On the Analytic Formalism of the Theory of Fuzzy Sets†

RICHARD BELLMAN

Departments of Mathematics, Electrical Engineering, and Medicine
University of Southern California, Los Angeles, California

AND

MAGNUS GIERTZ

Division of Mathematics, Royal Institute of Technology
Stockholm, Sweden

ABSTRACT

The extension of standard concepts of set theory (like union, intersection, etc.) to the theory of fuzzy sets is not obvious. Several policies for answering questions of this type have been discussed, but it appears that the operations of max and min play a central role in the arithmetic of fuzzy sets. Letting \( \mu_a(x) \) denote our willingness to accept \( x \) as a member in the fuzzy set \( A \), Zadeh defines intersection \( A \cap B \) and union \( A \cup B \) of the two fuzzy sets \( A \) and \( B \) by

\[
A \cap B = \{(x; \min \{\mu_a(x), \mu_b(x)\})\}
\]

and

\[
A \cup B = \{(x; \max \{\mu_a(x), \mu_b(x)\})\}.
\]

The object of this paper is to show that these definitions are not only natural, but under quite reasonable assumptions the only ones possible.

1. INTRODUCTION

An exact description of any real physical situation is virtually impossible. This is a fact we have had to accept and adjust to. As a result, one of the major problems in description (essential to communication, decision making, and, in a broader sense, to any human activity) is to reduce the necessary imprecision to a level of relative unimportance. We must balance the needs for exactness and simplicity, and reduce complexity without oversimplification in order to match the level of detail at each step with the problem we face.

The human brain, guided by experience, has developed a more or less intuitive capacity to cope with many problems of this type. An illustration of this is our everyday use of language, where very few words (if any) have an exact meaning. Nouns (house), verbs (walk), adjectives (tall) and so on all refer to concepts with imprecise borderlines. In order to narrow down these border-

† Supported by the National Institutes of Health under Grant No. GM 16197-03.

lines we must sacrifice simplicity, and in doing so we soon obtain a level of exactness unsuited for normal speech and writing, and eventually reach a point where this form of communication becomes impossible. But the requirement for precision does increase when we concentrate on specific scientific areas, and attains its highest level when we come to the foundations of mathematics, to abstract logic and set theory. By then we have reached the point where discussions in the usual sense are impossible, where communication is limited to a very small area; but admittedly, it is very exact in this area.

The point we want to make is that this makes the foundation of mathematics, in the form of formal logic and abstract set theory, very unsuited for the type of problems we normally face. The basic assumption of set theory, that each object either belongs to or does not belong to a given set, that there is no "in between," excludes practically all "sets of real objects." To illustrate the difficulty, a statement like, "He is an old man" implies from a set theoretic point of view that there is a (well defined) set of old men, and that a certain individual is identified as a member of that set. But it is difficult, if not impossible, to give a sensible definition of something like "the set of old men." The main problem is that we find any sharp borderline dividing all men into "old" and "not old" highly artificial. There are always men who are not precisely "old" but not exactly "not old" either. The same difficulty appears whenever we try to define any (nonmathematical) set of objects with a certain common property.

In order to obtain a more precise description of the sets we encounter in real life, Zadeh extended in a simple and elegant way the concept of "set" to that of "fuzzy set." This was done by replacing the rigid "either-or" membership relation of ordinary set theory by a more flexible one, allowing a "degree of membership" for each object in the form of a number indication how willing we are (subjectively) to accept that particular object as a member.

The extension of standard concepts of set theory (like union, intersection, etc.) to the theory of fuzzy sets is not obvious. To determine the intersection of two sets, for instance, we have to decide on questions like: If an object is accepted to 60% as a member of a fuzzy set $A$, and to 40% as a member in $B$, how willing should we be to accept it as a member in both $A$ and $B$?

Several policies for answering questions of this type have been discussed [1], but it appears that the operations of max and min play a central role in the arithmetic of fuzzy sets. Letting $\mu_A(x)$ denote our willingness to accept $x$ as a member in the fuzzy set $A$, Zadeh defines intersection $A \cap B$ and union $A \cup B$ of the two fuzzy sets $A$ and $B$ by

$$A \cap B = \{(x; \min [\mu_A(x), \mu_B(x)])\},$$

and

$$A \cup B = \{(x; \max [\mu_A(x), \mu_B(x)])\}.$$
The object of this paper is to show that these definitions are not only natural, but under quite reasonable assumptions the only ones possible.

2. FUZZY SETS AND FUZZY STATEMENTS

In a fuzzy set \( A = [x, \mu_A(x)] \) the "membership function" \( \mu_A \) is real-valued and defined on some (nonfuzzy) set \( X = \{x\} \). For each \( x \) in \( X \) the number \( \mu_A(x) \) reflects our (subjective) willingness to accept this particular \( x \) as a member in \( A \). By choosing \( \mu_A(y) > \mu_A(x) \) we indicate a greater willingness to accept \( y \) than to accept \( x \) as a member in \( A \), and conversely. We may (and shall) without loss of generality limit the range of \( \mu_A \) to the closed interval \([0, 1]\), agreeing to let \( \mu_A(x) = 1 \) signify our complete acceptance of \( x \) as a member and to let \( \mu_A(x) = 0 \) mean that we reject \( x \) completely. If the range of \( \mu_A \) contains only the numbers 0 and 1, \( A \) becomes an ordinary (nonfuzzy) set.

The number \( \mu_A(x) \) has a natural interpretation as our degree of acceptance of the statement "\( x \) is a member of \( A \)," and thus as a "truth value" associated with this statement. From this point of view it is just as meaningful to ask how willing we are to accept other more complicated statements about \( A \), like "both \( x \) and \( y \) are members in \( A \)" or "either \( x \) is a member of \( A \) or else both \( y \) and \( z \) are members."

When we assign the "degrees of membership" in \( A \) we are, in general, free to use subjective evaluation. Are we still free to do so in assigning "truth values" to statements of the above type, after the membership function in \( A \) has been determined? Certainly we could not be completely free, since assigning a larger acceptance to the statement "\( x \) and \( y \) are members of \( A \)" than to "\( x \) is a member of \( A \)," would seem inconsistent. Are there, perhaps, natural laws of consistency which, with a given membership function, completely determine the "truth value" of any such compound statement?

The problem we face when defining operations like union and intersection with fuzzy sets is of the same nature. Assuming that a certain object \( x \) has given degrees of membership in two fuzzy sets, \( A \) and \( B \), it is natural to let its degree of membership in \( A \cap B \) be our level of acceptance of the statement "\( x \) is a member of \( A \) and \( x \) is a member of \( B \)." Again the question is: given two statements \( S_1 \) and \( S_2 \), with given truth values, what restrictions do we have to observe when we wish to assign a truth value to compound statements like "\( S_1 \) and \( S_2 \)" or "\( S_1 \) or \( S_2 \)"?

In order to decide on these questions we consider, then, the seemingly more general situation where we have a collection of (arbitrary) statements, each with a truth value assigned to it. In fact, this brings us back to where we started since what we now have is a fuzzy set \( \{[S, \mu(S)]\} \) where the objects \( S \) are statements and the membership function \( \mu \) is defined by their truth values.
Using the connectives "and" and "or" we may combine the given statements to form compound statements. We begin the discussion of the restrictions on the truth values of these compound statements by stating, in words, a number of conditions we feel are necessary. Then we formalize some of these conditions as a set of axioms, and finally we prove that they imply that the only possible choice is

\[
\mu(S_1 \text{ and } S_2) = \min \{\mu(S_1), \mu(S_2)\},
\]

and

\[
\mu(S_1 \text{ or } S_2) = \max \{\mu(S_1), \mu(S_2)\}.
\]

3. BASIC ASSUMPTIONS

Let \( F = \{[S, \mu(S)]\} \) be a fuzzy set of statements, and consider compound statements, like "\( S_1 \) and \( S_2 \)," "\( S_1 \) or \( S_2 \)," "\( S_1 \) and (\( S_2 \) or \( S_1 \))" and so on, formed by combining the statements in \( F \) with the connectives "and" and "or."

The membership function \( \mu \), interpreted as assigning truth values to the statements in \( F \), imposes certain natural restrictions to the truth values we may assign to such compound statements. We shall discuss these restrictions assuming, just to avoid confusion at this stage, that the statements in \( F \) are unrelated in the sense that no statement in \( F \) already is such a combination of other statements in \( F \).

Our first assumption is that the truth value of a compound statement depends on the truth values of the statements in \( F \) which have been used to form it, but not on anything else. This means that a general definition of a truth value for, say "\( S_1 \) and \( S_2 \)" requires that we specify a real-valued function \( f(x, y) \) of two real variables, with \( 0 \leq x \leq 1 \) and \( 0 \leq y \leq 1 \), to give us

\[
\mu(S_1 \text{ and } S_2) = f[\mu(S_1), \mu(S_2)].
\]

In complete analogy

\[
\mu(S_1 \text{ or } S_2) = g[\mu(S_1), \mu(S_2)],
\]

requires the specification of a similar function \( g \).

Secondly, we assume that whenever truth values have been assigned to (arbitrary) statements \( S \) and \( T \), then the same functions \( f \) and \( g \) provide us with truth values for "\( S \) and \( T \)" and "\( S \) or \( T \)," i.e.

\[
\mu(S \text{ and } T) = f[\mu(S), \mu(T)],
\]

\[
\mu(S \text{ or } T) = g[\mu(S), \mu(T)].
\]

With the above assumptions and notations we now list a number of restrictions we feel should be imposed on the functions \( f \) and \( g \):

(i) \( f \) and \( g \) are nondecreasing and continuous in both variables. Our willingness to accept "\( S \) and \( T \)" or to accept "\( S \) or \( T \)," should not
decrease if our willingness to accept $S$ (or $T$) increases. With suitably small changes in the level of acceptance of $S$ or $T$ we should be able to change the truth value of "$S$ and $T$" or of "$S$ or $T$" by an arbitrarily small amount.

(ii) $f$ and $g$ are symmetric, i.e. $f(x, y) = f(y, x)$ and $g(x, y) = g(y, x)$. There is no reason to assign different truth values to "$S$ and $T$" and "$T$ and $S$".

(iii) $f(x, x)$ and $g(x, x)$ are strictly increasing in $x$. If $\mu(S_1) = \mu(S_2) > \mu(S_4) = \mu(S_3)$ we should be more willing to accept "$S_1$ and $S_2$" than to accept "$S_3$ and $S_4$".

(iv) $f(x, y) \leq \min\{x, y\}$ and $g(x, y) \geq \max\{x, y\}$. Accepting "$S$ and $T$" requires more, and accepting "$S$ or $T$" less, than accepting $S$ or $T$ alone. Thus $\mu(S$ and $T) < \mu(S)$ and so on.

(v) $f(1, 1) = 1$ and $g(0, 0) = 0$. If both $S$ and $T$ are completely accepted, we must accept "$S$ and $T$" completely as well, and if $S$ and $T$ are both completely rejected, then we must also reject "$S$ or $T$".

(vi) Logically equivalent statements have equal truth values. To illustrate with an example, the statement

$$S_1 \text{ and } (S_2 \text{ or } S_3)$$

is logically equivalent to

$$(S_1 \text{ and } S_2) \text{ or } (S_1 \text{ and } S_3).$$

We have no reason to be more willing to accept one of these statements over the other. Condition (ii) is another example of this principle.

We shall now single out a few of these restrictions on $f$ and $g$, sufficient to prove that the conditions give us only one choice, namely $f(x, y) = \min\{x, y\}$ and $g(x, y) = \max\{x, y\}$. The functions $f$ and $g$, having, as they do, domain $[0, 1] \times [0, 1]$ and range $[0, 1]$, define two binary compositions on $[0, 1]$. This interpretation is convenient in the following discussion, and we shall use the notations

$$f(x, y) = x \land y \text{ and } g(x, y) = x \lor y.$$
and further satisfy

\[ x \land y \text{ and } x \lor y \text{ are continuous and nondecreasing in } x \text{ [see (i)]} \]
\[ x \land x \text{ and } x \lor x \text{ are strictly increasing in } x \text{ [see (iii)],} \]
\[ x \land y \leq \min \{x, y\} \text{ and } x \lor y \geq \max \{x, y\} \text{ [see (iv)],} \]
\[ 1 \land 1 = 1 \text{ and } 0 \lor 0 = 0 \text{ [see (v)].} \]

Then, \( x \land y = \min \{x, y\} \) and \( x \lor y = \max \{x, y\} \).

**Proof.** We start by listing a few consequences of axioms (1)–(7) for \( \land \) and \( \lor \), and begin by considering the function \( h(x) = x \land x \). According to Eq. (6) \( h(0) = 0 \), and Eq. (7) states that \( h(1) = 1 \). Since \( h \) is continuous (Eq. (4)) and strictly increasing (Eq. (5)) it maps \([0, 1]\) one to one onto \([0, 1]\). If \( h(x) = a \), then (using Eqs. (3) and 6),

\[ x \land x = a \leq a \lor (a \land a) = (a \lor a) \land (a \lor a), \]

so from Eq. (5), \( x < a \lor a \). On the other hand

\[ x \geq x \land (x \lor x) = (x \land x) \lor (x \lor x) = a \lor a, \]

so we must have \( x = a \lor a \). Thus, for \( a \in [0, 1] \),

\[ x \land x = a \iff x = a \lor a. \]

Also, since the above inequalities are actually equalities, it follows that

\[ x \land (x \lor x) = x \lor (x \land x) = x \quad (x \in [0, \land 1]). \]

Putting \( x = a \lor a \) in the equality \( x = x \lor (x \land x) \) gives

\[ a \lor a = (a \lor a) \lor [(a \lor a) \land (a \lor a)] = (a \lor a) \lor a, \]

i.e. in combination with Eq. (2),

\[ a \lor a = [(a \lor a) \lor a] \lor a = (a \lor a) \lor (a \lor a), \]

and it follows from Eq. (5) that \( a \lor a = a \). The dual result \( a \land a = a \) is proved in the same way, and so

\[ a \lor a = a \land a = a \quad (a \in [0, 1]). \]

This result, Eq. (10), implies a strengthening of Eq. (9) to

\[ a \lor (a \land b) = a \land (a \lor b) = a \quad (a, b \in [0, 1]), \]

since

\[ a \leq a \lor (a \land b) = (a \lor a) \land (a \lor b) = a \land (a \lor b) \leq a. \]

Now, to prove the theorem, we assume \( a \) and \( b \) to be given in \([0, 1]\) with \( a \geq b \). Since \( a \land x \) is continuous and nondecreasing in \( x \) (Eq. (4)), with
ANALYTIC FORMALISM OF THEORY OF FUZZY SETS

\( a \land 0 = 0 \) (Eq. (6)) and \( a \land a = a \) (Eq. (10)), there must be some \( c \) satisfying \( a \land c = b \). But then we may use Eq. (11) to obtain

\[
a \lor b = a \lor (a \land c) = a = \text{max} \{a, b\},
\]

and finally apply Eq. (10) in

\[
a \land b = a \land (a \land c) = (a \land a) \land c = a \land c = b = \text{min} \{a, b\}.
\]

5. CONCLUDING REMARKS ABOUT FUZZY NEGATION

We conclude this note with a few remarks concerning the negation of fuzzy statements. Zadeh defines the level of acceptance of "not \( S \)" in terms of that of \( S \) by

\[
\mu(\text{not} \ S) = 1 - \mu(S)
\]
a very natural definition. Is it, on analogy with the definitions of "and" and "or," the only possible choice?

In looking for natural restrictions on a function giving the truth value of "not \( S \)" in terms of that of \( S \) we must not be misled by the intuition the normal sharp mathematical use of "not" has given us. We might, for instance, be inclined to reject logically contradictory statements like "\( S \) and (not \( S \))" completely. But in view of the definition of "and" in fuzzy set theory this would force us to assign the truth value 0 to "not \( S \)" whenever the level of acceptance of \( S \) is positive, and this is clearly not what we want.

The reason that sentences rejected by ordinary logic may have a positive level of acceptance in fuzzy set theory is clear. In the latter context a positive acceptance of a statement \( S \) does not exclude the acceptance of its negation, and therefore not of the statement "\( S \) and (not \( S \))" either. We may to some extent accept that a certain rose is red and at the same time not red. When we relax the rigid either-or condition of set theory the concept of negation necessarily becomes a fuzzy one. The borderline between \( S \) and "not \( S \)" is no longer sharp.

What conditions are, then, natural to impose on a function \( N \) connecting the truth values of \( S \) and not" \( S \)"

\[
\mu(\text{not} \ S) = N[\mu(S)]?
\]

Well, in the first place we should have

\[
N(0) = 1 \text{ and } N(1) = 0, \quad (\text{N1})
\]
in order that ordinary set theory remains a special case of fuzzy set theory. Secondly,

\( N \) is continuous and strictly monotonic decreasing \( \quad (\text{N2}) \)
since our acceptance of "not S" should become smaller when our acceptance of S increases. Also, in line with (vi) in Sec. 3, we would like to have

\[ N[N(\mu)] = \mu, \quad (N3) \]

or, equivalently \( N = N^{-1} \), i.e. we should be just as willing to accept "not (not S)" as we are to accept S.

The above conditions, Eqs. (N1)–(N3), do not determine \( N \) uniquely. Nor is \( N \) determined if we in addition require

\[ \mu(\frac{1}{2}) = \frac{1}{2}, \]

thus agreeing to put \( \mu(S) = (1/2) \) when we decide to give the same level of acceptance to S as to "not S."

In order to determine \( N(\mu) = 1 - \mu \) uniquely, it seems necessary to introduce a condition like: "A certain change in the truth value \( \mu(S) \) of S should have the same effect on the acceptance of "not S" regardless of the value of \( \mu(S) \)." But we do not feel entirely convinced that such a condition is really necessary. There might be situations where an increase in the truth value of S does not necessarily correspond to a numerically equal decrease in our acceptance of "not S." If this is so, there might not exist any general rule to determine \( \mu(\text{not } S) \) in terms of \( \mu(S) \). Still, the rule \( \mu(\text{not } S) = 1 - \mu(S) \) appears quite reasonable in practical applications.

REFERENCES


Received September 8, 1971