leq \phi(x)$ then the individual $x$ is definitely not-young. Hence, the 2-sharpened order $\leq_2$ with $r_1, r_2$ reflects the primary use of $Y$ in $X$, and any function $m : [0, 110] \to [0, 1]$ which is 1 on $[0, r_1]$, 0 on $[r_2, 1]$ and decreasing on $[r_1, r_2]$, is a measure for $([0, 110], \leq_2)$. For example, if $r_1 = 25$ and $r_2 = 55$, any of such measures verify:

$$m(\phi(x)) = \begin{cases} 1 & \text{if } \phi(x) \leq 25 \\ 0 & \text{if } \phi(x) \geq 55 \\ a \text{ non-increasing function of } \phi(x) & \text{if } 25 \leq \phi(x) \leq 55 \end{cases}$$
A model for \( m \) in the interval \([25, 55]\) can be found by checking a linear, quadratic, etc., function. That is, by checking which type of function better fits the use of \( Y \) when applied to individuals between 25 and 55 years. If a quadratic function seems to fit that behaviour of \( Y \), from \( y = x^2 + bx + c \) it follows \( 1 = 25^2a + 25b + c, \ 0 = 55^2a + 55b + c, \ 0 = 110a + b \) (as \( y'' = 2ax + b \) provided that \( a > 0 \) (as \( y''' = 2a \)). Then, solving all these equations, \( a = 1/900, \ b = -110/900, \ c = 55^2/900 \) are obtained. Consequently, it results

\[
m(\phi(x)) = \left( \phi(x)^2 - 110\phi(x) + 55^2 \right) / 900 \text{ if } 25 \leq \phi(x) \leq 55.
\]

Hence, the function

\[
\mu_Y(x) = \begin{cases} 
1 & \text{if } \text{Age}(x) \leq 25 \\
0 & \text{if } \text{Age}(x) \geq 55 \\
(\text{Age}(x)^2 - 110\text{Age}(x) + 55^2) / 900 & \text{if } 25 \leq \text{Age}(x) \leq 55
\end{cases}
\]

defines the fuzzy set \( Y \) of young people of \( X \) that corresponds to the use of \( Y \) translated by both \( \varepsilon_2 \) and \( m \).

**Example 6** Let \( X = \mathbb{R} \) and \( P = \text{Approximately } p, \) for some \( p \in \mathbb{R} \). The use of this predicate is captured by saying that there are positive real numbers \( \varepsilon_1 \) and \( \varepsilon_2 \) such that the statements “\( x \) is approximately \( p \)” are true if \( x \leq p - \varepsilon_1 \) or \( x \geq p + \varepsilon_2 \), and false if \( x = p \). Consequently, we can restrict our attention to the interval \([a, b] = [p - \varepsilon_1, p + \varepsilon_2]\) endowed with the 1-sharpened order \( \preceq_1 \) with \( p \). Any function \( m: [p - \varepsilon_1, p + \varepsilon_2] \to [0, 1] \) verifying:

1. \( m(p - \varepsilon_1) = m(p + \varepsilon_2) = 0, \ m(p) = 1 \)
2. \( m \) is non-decreasing on \([p - \varepsilon_1, p]\)
3. \( m \) is non-increasing on \([p, p + \varepsilon_2], \)

is a measure on \([p - \varepsilon_1, p + \varepsilon_2], \preceq_1 \).

Hence, some function like the following will define, once determined, the fuzzy set \( P \) of the real numbers that are approximately \( p \):

![Figure 8](image-url)
Remarks

1. Usually the linguistic labels of fuzzy sets considered in the applications correspond to linguistic variables with some numerical characteristics. From now on the most basic types of membership functions can be interpreted as measures of atomic statements “$x$ is $P$”, once these statements are ordered thanks to the primary use of $P$ on $X$. Of course, statements like “$x$ is $P$” are made equivalent to statements “$\phi(x)$ is $P^*$”, with $P^*$ a predicate on $[a, b]$. In example 5, for instance, the predicate $P^*$ can be considered as “low” and “A person $x$ is young” is made equivalent to the “The age of person $x$ is low”. When this equivalence is beforehand known, the study of the use of $P$ on $X$ can be reduced to the study of the use of $P^*$ on $[a, b]$.

2. It should be pointed out once more that the final plot of $\mu_P$ does not reflect the common name $P$ but some use of this name on $X$. The authors are following the Wittgenstein’s dictum “the meaning of a word is its use in language” (see [9]).

3. The generalization to the case where $P$ is used by means of several numerical characteristics merits further studies. If, for example, $P$ is characterized by $n$ variables, intervals $[a_i, b_i] \ (1 \leq i \leq n)$ will exist, endowed, respectively, with preorders $\preceq_i \ (1 \leq i \leq n)$, functions $\phi_i : X \rightarrow [a_i, b_i]$ and measures $m_i : [a_i, b_i] \rightarrow [0, 1]$. Then, a preorder $\preceq$, depending on the preorders $\preceq_i$, should be considered on $[a_1, b_1] \times \ldots \times [a_n, b_n]$ (for example, the product $\preceq = \prod \preceq_i$), as well as a convenient measure $m$ on $([a_1, b_1] \times \ldots \times [a_n, b_n], \preceq)$ based on the corresponding measures $m_i$, to finally obtain $\mu_P$ by aggregating everything.

5 Some Conclusions and Open Questions

This paper is only a first reflection on the idea that the values of membership functions are “measures” of the atomic or elementary statements made with the linguistic label of the fuzzy set. It is suggested that this goal can be achieved by means of a generalization of the concept of Fuzzy Measure, introduced by M. Sugeno in 1974. Also, by using a generalization of the De Luca and Termini sharpened order, the typical predicates “high”, “low”, “medium”, “approximately $P$”, etc., that are among the most frequent in the applications of Fuzzy Set Theory to Technology, are analyzed. In a close future the relationship between general measures and some concepts of Measurement Theory will be considered.

On the paper’s ground there is the belief that vague predicates can be only apprehended throughout their use on the corresponding universe of discourse, and that, consequently, any model of the use of a predicate should be based both on some theoretical considerations and on some experimentation leading to accept the model as good enough.

A descriptive theory of a predicate can help to a better understanding of the evolution of the meaning of predicates in language. For example, in the setting of
Mathematics, the use of $P = \text{probable}$ on a Boolean Algebra $X$ is currently reflected by identifying the degree up to which "$x$ is $P$" with $p(x)$, for some probability $p$ defined on $X$. But this happened after a historical process that took place during around two centuries.

If, without previously considering any probability, the primary use of $P$ on $X$ is reflected by identifying "$x$ is less probable than $y$" with $x \preceq y$ (the partial order of the Boolean Algebra), then any $\preceq$-measure $m : X \to [0, 1]$ is a candidate to give the degree up to which "$x$ is $P$", for each $x$ in $X$. A lot of time of using $P$ on different boolean algebras, indeed of experimenting with $P$, conducted to accept that $m$ should grow additively and today the word "probable" is only used in this restrictive sense. At this respect three remarks are to be made. First, that the additive law "If $x \wedge y = 0$ then $\mu_p(x \vee y) = \mu_p(x) + \mu_p(y)$" is a property imposed to the measure of atomic statements "$x$ is $P$" that does not directly follow from the primary use of $P$. It is a property of the secondary use of $P$. Second, that at each random experiment more information on the concrete use of $P$ is needed to determine the probability that better fits the experiment. Third, that once a probability $p$ is fixed, the meaning of the relation "less probable than" is enlarged as $p(x) < p(y)$ is implied by (but does not imply) $x \preceq y$.

Of course, a lot of open problems remain to be studied. This is the case, for example, if the preordering $\preceq_A$ has either no maximal or no minimal elements (that is, there are not either true or false atomic statements "$x$ is $P$" but a membership function $\mu_p$ exists such that $\text{Supp} \mu_p = 1$ or $\text{Inf} \mu_p = 0$. How to define a $\preceq_A$-measure in such cases? This is a relevant question because, as it is known, only functions $\mu : X \to [0, 1]$ verifying $\text{Supp} \mu = 1$ and $\text{Inf} \mu = 0$ (see [10]) can represent non-self-contradictory fuzzy sets. And those functions, like $\preceq_A$-measures, taking the values 1 and 0 are mere particular cases. Let’s consider the example $R = \text{Random}$ on the set $\mathbb{N} = \{1, 2, \ldots\}$ of natural numbers (see [11]), applied following the rule: "A natural number is called round if it is the product of a considerable number of comparatively small prime factors".

Interpreting "considerable number" by "big number", and with $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$, "$n$ is $R$" is made equivalent to

\[
\left[ \text{"} \alpha_1 + \cdots + \alpha_r \text{ is } \text{big} \quad \text{and} \quad \frac{\alpha_1}{n} \text{ is small}, \ldots, \text{ and} \quad \frac{\alpha_r}{n} \text{ is small} \right].
\]

Accepting that "$p$ is big" (with "big" predicated in $\mathbb{N}$) is equivalent to "$\frac{1}{p}$ is small" (with "small" predicated in $[0, 1]$), then "$n$ is $R$" is made equivalent to

\[
\left[ \text{"} \frac{1}{\alpha_1 + \cdots + \alpha_r} \text{ is } S, \text{ and } \frac{\alpha_1}{n} \text{ is } S, \ldots, \text{ and } \frac{\alpha_r}{n} \text{ is } S \right] \quad (\ast)
\]

with the predicate $S = \text{Small}$ on $[0, 1]$.

Provided that $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ and $m = q_1^{\beta_1} \cdots q_s^{\beta_s}$, a primary use of predicate $R$ on $\mathbb{N}$ can be given by:

"$n$ is less $R$ than $m$" if $\alpha_1 + \cdots + \alpha_r \leq \beta_1 + \cdots + \beta_s$, $r \geq s$, $\frac{\alpha_i}{n} \geq \frac{\beta_j}{m}$ ($1 \leq j \leq s$)

that induces the preorder $\preceq_R$ on $\mathbb{N}$ given by:
A Reflection on what is a Membership Function

\[ n \preceq_R m \text{ if } \sum_{i=1}^{r} \alpha_i \leq \sum_{i=1}^{s} \beta_i, r \geq s, \frac{p_j}{n} \geq \frac{q_j}{m} (1 \leq j \leq s) \]

In \((N, \preceq_R)\) any prime number is minimal. But there are not maximal elements, as for any \( n = p_1^{\alpha_1} \ldots p_r^{\alpha_r} \) the number \( p^{r+1} \), with \( \alpha = \sum_{i=1}^{r} \alpha_i \) and \( p = \text{Max}(p_1, \ldots, p_r) \), verifies \( n \preceq_R p^{r+1} \). How to define a \( \preceq_R \)-measure on \( N \)? Let’s call a \( \preceq_R \)-measure any function \( \mu : N \rightarrow [0,1] \) such that:

1. \( \mu(p) = 0 \) if \( p \) is prime
2. \( \lim_{n \rightarrow \infty} \mu(p^n) = 1 \) if \( p \) is prime
3. If \( n \preceq_R m \), then \( \mu(n) \leq \mu(m) \).

A natural way to find out an adequate \( \mu \) for translating the degrees up to which “\( n \) is \( R \)” follows from \((*)\), an expression on which there are only two variables, \( S \) and \( \text{and} \). The predicate \( S \) is to be represented by a function \( \mu_S : [0,1] \rightarrow [0,1] \), strictly decreasing and such that \( \mu_S(0) = 1 \) and \( \mu_S(1) = 0 \). The connective \( \text{and} \) can be represented, as is usual, by a continuous t-norm \( T \). Hence, provided that \( n = p_1^{\alpha_1} \ldots p_r^{\alpha_r} \), it is:

\[
\text{Degree up to which “} n \text{ is } R \text{”} = \mu_R(n) = T \left( \mu_S \left( 1/ \sum_{i=1}^{r} \alpha_i \right), \mu_S (p_1/n), \ldots, \mu_S (p_r/n) \right).
\]

\((**)\)

Now, if \( p \) is a prime number, it is \( \mu_R(p) = T(\mu_S(1/1), \mu_S(p/p)) = T(0,0) = 0 \)， and
\[
\lim_{n \rightarrow \infty} \mu_R(p^n) = \lim_{n \rightarrow \infty} T(\mu_S(1/n), \mu_S(1/p^{n-1})) = T(\mu_S(0), \mu_S(0)) = T(1,1) = 1.
\]

Moreover, it is not difficult to show that if \( n \preceq_R m \) it follows \( \mu_R(n) \leq \mu_R(m) \). Hence, for any continuous t-norm \( T \) and any strictly decreasing function \( \mu_S \) such that \( \mu_S(0) = 1 \) and \( \mu_S(1) = 0 \), function \( \mu_R \) is a \( \preceq_R \)-measure on \( N \). Once both \( T \) and \( \mu_S \) are fixed, the corresponding triplet \((N, \preceq_R, \mu_R)\) will be a model for the use of \( R \) given by \((**)\). Of course, to fix \( T \) and \( \mu_S \), more information on both \( S \) and \( \text{and} \) is required. For example, taking \( \mu_S(x) = 1 - x \) and \( T = \text{Prod} \), it follows (see [12]):

\[
\mu_R(n) = \left( 1 - 1/ \sum_{i=1}^{r} \alpha_i \right) \prod_{j=1}^{r} (1 - p_j/n),
\]

a function such that \( \text{Inf} \mu_R = 0 \) and \( \text{Sup} \mu_R = 1 \).

Another open problem refers to the additivity of \( \preceq_R \)-measures, a concept that seems meaningful not only for Boolean Algebras but also for lattices. For example, if \( S = [a, b] \) is endowed with the usual total order of \( R \), \(([a, b], \leq) \) is a lattice with
\[ x \land y = \text{Min}(x, y) \text{ and } x \lor y = \text{Max}(x, y), \] lowest element \( a \) and greatest element \( b \). This lattice (a chain) is a De Morgan Algebra but not a Boolean Algebra, and any non-decreasing function \( m : [a, b] \rightarrow [0, 1] \) such that \( m(a) = 0 \) and \( m(b) = 1 \) is a \( \leq \)-measure on \([a, b] \). In addition, all these functions are additive \( \leq \)-measures, because 
\[ x \land y = a \text{ if } x = a \text{ or } y = a, \] and then if \( x \land y = a \) it follows \( m(x \lor y) = m(x) + m(y) \), as either \( m(x) = 0 \) or \( m(y) = 0 \). Clearly, the same reasoning holds for any \( \leq \)-measure on \([a, b]\) whose preorder \( \leq \) has \( a \) as a maximum and \( b \) as a minimum, again if \( ([a, b], \leq) \) is not a lattice, as it is in the case of crisp predicates.

Last but not least, when and how a model for the use of a compound predicate can be established by means of the respective models of its components? For example, if

- \( P \) is used on \( X \) by \( \phi_P : X \rightarrow S_P, p \leq_P \) and \( m_P \)
- \( Q \) is used on \( X \) by \( \phi_Q : X \rightarrow S_Q, p \leq_Q \) and \( m_Q \)
- \( P \& Q \) is used on \( X \) by \( \phi_{P \& Q} : X \rightarrow S_{P \& Q}, p \leq_{P \& Q} \) and \( m_{P \& Q} \)

then to know if there exists a relation between:

- \( S_{P \& Q} \) and \( S_P, S_Q \)
- \( \phi_{P \& Q} \) and \( \phi_P, \phi_Q \)
- \( \leq_{P \& Q} \) and \( \leq_P, \leq_Q \)
- \( m_{P \& Q} \) and \( m_P, m_Q \)

is a crucial question to legitimize any algebra of fuzzy sets. At this point it should be remembered that a connective \( \& \) is called functionally expressible if there is a function \( F : [0,1] \times [0,1] \rightarrow [0,1] \) such that \( \mu_{P \& Q} = F \circ (\mu_P \times \mu_Q) \). Hence, it can be said that the use of a compound predicate \( P \& Q \) is decomposable if \( m_{P \& Q} \circ \phi_{P \& Q} = F \circ (m_P \circ \phi_P \times m_Q \circ \phi_Q) \) for some convenient function \( F \).

Acknowledgment. Authors thank Michio Sugeno (Yokohama) for many suggestive discussions on the subject during his stay in Spain as a Visiting Professor at the Department of Artificial Intelligence (Technical University of Madrid).

References


