Proceedings of the 1980 Winter Simulation Conference
T.I. Ören, C.M. Shub, P.F. Roth (eds.)

# A FUZZY SETS BASED LINGUISTIC APPROACH: THEORY AND APPLICATIONS

Piero P. Bonissone
General Electric Company
Corporate Research and Development
Building 37-579
1 River Road
Schenectady, NY 12345

Fuzzy sets theory and fuzzy logic constitute the basis for the linguistic approach. Under this approach, variables can assume linguistic values. Each linguistic value is characterized by a label and a meaning. The label is a sentence of a language. The meaning is a fuzzy subset of a universe of discourse. Models, based on this approach, can be constructed to simulate approximate reasoning. The implementation of these models presents two major problems, namely how to associate a label to an unlabelled fuzzy set on the basis of semantic similarity (linguistic approximation) and how to perform arithmetic operations with fuzzy numbers. For each problem a solution is proposed. Two illustrative applications are discussed.

## 1. INTRODUCTION

There are situations where it is more natural to handle uncertainty by fuzzy set theory (Zadeh 1965) than by probability theory.* Such is the case when dealing with the inherent imprecision of concepts involved in human reasoning and natural language. (Lakoff 1973, Zadeh 1975b, 1975c, Hersch 1975, Gaines 1976, Gupta 1977).

The theory of fuzzy sets has been the basis for the development of the linguistic approach (Zadeh 1975b) and its corresponding fuzzy logic (Bellman 1976).

In this approach any variable is treated as a linguistic variable, i.e., it can assume linguistic values. A linguistic value is composed of its syntactic value or label, a sentence belonging to a term set, and its semantic value, the membership distribution of a fuzzy set defined on a universe of discourse. Fuzzy logic is a logic whose truth-values are linguistic. More precisely, fuzzy logic is a fuzzification of Łukasiewicz infinite - valued logic or $L_{aleph-1}$ (Rescher 1969, Tsukamoto 1979).

An important feature of fuzzy logic is its ability to deal with approximate causal inference. Given an inference scheme "if P then Q", where P and Q are fuzzy propositions, it is possible from a proposition P', which approximately matches P, to deduce Q', which is similar to Q, by means of a logical interpolation referred to as "generalized modus ponens". Thus, a decision table can be constructed by taking the union of several conditional fuzzy rules "IF Pi THEN Qi". When a fuzzy proposition P' is given, the decision table will produce an inference _without the need of an exact pattern-matching_ between P' and the premises Pi of each rule. (Zadeh 1975c)

This type of inference, which is impossible in ordinary logic systems, enables us to build a model to simulate approximate reasoning. This model can accept a set of sentences as input and, based on the information contained in them, can upon request derive conclusions or answer questions.

Fuzzy logic provides the logical operators ($\cup, \cap, \neg, \rightarrow$) used to build the model. The extension principle (Zadeh 1975b) provides the mathematical tool to perform any arithmetic function on fuzzy sets required by the model. Models of this type have been tested in fields such as process control (Zadeh 1973,

*Probability related to randomness, deals with the uncertainty of whether a given element belongs or does not belong to a well defined set. On the other hand, fuzzy set theory and its associated possibility theory (Zadeh 1978), are related to fuzziness. They deal with the uncertainty derived by the partial membership of a given element to a set whose boundaries are not sharply defined.

Kickert 1976, 1978a, King 1977, Procyk 1977, Pappis 1977), decision analysis (Bellman 1970, Zadeh 1975a, 1977, Baas 1977, Kickert 1978b, Tong 1979, 1980), artificial intelligence (Bonissone 1978a), etc.

## 2. THE LINGUISTIC APPROACH: THEORY

For any given problem context, a data base, composed of a term set, is generated. The term set is a set of sentences belonging to a language $\mathcal{L}$. This language can be generated by a context free grammar G as proposed by Zadeh (Zadeh 1973) or by a regular grammar. Since the generative grammar G is a 4-tuple ($V_N$, $V_T$, S, P) where $V_N$ is the set of nonterminals, $V_T$ is the set of terminals, S is the starting symbol and P the productions then our choice of these will determine the size and form of the term set. Obviously, this will be problem dependent, but in general $\mathcal{L}$ should be large enough such that any possible situation of the problem context can be described.

However, in most practical cases, $\mathcal{L}$ does not have to be infinite, since only an approximate description of each particular situation is required. Moreover, $\mathcal{L}$ must be easily understandable. Thus, complex syntactic structures, such as the illimited recursive use of the same production rule by means of a cyclic nonternimal (which yields to an infinite language) should be avoided.

The syntactic representation is given by a set of labels. Its semantic characterization is provided by a set of membership distributions. A distribution is associated with each label. Therefore, a language can be seen as a fuzzy relation $\mathcal{L}$ from a set of terms S to a universe of discourse U, which assigns to pair (s, u), element of SxU, a grade of membership $\mu_{\mathcal{L}}(s,u)$.

If we fix s, the membership function $\mu_{\mathcal{L}}(s,u)$ determines a fuzzy subset A(s) of U whose membership function is:

$$\mu_{A(s)}(u) = \mu_{\mathcal{L}}(s,u) \qquad u \in U \, , \, s \in S$$

The fuzzy subset A(s) of U is taken as the <u>meaning</u> of s.

The term s is the <u>label</u> of A(s).

Among the terminals of G we find primary terms (e.g., <u>young</u>, <u>middle-aged</u>, <u>old</u>), hedges (e.g., <u>not</u>, <u>much</u>, <u>very</u>, <u>rather</u>, <u>more or less</u>), relations (e.g., <u>younger than</u>, <u>older than</u>), conjunctions (e.g., <u>and</u>, <u>but</u>), and disjunctions (e.g., <u>or</u>). While the primary terms are labels of primary fuzzy sets, the rest can be seen as labels of different kinds of operators which act on the primary fuzzy sets, modifying their original membership distributions (Lakoff 1973, Hersch 1975, Zadeh 1975b). Thus, any sentence s of language $\mathcal{L}$, regardless of its syntactic complexity, is associated with a fuzzy set A(s), characterized by its membership function $\mu_{A(s)}(u)$ or, for simplicity of notation, $\mu_s(u)$.

## 3. COMPUTATIONAL PROBLEMS

Various models, based on this approach, have been implemented on the computer (Wenstop 1975, Procyk 1976, Bonissone 1978a, 1978b, 1979, Eshragh 1978, Tong 1979). In most of them, a context free grammar generated the language.* The membership distribution representing the meaning of each sentence, was obtained by proper modification of primary fuzzy sets. It was sampled and stored as a vector associated with the sentence. APL was the high level language used to construct the models.

However, the computational burden of the implementation of these models has been considerable. The flexibility introduced by the linguistic approach is a valuable tool only if it becomes efficient. In order to achieve this efficiency, two computational problems must be solved.

The first problem arises when the output produced by the model is required to have the same nature as the input, i.e. to be linguistic. The model, by operating on the fuzzy sets associated with the input, generates as output another fuzzy set. The problem, also referred to as linguistic approximation (Procyk 1975, Eshragh 1978, Bonissone 1978b, 1979), consists of finding a label, generally a sentence in a language, whose meaning is the same or the closest to the meaning of the unlabelled fuzzy set generated by the model. Because of the inefficiency of performing pairwise comparisons for all the sentences of the language, the solution to this problem is based on pattern recognition techniques (Bonissone 1978b, 1979).

The second computational problem is related to the implementation of the extension principle. This principle enables any nonfuzzy function to accept fuzzy sets as arguments. The resulting function value is also a fuzzy set with a uniquely defined membership distribution. This principle is invoked every time that an arithmetic operation is required by the model.

---

*The language used in these implementations was a finite one, since no cyclic no nterminals were allowed in the grammar. Thus, a regular grammar could have been used to generate the same language.

However, the use of the extension principle with sampled membership distributions generates a considerable increase in the finite discrete support of the result. Furthermore, the membership distribution of the result is no longer sampled at regular intervals and only some of its membership values are actually correct (Bonissone 1980).

The solution to this problem consists of using a parameter-based representation for fuzzy sets. The arithmetic operations are defined as function of these parameters. Sampling is performed on the output set. Existing studies on this problem (Dubois 1978, 1979) are expanded in this paper and a table of formulae of basic arithmetic operations is provided.

## 4. SOLUTION TO THE FIRST PROBLEM (LINGUISTIC APPROXIMATION)

As briefly mentioned in section 3., the problem of linguistic approximation consists of associating a label with a membership distribution on the basis of semantic similarity. This can be seen as a mapping from the crisp set of all fuzzy subsets of the universe of discourse U into the language $\mathcal{L}$. The proposed solution to this problem relies on feature extraction and other pattern recognition techniques.

It is based on the assumption that the cardinality of the term set is finite, i.e. $|\mathcal{L}| \leq M < \infty$ . This assumption was heuristically justified by the requirements imposed on the language $\mathcal{L}$ in section 2.

The following is a more detailed explanation of this solution. The semantic part of the term set is mapped from the space of membership distributions into a feature space, by evaluating four weekly correlated features for each distribution. Assume that the universe of discourse, U, on which the fuzzy sets are defined, is finite and discrete and that $|U| = D$. Then we define a function F such that

$$F: \quad [0,1]^D \to R^N, \quad N \ll D$$

which maps each fuzzy set $L_i$ onto the N-dimensional space $\mathcal{P}$. Each element in $\mathcal{P}$ is a point (denoted by the vector $\underline{P}_i$) corresponding to the values of the characteristic features of $L_i$. Thus

$$F(\mu_{L_i}(x)) = \underline{P}_i \hat{=} (p_i^1, p_i^2, \ldots, p_i^N)$$

This is a crucial step, since the correct selection of features determines the success or failure of almost any pattern recognition process. In choosing these, we try to have the minimum number consistent with a good representation of the original data. Four features have proved themselves to be efficient in practice (Bonissone 1978b, 1979).

The first feature is the power of the fuzzy set. That is

$$p_i^1 = \sum_{k=1}^{D} \mu_{L_i}(x_k)$$

The second feature is a measure of the fuzziness of the set. Using a definition proposed by Loo (Loo 1977)

$$p_i^2 = \left[\sum_{k=1}^{D} H^2(\mu_{L_i}(x_k))\right]^{0.5}$$

where $H(h) = h$, $0 \leqslant h \leqslant 0.5$, and $H(h) = 1 - h$, $0.5 \leqslant h \leqslant 1$.

The third and fourth features are the first moment of the membership distribution of the fuzzy set and its skewness. They are respectively a measure of the "center of gravity" and of the asymmetry.

Then, each sentence in the language is represented by a pattern of characteristic features. The proposed method proceeds in two main steps. The first step consists of evaluating the four features of the unlabelled fuzzy set generated by the model. In our notation this fuzzy set will be referred to as Z. A quadratic weighted distance in the feature space is used as a metric to evaluate similarity between two fuzzy sets. That is

$$d_1(P_1,P_2) = \{\sum_{i=1}^{N} W_i^2(p_1^i - p_2^i)^2\}^{1/2}$$

The weights $W_i$ play an important role since they allow different features to be emphasized. They are defined as $W_i = I_i/R_i$, where $R_i$ is the length of the range of values that feature $p^i$ takes over all the points in the data set, and $I_i$ is the relative importance of parameter $p^i$ in evaluating semantic similarity. Clearly, $I_i$ depends on the user's subjective perception of similarity, for the given problem context. It is suggested to obtain the value of $I_i$ from the user by means of pairwise comparison tests, using the concept of cardinal ratio scale to establish a preference cardinal ordering (Saaty 1974, 1979, Bonissone 1979).

A search in this low order pattern space is performed. The result of this pre-screening process of the language is a small crisp subset of sentences, LA(Z), all of them close in meaning to the unlabelled fuzzy set Z (according to some tolerance parameter).

In the second step, the final assignment of the label is obtained when the membership distributions of these preselected sentences are compared with the one of the unlabelled fuzzy set. For this purpose, a measure to evaluate the semantic similarity of two fuzzy sets is proposed. The metric we use is a modified form of Bhattacharya distance that is defined by

$$d_2(L_i,Z) = [1-R(L_i,Z)]^{1/2} \quad , \quad L_i \in LA(Z)$$

where R is called the Bhattacharya coefficient. In the discrete case this is given by

$$R(L_i Z) = \sum_{k=1}^{D} \left[\frac{\mu_{L_i}(x_k)\mu_Z(x_k)}{\text{power}(L_i)\text{power}(Z)}\right]^{1/2}$$

5. SOLUTION TO THE SECOND PROBLEM (ARITHMETIC OPERATIONS)

Before we use the proposed parameter-based representation for fuzzy sets, let us introduce the definition of fuzzy number. A fuzzy number is simply a fuzzy subset of the real line and is completely defined by its membership function such that

$$\mu: \mathbb{R} \to (0,1)$$

For our purposes, we further restrict this definition to those fuzzy numbers which are both normal and convex. Thus in addition to the above constraint we have

$$\text{normality:} \quad \bigvee_x \mu(x) = 1.0 \qquad x \in \mathbb{R}$$

$$\text{convexity:} \quad \mu(\lambda x_1 + (1-\lambda)x_2) \geq \mu(x_1) \wedge \mu(x_2) \qquad \forall x_1,x_2 \in \mathbb{R}, \ \forall \lambda \in [0,1]$$

where V and Λ indicate supremum and minimum, respectively.

With the requirement of convexity, a piecewise continuity in the membership distribution is assured. This requirement also implies that the points of the real line, with the highest membership values, are clustered around a given real interval (or point). This fact allows us to easily understand the semantics of a fuzzy number by looking at its distribution and to associate it with a properly descriptive syntactic label (e.g., "approximately 100").

The requirement of normality implies that, among the points of the real line with the highest membership value, there exists at least one which is completely compatible with the predicate associated with the fuzzy number. In other words, it is totally possible for the fuzzy number to take that (or those) particular value(s) on the real axis.

It should be noted that fuzzy numbers do not form a ring, since they lack the inverse element for the operations of "+"(sum) and "X"(multiplication). Moreover, only positive convex fuzzy numbers form a commutative semiring. (Mizumato 1976, Dubois 1978). In fact, negative fuzzy numbers are not closed under "x". Any other type of fuzzy numbers do not satisfy the distribuitive law.

Any normal convex fuzzy number may be characterized by a 4-tuple $(a,b,\alpha,\beta)$ where $[a,b]$ is the closed interval on which the membership function is equal to 1.0, $\alpha$ is the "left bandwidth" and $\beta$ is the "right bandwidth".* Figure I illustrates this characterization.

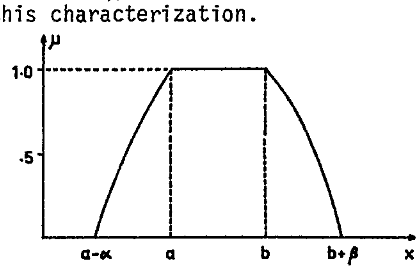


Fig. I. Characterization of fuzzy number $A = (a,b,\alpha,\beta)$

Notice that crisp numbers can be represented in this form by $(a,b,0,0)$ and that interval-numbers may be written as $(a,b,0,0)$. Thus, this characterization allows for the use of "mixed" arithmetics, where fuzzy numbers can be easily combined with interval numbers or crisp numbers in the cases where more precise information is available.

*Refer to Appendix 1 for a formal definition of left and right bandwidths.

If we now limit the shape of the left and right slopes of the mambership function to be an even function, $S(*)$, such that $S(-x) = S(x)$ and $S(0) = 1$, then, if $S(*)$ is also monotonically decreasing on $[0,+\infty)$, the simple algebraic operations can be written as formulae involving the parameters in the 4-tuple.

So, if $\tilde{m} \triangleq (a,b,\alpha,\beta)$ and $\tilde{n} \triangleq (c,d,\gamma,\delta)$ with the understanding that the left slopes are given by $S(a-x/\alpha)$ and $S(c-x/\gamma)$, and the right slopes by $S(x-b/\beta)$ and $S(x-d/\delta)$, then Table 1 gives the formulae for addition, subtraction, multiplication, division and power.

Table 1: Basic operations with fuzzy numbers

| Operation | Result | Conditions |
|---|---|---|
| $\tilde{m}+\tilde{n}$ | $(a+c,\ b+d,\ \alpha+\gamma,\ \beta+\delta)$ | all $\tilde{m}$, $\tilde{n}$ |
| $\tilde{m}-\tilde{n}$ | $(a-d,\ b-c,\ \alpha+\delta,\ \beta+\gamma)$ | all $\tilde{m}$, $\tilde{n}$ |
| $\tilde{m}\times\tilde{n}$ | $(ac,\ bd,\ a\gamma+c\alpha-\alpha\gamma,\ b\delta+d\beta+\beta\delta)$ | $\tilde{m}>0$, $\tilde{n}>0$ |
| | $(ad,\ bc,\ d\alpha-a\delta+\alpha\delta,\ -b\gamma+c\beta-\beta\gamma)$ | $\tilde{m}<0$, $\tilde{n}>0$ |
| | $(bd,\ ac,\ -b\delta-d\beta-\beta\delta,\ -a\gamma-c\alpha+\alpha\gamma)$ | $\tilde{m}<0$, $\tilde{n}<0$ |
| $\tilde{m}\div\tilde{n}$ | $\left(\frac{a}{d},\ \frac{b}{c},\ \frac{a\delta+d\alpha}{d(d+\delta)},\ \frac{b\gamma+c\beta}{c(c-\gamma)}\right)$ | $\tilde{m}>0$, $\tilde{n}>0$ |
| | $\left(\frac{a}{c},\ \frac{b}{d},\ \frac{c\alpha-a\gamma}{c(c-\gamma)},\ \frac{d\beta-b\delta}{d(d+\delta)}\right)$ | $\tilde{m}<0$, $\tilde{n}>0$ |
| | $\left(\frac{b}{c},\ \frac{a}{d},\ \frac{-b\gamma-c\beta}{c(c-\gamma)},\ \frac{-a\delta-d\alpha}{d(d+\delta)}\right)$ | $\tilde{m}<0$, $\tilde{n}<0$ |
| $\tilde{m}^{\tilde{n}}$ | $(a^c, b^d, a^c-(a-\alpha)^{c-\gamma}, (b+\beta)^{d+\delta}-b^d)$ | $\tilde{m}\in[1,\infty)$ $\tilde{n}>0$ |
| | $(b^c, a^d, b^c-(b+\beta)^{c-\gamma}, (a-\alpha)^{d+\delta}-a^d)$ | $\tilde{m}\in[1,\infty)$ $\tilde{n}<0$ |
| | $(a^d, b^c, a^d-(a-\alpha)^{d+\delta}, (b+\beta)^{c-\gamma}-b^c)$ | $\tilde{m}\in[0,1]$ $\tilde{n}>0$ |
| | $(b^d, a^c, b^d-(b+\beta)^{d+\delta}, (a-\alpha)^{c-\gamma}-a^c)$ | $\tilde{m}\in[0,1]$ $\tilde{n}<0$ |

where $\tilde{m} \triangleq (a,b,\alpha,\beta)$ and $\tilde{n} \triangleq (c,d,\gamma,\delta)$

Note that the last ten formulae are only approximate in that the left and right bandwidths of the result are not exact. However, they introduce very little error and in practice have proved themselves to be of great value (Bonissone 1980).

We now have a compact way of representing the kind of fuzzy sets in which we are interested. Furthermore, because we can perform algebraic operations with these representations we do not need to use the extension principle but we can compute the output set induced by a mapping in a direct way. Sampling needs only then to be performed on the output set, at which stage we can fix the number of sample points in accordance with our requirements.

## 6. APPLICATION OF THE SOLUTIONS

### 6.1 Example 1: An Application in Decision Analysis

The validity of the proposed solutions to the above problems is illustrated by a short example. The example summarizes the author's recent work in decision analysis (Tong and Bonissone 1979, 1980) where the two solutions were extensively used.

A one stage multichoices multicriteria decision problem, simulated by a computer implemented system, describes the following investment situation.

A private citizen has a moderately large amount of capital which he wishes to invest to his best advantage. He has selected five possible investment areas, $\{a_1,a_2,a_3,a_4,a_5\}$, and has four criteria, $\{c_1,c_2,c_3,c_4\}$, by which to judge them. These are

$a_1$ - the commodity market
$a_2$ - the stock market
$a_3$ - gold and/or diamonds
$a_4$ - real estate
$a_5$ - long term bonds

and

$c_1$ - the risk of losing the capital sum
$c_2$ - the vulnerability of the capital sum to modification by inflation
$c_3$ - the amount of interest received
$c_4$ - the cash realisability of the capital sum.

His rating of the alternatives with respect to the criteria ($r_{ij}$: $i=1,\ldots,5$, $j=1,\ldots,4$) is expressed linguistically as shown in Table 2. His problem is to select one of the $a_i$ with the additional constraint that he does not consider the criteria to be equally important but gives them linguistic weights, $(\alpha_1,\alpha_2,\alpha_3,\alpha_4)$ as shown in Table 3.

Table 2

| | $c_1$ | $c_2$ | $c_3$ | $c_4$ |
|---|---|---|---|---|
| $a_1$ | high | moreorless high | very high | fair |
| $a_2$ | fair | fair | fair | moreorless good |
| $a_3$ | low | from fair to moreorless low | fair | good |
| $a_4$ | low | very low | moreorless high | bad |
| $a_5$ | very low | high | moreorless low | very good |

Table 3

| $\alpha_1$ | $\alpha_2$ | $\alpha_3$ | $\alpha_4$ |
|---|---|---|---|
| indifferent | moreorless important | very important | moreorless unimportant |

In both cases, describing the weights and the ratings, the system provides the user with a term set of linguistic values. These linguistic values are fuzzy numbers defined on the closed interval [0,1] of the real line. Furthermore, S(*) is linear, giving a particularly simple form. Only seven basic set shapes are used to represent the range of linguistic values (see Figure II). This means that each set has several interpretations. Table 4 illustrates the definitions and interpretations of the linguistic values

Table 4

| | | Interpretation when used with | | | | |
|---|---|---|---|---|---|---|
| Shape | Fuzzy Numbers | $c_1$ | $c_2$ | $c_3$ | $c_4$ | Weights |
| 1 | (0,0,0,.2) | Very high | Very high | Very low | Very bad | Very unimportant |
| 2 | (0,.1,0,.2) | High | High | Low | Bad | Unimportant |
| 3 | (.2,.2,.2,.2) | More or less high | More or less high | More or less low | More or less bad | More or less unimportant |
| 4 | (.5,.5,.2,.2) | Fair | Fair | Fair | Fair | Indifferent |
| 5 | (.8,.8,.2,.2) | More or less low | More or less low | More or less high | More or less good | More or less important |
| 6 | (.9,1,.2,0) | Low | Low | High | Good | Important |
| 7 | (1,1,.2,0) | Very low | Very low | Very high | Very good | Very important |

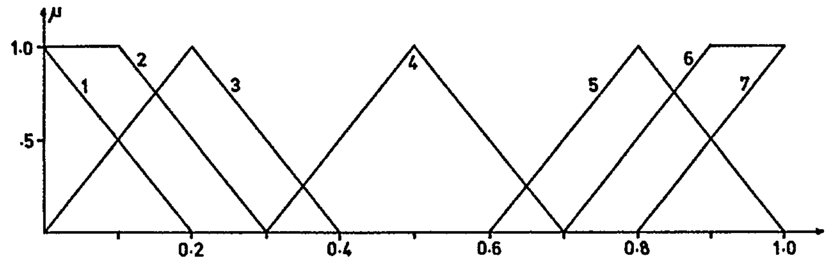


Fig. II. Linguistic ratings and weights

The first step in solving this problem is to compute a suitability set for each of the alternatives. Since the linguistic ratings and weights are appropriately defined fuzzy numbers, we just use a fuzzy weighted sum to give

$$S_i = \sum_{j=1}^{4} \alpha_j r_{ij}$$

These operations are executed by means of the formulae of Table 1. The results are as shown in Figure III.

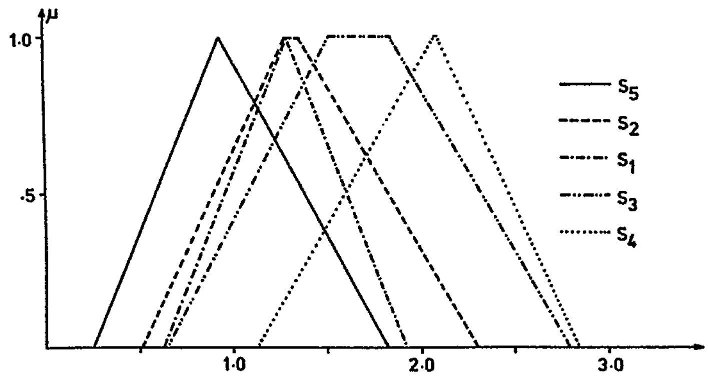


Fig. III. Suitability sets

Applying the concept of dominance relation (Tong and Bonissone 1980) to the suitability sets, we obtain the overall degree to which each of the alternatives dominates the others. A difference fuzzy set Z between the best alternatives and a weighted average of the remaining ones constitute the "degree of preferability" of the chosen alternative. The weights used in this average are the degree of dominance of the selected alternative.

The obtained fuzzy set Z represents the amount by which the chosen solution is preferable to the other ones. By using the linguistic approximation technique, this is interpreted and labelled.

In the case of this example, alternative 4 (real estate) was selected, since $S_4$ dominates the other suitability sets. The preferability of the solution was labelled "from indifferent with to marginally better than the other alternatives". This weak statement about the strength of the decision was due to the proximity of $S_3$ (gold/diamonds) to the chosen $S_4$ (real estate).

6.2 Example 2: An Application in Artificial Intelligence.

The simple example, used in this section, deals with a typical problem of Artificial Intelligence (A.I.), namely natural language understanding and approximate reasoning.

In this problem, the input is a sequence of sentences belonging to a subset of English, limited by a lexicon of 96 words. An Augmented Finite State Transition Network parser (Woods 1970), implemented in APL, analyzes each sentence, verifies its syntactic correctness and interprets its associated meaning. The meaning, as mentioned in Section 2, is represented by a fuzzy set.

For example, the sentence "Ann is very tall and Nancy is much shorter than Ann" is transformed into:

Height (Ann) = very tall
Height (Nancy) = much shorter than very tall

where "Tall" is a primitive fuzzy subset of a universe of discourse U representing values of height. "Very tall" is a fuzzy set obtained by applying the modifier "very" to the fuzzy set "tall" (Lakoff 1973, Zadeh 1975a). This is illustrated in Fig. IV. In this example the linguistic value assigned to Height (person) has the same meaning for both males and females.

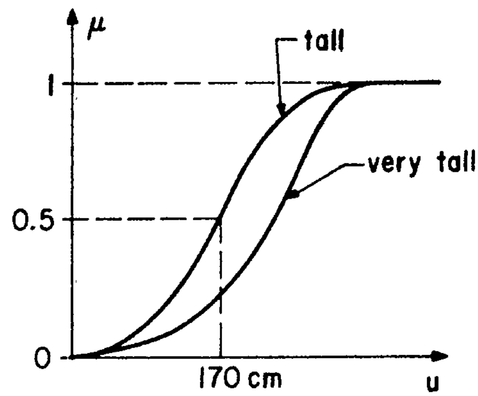


Fig. IV. Representation of "Tall" and "Very tall"

We could consider the entire procedure as the assignment of a linguistic value, e.g., very tall, to a linguistic variable, e.g., Height (Ann). Each time a sentence is parsed, the linguistic variables, present or identifiable in the sentence, are assigned their corresponding values. After parsing the sentences, the system is ready to answer questions related to the data. The answer to these questions can be viewed as finding the final value of a particular linguistic variable.

To clarify this presentation, we list the sentences used as input to the parser. Given that:

> Ann is very short and Nancy is much taller than Ann.
> Gail is indeed very tall.
> Nancy is not much shorter than Gail.
> Paola is medium-height.
> Mary is taller than Ann but Mary is taller than Paola is false.
> John is very tall is rather false but John is taller than Paola.
> Dennis is shorter than John but Dennis is taller than Mary.
> Piero is shorter than Nancy.
> Piero is taller than Paola.

we may want to ask "What is the height of Dennis" or "What is the height of John". The answers are obtained by taking the intersection, e.g. AND, of the assignments (clauses) associated with the linguistic variables Height (Dennis) or Height (John).

This is a purely syntactic procedure. Therefore some answers could be lengthy and incomprehensible as the answer Height (Dennis) = ((SHORTER THAN (((VERY TALL) IS RATHER FALSE)) AND (TALLER THAN MEDIUM-HEIGHT))) AND ((TALLER THAN ((TALLER THAN VERY SHORT)) AND ((TALLER THAN MEDIUM-HEIGHT) IS FALSE))) which would be obtained by following the described syntactic procedure.

At this point, the linguistic approximation comes into the picture. The meaning of each answer is determined (by obtaining the corresponding membership distribution) and a more understandable label is associated with it. In our experiment, we interrogated the system about the height of Nancy, Mary, Piero, John and Dennis, with the following results:

| | |
|---|---|
| Height (Nancy) | = Very tall |
| Height (Mary) | = From sort of short to medium-height. |
| Height (Piero) | = Between medium-height and rather tall. |
| Height (John) | = Sort of tall. |
| Height (Dennis) | = Between medium-height and sort of tall. |

The details of the implementation of the parser are omitted in this discussion. They can be found in (Bonissone 1978a). Details of the computer implementation of this example can be found in (Bonissone 1979).

It is important to note that the obtained answers are not precise in meaning, since they do not specify a given height in units such as meters or feet. On the contrary, they reflect the fuzziness of the input data. Nevertheless, the answers are perfectly consistent with the ones provided by a human on the basis of his/her intuitive understanding of the data and approximate reasoning capabilities.

## APPENDIX 1

Basic concepts of Fuzzy Sets.

Given a collection of objects, referred to as the universe of discourse $U$, a fuzzy subset $A \subset U$ is a class of objects of the collection with not well defined boundaries. Since the predicate which characterizes the fuzzy subset A does not cause a dichotomy on the universe $U$, each element of $U$ may have a grade of membership which ranges between zero and full membership. Therefore, we can say that the basic property of a fuzzy subset is its gradual rather than abrupt transition from membership to non-membership.

The function which assigns these membership values is referred to as $\mu_A(u)$ and it is a mapping from the universe $U$ into the interval $[0,1]$* i.e.

$$\mu_A(u) : U \dashrightarrow [0,1]$$

The membership function is context dependent. The grades of membership reflect an "ordering" of the objects in $U$ induced by the predicate associated with A. Therefore $\mu_A(u_i)$ can be interpreted as the degree of compatibility of the predicate associated with A and the object $u_i$. We can represent A as:

$$A = \bigcup_u \mu_A(u)$$

Given a universe $U$, we say that $A \subset U$ is fully characterized by its membership function $\mu_A(u)$.

OPERATIONS

We define the following operations on fuzzy sets , based on their corresponding membership functions. Let A, B, C be three fuzzy subsets of $U$.

Equality ( = )

$$A = B \text{ iff } \mu_A(u) = \mu_B(u) \quad \text{for all } u \in U$$

Containment ( $\subset$ )

$$A \subset B \text{ iff } \mu_A(u) \leqslant \mu_B(u) \quad \text{for all } u \in U$$

Union ( OR )

$$C = A \text{ OR } B \text{ iff } \mu_C(U) = \underset{u \in U}{\text{Max}} (\mu_A(u), \mu_B(u))$$

Intersection ( AND )

$$C = A \text{ AND } B \text{ iff } \mu_C(u) = \underset{u \in U}{\text{Min}} (\mu_A(u), \mu_B(u))$$

Complementation ( NOT )

$$B = \text{NOT } A \text{ iff } \mu_B(u) = 1 - \mu_A(u)$$

Bounded Sum ( $\oplus$ )

$$C = A \oplus B \dashrightarrow \mu_C(u) = \underset{u \in U}{U} (1 \text{ AND } (\mu_A(u) + \mu_B(u)))$$

Bounded Difference ( $\ominus$ )

$$C = A \ominus B \dashrightarrow \mu_C(u) = \underset{u \in U}{U} ( 0 \text{ OR } (\mu_A(u) - \mu_B(u)))$$

*Goguen showed that in a more general case the membership function is a map from $U$ into a distributive lattice L. (Goguen 1967) However for all the proposed applications we will use the particular case of $L = [0,1]$

Power

Power is used here in its arithmetic sense. Raising a fuzzy set A to a real number power "b" is defined as:

$$B = A^b \quad \text{-->} \quad \mu_B(u) = (\ \mu_A(u)\ )^b$$

Implication ( ==> )

Several ways of defining implication could be listed. We will limit ourselves to the two most common definitions:

$$A \Longrightarrow B \quad = (A \times B) \ \text{OR} \ ((\text{NOT } A) \times V)$$

or

$$A \Longrightarrow B \quad = (\text{NOT } (A \times V) + (U \times B))$$

where

$U$ and $V$ are the universe of discourse of A and B respectively.

x is the Cartesian product which is defined as:

$$C = A \times B \ \text{-->} \ \mu_C(u_i, u_j) = \bigcup_{u \in U} (\ \mu_A(u_i) \ \text{AND} \ \mu_B(u_j))$$

Union, intersection and complementation satisfy De Morgan's laws, as well as the associative and distributive properties.

Extension Principle

This principle allows any non-fuzzy function to be fuzzified in the sense that if the function arguments are made fuzzy sets, then the function value is also a fuzzy set whose membership function is uniquely specified.

More formally, if the scalar function, f, takes n arguments $(\chi_1, \chi_2, \ldots, \chi_n)$ denoted by $\underline{X}$, and if the membership functions associated with each of these is given by $\mu_1(\chi_1), \mu_2(\chi_2), \ldots, \mu_n(\chi_n)$ then

$$\mu_{f(\underline{X})}(y) = \bigvee_{\underline{X}} \left[ \bigwedge_{i=1}^{n} \mu_i \ (\chi_i) \right]$$
$$\text{s.t. } f(\underline{X}) = y$$

where V and Λ indicates supremium and minimum, respectively.

## DEFINITIONS

a-level

An a-level set of a fuzzy set A is defined as:

$$A_a = \{ u \mid \mu_A(u) \geq a \}$$

Then A satisfies the resolution identity

$$A = \bigcup_{a \in [0,1]} (\ a A_a\ )$$

Bandwidth

The bandwidth of A is the 0.5 level-set of A.

Support

The support of A is defined as

$$\text{Support}(A) = S_A \; \{ u \mid \mu_A(u) > 0 \}$$

Left Bandwidth and Right Bandwidth

This terminology is only used when dealing with parametrized fuzzy numbers (see Section 5). Therefore, given a fuzzy number $A \in \mathbb{R}$ (A is convex and normal by assumption), we define

$$\text{Left Bandwidth}(A) = L_A = (a-\alpha, a)$$
$$\text{Right Bandwidth}(A) = R_A = (b, b+\beta)$$

$$(a-\alpha) = \min \{ r \in \mathbb{R} \mid \mu_A(r) > 0 \}$$
$$a = \min \{ r \in \mathbb{R} \mid \mu_A(r) = 1 \}$$
$$b = \max \{ r \in \mathbb{R} \mid \mu_A(r) = 1 \}$$
$$(b+\beta) = \max \{ r \in \mathbb{R} \mid \mu_A(r) > 0 \}$$

REFERENCES

Bass S. J. & Kwakernaak H. (1977) Rating and ranking of multiple-aspect alternatives using fuzzy sets. Automatica, 13, pp. 47-58

Bellman R. E. and Zadeh L.A. (1970) Decision making in a fuzzy environment. Management Science, 17, No. 4, pp. 141-164

Bellman R. E. & Zadeh L. A. (1976) Local and fuzzy logic. Memorandum No. ERL-M584, University of California, Berkeley. Also in Modern Uses of Multiple-Valued Logic, pp. 103-165 J. M. Dunn and G. Epstein (eds.) D. Reidel Publishing Co., Dordrecht-Holland, 1977.

Bonissone P. P. (1978a) Report of final project for course CS292P "Natural Language Process", U.C. Berkeley, Fall 1978 (R. Wilensky)

Bonissone, P. P. (1978b) A pattern recognition approach to the problem of linguistic approximation in system analysis. Memorandum No. UCB/ERL M78/57 University of California, Berkeley. Also presented at IEEE 1979 International Conference on Cybernetics and Society, Denver, Colorado, Oct. 79.

Bonissone, P. P. (1979) The problem of linguistic approximation in system analysis. Ph. D. Thesis, Department of Electrical Engineering and Computer Sciences, University of California, Berkely.

Bonissone, P. P. (1980) Notes on fuzzy numbers. Technical Information Series, General Electric C.R.D. (in preparation).

Dubois, D. & Prade, H. (1978) Operations on fuzzy numbers. Int. J. Systems Sci., , pp 613-626.

Dubois, D. & Prade, H. (1979) Fuzzy real Algebra: Fuzzy Sets and Systems, 2, 4, pp. 327-348

Eshragh F. & Mamdani E. H. (1978) A general approach to linguistic approximation. Department of Electrical & Electronic Engineering, Queen Mary College, London (internal report)

Gaines B. R. (1976) Foundations of fuzzy reasoning. Int. J. Man-Machine Studies, 6, pp. 623-668.

Goguen, J. A. (1967) L Fuzzy Sets, J. of Mathematical Analysis and Applications, 18, pp. 145-174.

Gupta M. M. & Saridis G. N. (eds.) (1977) Fuzzy automata and decision processes. North-Holland Publishing Company, Amsterdam

Hersch H. M. & Caramazza A. (1975) A fuzzy set approach to modifiers and vagueness in natural language. J. Experimental Psychology, 105, No. 3, pp. 254-276

Kickert W. J. M. & Van Nauta Lemke H. R. (1976) Application of a fuzzy controller in a warm water plant. Automatica, 12, 301-308

Kickert W. J. M. & Mamdani E. H. (1978a) Analysis of a fuzzy controller. Fuzzy Sets and Systems, 1, pp. 29-44

Kickert W. J. M. (1978b) Fuzzy theories on decision-making. Martinus Nijhoff Social Sciences Division Leiden, The Netherlands.

King P. J. & Mamdani E. H. (1977) The application of fuzzy control system to industrial processes. Automatica, 13, 235-242

Lakoff G. (1973) Hedges: a study in meaning criteria and the logic of fuzzy concepts. Journal of Philosophical Logic, 2, pp. 458-508

Loo S. G. (1977) Measures of fuzziness Cybernetica, 20, 3, pp. 201-210

Mizumoto M. Tanaka K. (1976) Algebraic properties of fuzzy numbers, IEEE 1976 Internation. Conference on Cybernetics ans Society, Washington D.C.

Pappis C. P. & Mamdani E. H. (1977) A fuzzy logic controller for a traffic junction. IEEE Trans. on System, Man, and Cybern. vol. SMC-7, No. 10, 707-717

Procyk T. J. (1976) Linguistic representation of fuzzy variables. Fuzzy Logic Working Group, Queen Mary College, London

Procyk T. J. (1977) A self-organizing controller for single-input single-output systems. Fuzzy Logic Working Group, Research Rep. No. 6, Queen Mary College, London

Rescher N. (1969) Many-Valued Logic, McGraw-Hill

Saaty, T. L. (1974) Measuring the fuzziness of sets. Journal of Cybernetics, 4, 4, pp. 53-61

Saaty, T. L. (1979) Exploring the interface between hierarchies, multiple objectives and fuzzy sets. Fuzzy Sets and Systems, 1, pp. 57-68

Tong, R. M. (1977) A control engineering Review of fuzzy systems. Automatica, 13, pp. 559-569

Tong, R. M. and Bonissone P. P. (1979) Linguistic decision analysis using fuzzy sets. Memorandum No. UCB/ERL M79/72, Electronic Research Laboratory, U. C. Berkeley, California

Tong, R. M. & Bonissone, P. P. (1980) A linguistic approach to decision making with fuzzy sets. IEEE Transactions on Systems Man and Cybernetics, SMC-10, 11

Tsukamoto, Y. (1979) Fuzzy Logic Based on Łukasiewicz Logic and Its Applications to Diagnosis and Control. Ph.D. Thesis, Tokyo Institute of Technology, Tokyo, Japan

Wenstop. F. (1975) Application of linguistic variables in the analysis of organizations. Ph.D. Dissertation, School of Business Administration U.C. Berkeley

Woods, W. A. (1970) Transition Network Grammars for Natural Language Analysis. Comm. ACM, 13, #10, pp. 591-606

Zadeh, L. A. (1965) Fuzzy sets. Information and Control, 8, 338-353

Zadeh, L. A. (1973) Outline of a new approach to the analysis of complex systems and decision processes. IEEE Trans. on Systems, Man, and Cybern. vol. SMC-3, No. 1, 28-44

Zadeh, L. A. (1975a) The linguistic approach and its application to decision analysis. Proceedings of Conference on Directions in Decentralized Control, Many-person Optimization and Large Scale Systems, U. C. Ho and S. M. Kitter (eds.), Plenum Press

Zadeh, L. A. (1975b) The concept of a linguistic variable and its application to approximate reasoning
Part I Information Sciences 8, 199-249
Part II Information and Sciences, 8, 301-357
Part III Information and Sciences, 9, 43-80

Zadeh, L. A. (1975c) Fuzzy logic and approximate reasoning (In memory of Grigore Moisil). Synthese, 30, pp. 407-428

Zadeh, L. A. (1977) Linguistic characterization of preference relations as a basis for choice in social systems. Memorandum No. UCB/ERL M77/24 Electronics Research Laboratory, U.C. Berkeley

Zadeh, L. A. (1978) Fuzzy Sets as a Basis for a Theory of Possibility, Fuzzy Sets and Systems, 1, 1, pp. 3-28.