A GEOMETRICAL FUZZY PARTITIONS APPROACH TO FUZZY QUERY AND FUZZY DATABASE RETRIEVAL.

E. G. Mtalo and E. Derenyi

Department of Surveying Engineering
University of New Brunswick
Fredericton, N.B. CANADA
COMMISSION III

ABSTRACT

Real world knowledge is usually vague and ambiguous and human beings generally think and communicate in fuzzy non-precise terms. The fuzzy sets theory introduced by Zadeh in 1965 forms the mathematical and practical basis for the representation and manipulation of such fuzzy information.

Conventional databases, however, must contain precisely defined facts or data because database query languages such as the SQL impose strict formats for data entry and query. Conventional query languages do not permit ambiguous or non-precise queries.

This paper presents a new method for the representation of fuzzy numerical quantities in a way favorable to the storage and retrieval of fuzzy values or vague expressions in a database.

Key Words: Fuzzy Sets, Fuzzy Retrieval, Geometric Partitions, GIS, Database.

1. INTRODUCTION.

Conventional databases must contain precisely defined facts and numeric values because database query languages such as the SQL impose strict formats for data entry and query. Conventional query languages do not permit ambiguous or non-precise queries. Real world data and knowledge is usually vague and ambiguous and human beings generally think vaguely and communicate in fuzzy non-precise terms (Zadeh et al, 1975; Zadeh, 1989). However vague or fuzzy real world knowledge does not lend itself to easy manipulation by conventional database management systems.

The fuzzy sets theory introduced by Zadeh in 1965 (Zadeh et al. 1975; Kandel, 1986) forms the mathematical and practical basis for the representation and manipulation of fuzzy information.

1.1 Representation of Vague Information by Fuzzy Sets.

Consider the statement that the distance from the Earth to the Sun is "very great", or the statement "100 is much greater than 5". Using Zadeh's fuzzy sets theory, the terms "great" and "greater" may be regarded as *fuzzy sets* and can therefore be defined in terms of fuzzy membership functions. Once these fuzzy sets have been defined the modifiers, "very" and "much", can be applied to transform them into the corresponding fuzzy sets *very great* and *much greater* respectively(see Zadeh et al. 1975; Kaufmann and Gupta, 1988; Shmucker, 1984; Kandel, 1986; Zadeh, 1989).

Typically the fuzzy set *great* will be represented by a *fuzzy membership function* $\mu_{great}(x)$: X->[0,1] and *greater* can be represented by the membership function $\mu_{greater\ than}(x)$: X->[0,1]. Generally $\mu_A(x)$ denotes the membership function of the elements of the universe X in the fuzzy set A such that the elements, x, take value in the universe X. The membership values, $\mu_A(x)$ on the other hand take value in the *evaluation space* of the fuzzy set, generally considered to be the continuous interval [0,1] (Zadeh et al, 1975; Dubois and Prade, 1980; Kandel, 1986). The membership value,$\mu_A(x)$, is a real number between 0 and 1 inclusive, expressing the strength of the membership of an element, x, of the universe X in a the fuzzy set A.

Practical specification of a membership function involves the assignment of the parameters of some *standard membership function*, such as the S-function and π-function (Dubois and Prade 1980; Kandel 1986; Klir and Folger 1988). Important characteristic points of standard membership functions include; the *cut-out points* (points with zero membership value), the *peak point* (where the membership function attains the maximum value of 1) and the *turnover points* (points where the membership function has the value of 0.5). Dubois and Prade (1988, 1990) suggest that for most applications simpler piece-wise linear membership functions such as triangular functions and trapezoidal functions provide satisfactory results.

The method proposed in this paper differs from the standard approach in that, fuzzy expressions such as "greater than five" are characterized by geometric partitions induced by them in the real numbers domain. The paper begins by defining the concept of fuzzy geometric partitions and then goes on to show how it can be applied to the problem of storing and querying fuzzy database objects. For the purposes of fuzzy database retrieval, the universe of discourse is some search space X containing crisp and fuzzy objects. To facilitate database retrieval each fuzzy query and fuzzy database object can be associated with a fuzzy partition defined over the search space. The condition for a successful search is obtained when the partition induced by the query object contains the partition generated by the fuzzy database object.

The paper also looks at the potential use of the fuzzy geometrical partitions method in the construction of membership functions, which may then be used to represent fuzzy numerical sets in the usual way(Zadeh ,1975, 1989).

2. THE CONCEPT OF FUZZY GEOMETRIC PARTITIONS.

The concept of fuzzy partitions is not new, what is new is, however, the manner in which this concept is used to characterize and represent fuzzy valued expressions or *fuzzy numbers*. An earlier use of the term radial partition to characterize fuzzy sets can be found in Kaufmann (1975). Kaufmann shows that a fuzzy set, M, induced by the binary relation y >> x , for y = kx, k > 1, in the two dimensional real space constitutes a radial partition of the real numbers space.

In Dowsing et al.(1986) the concept of the *diagonal set* of the universe of discourse, is used to define and characterize the equality operator. This research extends and generalizes the concept of the diagonal set generated by the equality operator, defined in Dowsing et al.(1986), and uses it to define general fuzzy comparison operators or fuzzy binary relations, R(x,y), in terms of the radial sets or partitions induced by them in a two dimensional real space X.

Intuitively the operators "equal", "greater", and "less" define basic geometric partitions of the two dimensional space (Figure 1). The partition generated by the *equality operator*, =, is called the *diagonal set* (Dowsing et al, 1986) or diagonal partition. By extension the partition corresponding to the operator *greater* (>) is called the *upper diagonal partition*, while the partition of the operator *less* (<) is called the *lower diagonal partition* (Figure 1).

Using this scheme partitions corresponding to the fuzzy modifications of the operators *equal*, *greater*, and *less* , may be assigned in the search space as shown in Figure 2. This is possible because, by common sense reasoning, the fuzzy expression *about* modifies the expression *equal to* by generating a narrow band around the value x. The width of the band or size of the partition generated also depends on the magnitude of the crisp value x. By common sense the partition generated by "about x", "more or less x" , and "roughly equal to x" are radial partitions stretching over both sides of the line of equality (Figure 2). Naturally the partition induced by "roughly equal to x" must be somewhat wider than those generated by "more or less x", or "about x".

When the fuzzy modifier *much* is applied to the operators *great* and *less* it induces partitions which exclude all values close to the line of equality. It is logical, therefore, to place the lower boundary of the partition of *much greater than x* as far away as possible in the upper diagonal space. This will create a radial partition enclosing a "very wide" angle with the line of equality. Similarly the partition for *much less than x* encloses a "very wide" angle with the line of equality in the lower diagonal space. Note that the partitions generated by *greater than x* and *less than x* are supersets of the partitions induced by *much greater than x* and *much less than x* respectively.

The fuzzy expressions *slightly greater than x* and *slightly less than x* give rise to asymmetric narrow, radial partitions "very close" to the line of equality. Naturally, the partition induced by *slightly less than x* lies in the lower diagonal space. The partition induced by *slightly more than x* lies in the upper diagonal space, very close to the line of equality.

The common sense interpretation of the fuzzy expressions introduced above, must now be defined mathematically, in order to form the basis for the proposed fuzzy geometric partitions based representation of fuzzy objects and comparison of fuzzy database objects for retrieval purposes.

2.1 Specification of the Fuzzy Partitions Induced by Fuzzy Restrictions.

Fuzzy expressions, such as *about x* and *more or less x* where *x* is a number, are said to constitute elastic constraints on the set of admissible real numbers (Dubois and Prade, 1980, Kandel, 1986). An arbitrary real number **x** which satisfies the elastic constraint is said to be a *generic value* of the fuzzy expression (Kandel, 1986; Dubois and Prade, 1980).

The term *generic value* is used, in this study, to characterize crisp values lying within a vague interval or the fuzzy partition associated with a fuzzy number or fuzzy restriction. The basic idea upon which the concept of fuzzy geometric partitions is founded, is the simple, intuitive, idea that in common sense reasoning a vague expression, such as *about 5*, invokes a mental band of uncertain but narrow width around the crisp number 5 as explained in the previous section. The main assumption is that the human mind realizes this vague interval by a process in which values picked out from the domain of real numbers are subconsciously compared with the crisp value 5 and rejected if they differ "too much" from it. In this respect the vague expression, *about 5*, is equivalent to the generic binary relation, *about(x, 5),* where *x* is an arbitrary number which may or may not be an acceptable member of the vague set *about 5*, depending on its "distance" from the crisp value 5 (see Figures 2 and 6).

Let the collection of the fuzzy restrictions: *equal to, greater than, less than, much greater than, much less than, slightly greater than, slightly less than, about, more or less, roughly equal to,* be denoted by PRED.

Using the idea of the diagonal subset generated by the equality operator in Dowsing et al. (1986), the definition of the equality operator can be extended and generalized to a general

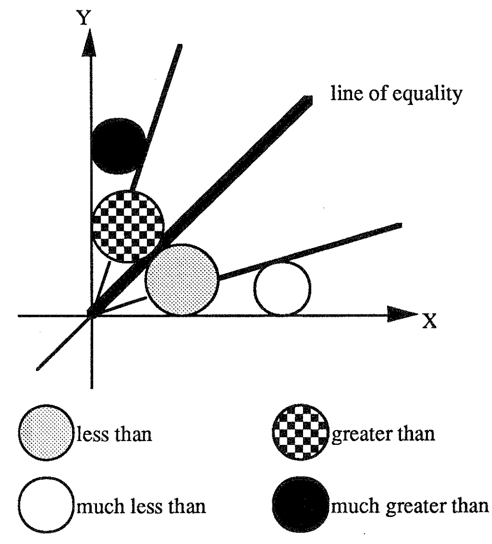


Figure 1: Basic partitions of the search space.

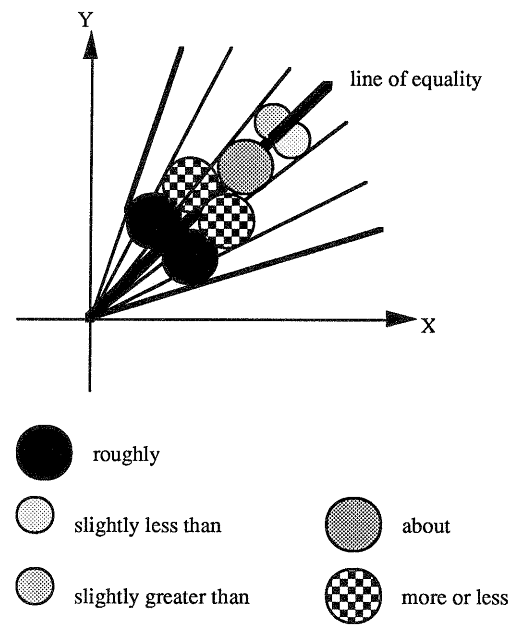


Figure 2: Partitions induced by some common fuzzy expressions.

comparison operator $R \in$ PRED. Let the partition induced by the general fuzzy comparison operator, R, in a two dimensional search space be denoted by P_R. The generalized extension of the definition of equality operator (Dowsings et al., 1986) is as follows:

For an interpretation I, with the universe X, the set R_I on which R is to be true must be a radial or sectoral subset:

$\{(x,y) \mid x \in X, y \in X, x\, R\, y\}$ of $X \times X$

Based on this generalized definition, individual members of the PRED set can now be defined. If we let R = *much greater than* we have:

$$\text{much greater than}_I = \{(x,y) \mid x \in X, y \in X, x >> y\} \text{ for all } (x,y) \in X \times X. \quad (1)$$

similarly for R = *much less than* we have:

$$much\ less\ than_I = \{(x,y) \mid x \in X,\ y \in X,\ x << y\} \text{ for all } (x,y) \in X \times X. \qquad (2)$$

where, >> and << have the usual meaning, much greater than, and much less than, respectively.

Definitions of all the relations in the PRED set can be produced by a similar process. For practical purposes it is instructive to represent the partitions in Eqs. (1) and (2) geometrically. This is achieved by transforming the definitions in Eqs. (1) and (2) into their polar counterpart. Figure 3 defines the important geometrical parameters needed for this. Thus referring to Figure 3, if the partition generated by the general comparison operator, R, is denoted by P_R it can be defined as:

$$P_R = \{[\alpha_u, \alpha_l] \mid \alpha_u \in [0, \pi/2],\ \alpha_l \in [0, \pi/2],\ \alpha_l \geq \alpha_u\} \qquad (3)$$

Now substituting the operator *equal*, for R, in Eq. (3) we get the following definition for the equality operator (Eq. 4):

$$P_= = \lim_{\alpha_u \to \pi/4,\ \alpha_l \to \pi/4} \{[\alpha_u, \alpha_l]\} = [\pi/4, \pi/4] \qquad (4)$$

The equality of the upper and lower bounds in Eq. (4) means that the equality operator partitions the search space diagonally. This definition is therefore equivalent to the one given in Dowsing et al(1986). The partitions induced by the operators *greater* and *less* are defined by Eqs. (5) and (6) below.

$$P_> = \lim_{\alpha_u \to 0,\ \alpha_l \to \pi/4-} \{[\alpha_u, \alpha_l]\} = [0, \pi/4-] \qquad (5)$$

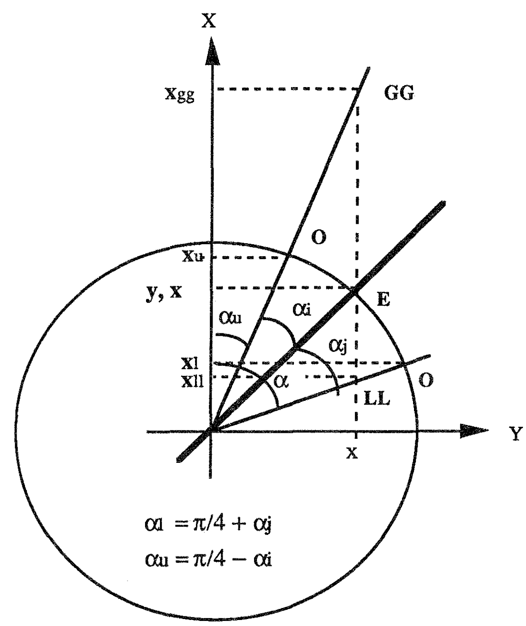
$$P_< = \lim_{\alpha_u \to \pi/4+,\ \alpha_l \to \pi/2} \{[\alpha_u, \alpha_l]\} = [\pi/4+, \pi/2] \qquad (6)$$

where $\pi/4-$ and $\pi/4+$ mean infinitesimally smaller than, and greater than, respectively.

The above strategy can be used to define the partitions corresponding to the fuzzy restrictions "much greater than" and "much less than" as given in Eqs. (7) and (8) below.

$$P_{>>} = \lim_{\substack{\alpha_u \to 0,\\ \alpha_l \to \pi/4 - \text{VERYWIDE}'}} \{[\alpha_u, \alpha_l]\} \qquad (7)$$

$$P_{<<} = \lim_{\substack{\alpha_u \to \pi/4 + \text{VERYWIDE}',\\ \alpha_l \to \pi/2}} \{[\alpha_u, \alpha_l]\} \qquad (8)$$



GG: $x = x_{gg} = x/\tan(\tan(\pi\iota/4 - \alpha_i)$ for **much more than(x)**
E: $y = x = x$, for **equal to(x)**
LL: $x = x_l = x/\tan(\pi\iota/4 + \alpha_j)$ for **much less than (x)**
O_u, O_l :
$x_u = x = x/\cos(\pi\iota/4)\cos(\pi\iota/4 - \alpha_i)$,
$x_l = x = x/\cos(\pi\iota/4)\cos(\pi\iota/4 + \alpha_j)$, for all other cases.

Figure 3: Geometrical parameters of the partitions induced by fuzzy expressions.

In Eqs. (7) and (8) the terms WIDE, VERYWIDE, and others listed in Table 1, represent a fuzzy constant whose value may be subjectively assigned to reflect individual conception of the vague expressions "verywide", "wide", etc. They may be interpreted as generic band width of the fuzzy intervals associated with the fuzzy expressions. Based on these parameters, generic band widths associated with all fuzzy expressions in PRED can be assigned as shown in Table 2.

When assigning values to the fuzzy constants in Tables 1 and 2 two important criteria must be considered:

1. It is important for the generated partitions to be compatible with common sense in accordance with the criteria for suitable characteristics of membership functions (Magrez and Smets, 1989).

2. The usefulness of fuzzy sets in modelling fuzzy concepts, class, or linguistic variables depends on the appropriateness of the selected membership functions(Kandel, 1986).

Because in this experiment membership functions are constructed from subjectively assigned partitions, the second criteria will be used to test the validity of the assigned partitions. Accordingly the partitions assigned to fuzzy variables will be deemed proper if membership functions constructed from them are similar to or comparable with those assigned by conventional fuzzy sets methods.

3. COMPUTATION OF FUZZY GENERIC VALUES AND CONSTRUCTION OF MEMBERSHIP FUNCTIONS.

When appropriately specified, fuzzy geometric partitions provide a means for performing direct comparison of fuzzy objects for data base search purposes. Alternatively fuzzy membership functions may be constructed from the assigned partitions and used to compare fuzzy objects based on existing theories of fuzzy set inclusion, equality and composition of fuzzy sets as outlined in Zadeh et al.(1975); Dubois and Prade (1980); Kandel (1986); Dubois and Prade (1988); Klir and Folger (1988); and Zadeh (1979, 1989);.

In section 2.1 the vague expression *about 5* was said to be equivalent to a binary fuzzy relation *about(x, 5)* such that x is a generic value satisfying the fuzzy restriction. Based on this the concept of the generic value of a fuzzy number may be defined as follows:

> *Let R be any fuzzy predicate in PRED, then the unary fuzzy expression R(x) where $x \varepsilon X$, is said to induce a generic value $x \varepsilon X$ such that the expression, $x = R(x)$, or equal(x, R(x)), evaluated over the universe of discourse is true.*

Using this definition the fuzzy restriction *greater than 5* has a corresponding binary fuzzy relation *greater_than(x, 5)* where x is a generic value satisfying the fuzzy restriction.

To characterize generic values in a mathematically meaningful way, tentative values for the parameters α_i and α_j defined in Figure 3 are given in Table 2. The angular parameters α_i and α_j represent the band width of left and right tailed fuzzy sets(Dubois and Prade, 1980) respectively. For symmetric fuzzy sets, α_i and α_j represent the left and right half-band widths. The term band width is used in the same sense as it is used to characterize standard membership functions(Kandel, 1986).

Notice that in Table 2 the parameters α_i and α_j are assigned values by logarithmically weighting the fuzzy constants defined in Table 1. This is necessary to preserve the fact that perception of changes in numerical magnitudes vary as the difference between the numbers involved change from very small to very large. It is intuitive to use logarithmic weighting since logarithmic functions are also used in modelling image intensities in natural vision, photography and image processing to reflect the human physiological response to increasing light stimulus (Land et al, 1989).

Based on the values in Table 2, functions for computing arbitrary generic values for the fuzzy expressions in the PRED set can be derived. These functions are summarized in Table 3.

3.1 Comparison of Fuzzy Objects Using Generic Values.

The equations required for computing upper and lower bounding generic values for all the fuzzy predicates in the PRED set are summarized in Table 3. Generic values computed by these equations can be used to facilitate direct comparison of fuzzy objects for the purposes of database searching. For example the query object *more or less x* can be interpreted as a request to retrieve all database objects satisfying the elastic constraint *more_or_less(y,x)*. Valid generic objects y must therefore have values lying close to the crisp value x. This condition may be expressed as

$$x_u \geq y \geq x_l \quad (9)$$

where x_u and x_l are generic values corresponding to the upper and lower bounds of the partition induced by the fuzzy restriction *more_or_less(x)*. The values of x_u and x_l can be computed from Eqs. (10) and (11) respectively, where the term CLOSE' is as defined in Table 2.

Table 1: DEFINITION OF THE FUZZY CONSTANTS DETERMINING THE BAND WIDTH OF THE FUZZY PATITIONS.

LABEL	VERYWIDE	WIDE	CLOSE1	CLOSE	VERYCLOSE
GENERIC WIDTH	$\pi/3$	$\pi/6$	$\pi/12$	$\pi/24$	$\pi/48$

Table 2: FORMULAE FOR COMPUTING THE GENERIC BAND WIDTH OF THE PARTITIONS.

Fuzzy Predicate	Left and Right Tails:	Symbolic Value
much_greater_than(x)	WIDE+VERYCLOSE/(4+log(x))	VERYWIDE'
much_less_than(x)	WIDE+VERYCLOSE/(4+log(x))	VERYWIDE'
slightly_more_than(x)	3*VERYCLOSE/(1+0.5log(x))	VERYCLOSE'
slightly_less_than(x)	3*VERYCLOSE/(1+0.5log(x))	VERYCLOSE'
more_or_less(x)	5*CLOSE1/(8+log(x)) if x<=5 CLOSE1/(2+log(x)) if x<=10 CLOSE1/(3+log(x)) if x>10	CLOSE1'
about(x)	4*CLOSE/(3+3log(x)) if x<=5 4*CLOSE/(6+3log(x)) if x<=10 4*CLOSE/(9+3log(x)) if x>10	CLOSE'
roughly(x)	CLOSE1/(1+log(x)) if x <= 5 CLOSE1/(2+log(x)) if x <=10 CLOSE1/(3+log(x)) if x >10	CLOSE"

The meaning of the trigonometric expressions appearing in Eqs. (10) and (11) is obvious from Figure 3. The condition in Eq. (9) may now be used by the query processor for the approximate selection of database objects which satisfy the fuzzy query.

$$x_u = \frac{x}{\cos\left(\frac{\pi}{4}\right)} \cos\left(\frac{\pi}{4} - \text{CLOSE}'\right) \quad (10)$$

$$x_l = \frac{x}{\cos\left(\frac{\pi}{4}\right)} \cos\left(\frac{\pi}{4} + \text{CLOSE}'\right) \quad (11)$$

Thus by means of the concept of fuzzy partitions and generic values the original fuzzy query is transformed into an interval comparison problem in which the interval bounds correspond to the *cut-out points*, *peak points*, *turnover points*, or any other desirable characteristic points of the membership function of the fuzzy object. It must however be noted that manipulation of fuzzy numeric data by interval arithmetic can only be tolerated for low accuracy requirements (Kandel, 1986, Klir and Folger, 1988).

3.2 Construction of Fuzzy Membership Functions From Geometric Fuzzy Partitions.

Construction of membership functions is a necessary step towards verification and validation of the assigned partitions in line with the second criteria in section 2.1. In addition the membership functions are useful in themselves, as theoretically well founded tools for representing and querying fuzzy knowledge(Schmucker, 1984; Dubois and Prade, 1988; Kandel, 1986).

The construction of membership functions must be done subject to certain desirable characteristics of membership functions (Zadeh et al. 1975; Kandel, 1986; Klir and Folger, 1988):

(i) The membership function must map the set of objects in the universe of discourse into the interval [0,1].

(ii) The membership function must satisfy necessary fuzzy set theoretic properties with respect to fuzzy union, intersection, complementation etc.

Let R(x) represent a general fuzzy predicate or fuzzy comparison operator in the PRED set, where $x \in X$ is some crisp numeric value in the real numbers universe. If the generic value induced by R on x is denoted by **x**, and **x** is selected such that it is a bounding value, then it lies on the boundary of the partition generated by R in the universe of discourse.

Denote the amount by which R "stretches" x by D. Then $D = |\mathbf{x} - x|$ represents the width of the partition generated by R in $X \times X$. The set $\{\mathbf{x} \mid \mathbf{x} = R(x)\}$, for all $x \in X$, defines the partition induced by R. Let **y** represent some crisp value (or the crisp generic value of a fuzzy number) in the search space. Denote the difference between **y** and x by **d**. Then by set membership definition (Klir and Folger, 1988), **y** is contained in the partition induced by R(x) if the condition(see Figure 4):

$$d = |y - x| < D \quad (12)$$

is satisfied. This condition means that **y** falls within the "stretch" of R(x).

To provide a fuzzy set theoretic basis for the comparison of fuzzy values, membership functions for the general fuzzy restriction, R(x), may be constructed by the following procedure(Figure 4):

1. Set the width D of the partition generated by R(x) on the v-axis (horizontal axis) as shown in Figure 4.
2. Draw a line of unit length along the u-axis(vertical axis).
3. Link the end of the unit line with point D on the v-axis.
4. Plot the distance d, of **y** from x, along the v-axis.
5. Mirror project d perpendicularly on to the u-axis and denote its image by μ_R.
6. The distance, μ_R, of the projection point is proportional to the strength of the membership of x_1 in the fuzzy set represented by R(x). It may therefore be regarded as a first approximation to its fuzzy membership value.
7. Letting **y** cover the range of all values in X modify μ_R by applying intensification, dilation, normalization, concentration (Schmucker, 1984; Kandel, 1986), or any other fuzzy set theoretic transformation function, F, to arrive at a suitable shape of the membership function.

From Figure 4 an approximate formula for computing the fuzzy membership value is obtained(Eq. 13).

$$\mu_R = \frac{d}{D} \quad (13)$$

To enforce the condition that fuzzy membership values must be in the range [0,1] Eq.(13) is rewritten as

$$\mu_R(y) = \begin{cases} 0, & \text{if } \mathbf{y} \neq R(x) \\ F\left(\frac{d}{D}\right), & \text{if } \mathbf{y} = R(x) \end{cases} \quad (14)$$

Table 3: COMPUTATION OF GENERIC VALUES FROM GENERIC BAND WIDTH.

Fuzzy Predicate	Generic Value Equation
much_greater_than(x)	x" = x/tan(pi/4 - VERYWIDE')
much_less_than(x)	x" = x/tan(pi/4 + VERYWIDE')
slightly_more_than(x)	x" = x/cos(pi/4)cos(pi/4 - VERYCLOSE')
slightly_less_than(x)	x" = x/cos(pi/4)cos(pi/4 + VERYCLOSE')
more_or_less(x)	x" = x/cos(pi/4 +- CLOSE1')
about(x)	x" = x/cos(pi/4 +- CLOSE')
roughly_equal_to(x)	x" = x/cos(pi/4 +- CLOSE")
more_than(x)	x" = x + D; D > 0 real number
less_than(x)	x" = x + D; D < 0 real number

where F is some appropriate function.

Practical membership functions for the fuzzy constraints in the PRED set can now be obtained by substituting for D and d based on the equations of the generic values defined in Table 3. For example to get the membership function for the fuzzy expression "much more than x", the width of the partition induced by the fuzzy restriction is computed from

$$D = \mathrm{abs}(\mathbf{x} - x) \tag{15}$$

where $\mathbf{x}$ is a generic value lying on the boundary of the partition induced by *much_more_than(x)*, and its value can be computed using the equations in Table 3 as

$$\mathbf{x} = \frac{x}{\tan\left(\frac{\pi}{4} - \mathrm{VERYWIDE'}\right)} \tag{16}$$

From Table 3 VERYWIDE' is defined as

$$\mathrm{VERYWIDE'} = \mathrm{WIDE} + \frac{\mathrm{VERYCLOSE}}{(4 + \log x)} \tag{17}$$

and after appropriate substitutions and simplification we get

$$\mathbf{x} = \frac{x}{\tan\left(\frac{\pi}{4} - \frac{\pi}{48}\left(8 + \frac{1}{4 + \log x}\right)\right)} \tag{18}$$

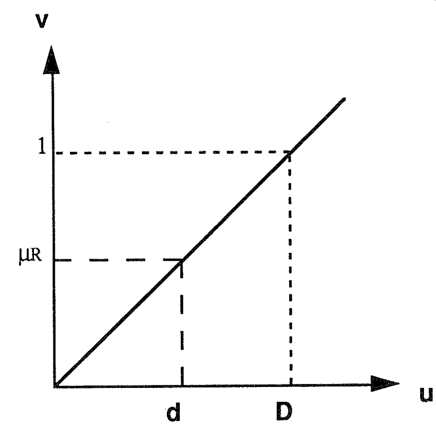
The problem of determining the compatibility between $\mathbf{y}$ and *much_more_than(x)* is equivalent to the problem of finding the membership value of $\mathbf{y}$ in the fuzzy set induced by the fuzzy restriction "much more than" on x. From Eq. (14) an approximate value of the membership of $\mathbf{y}$ in the fuzzy relation R(x) is given by Eq. (19).

$$\mu_R(y) = \frac{\mathrm{abs}(\mathbf{y} - x)}{\mathrm{abs}(\mathbf{x} - x)} \tag{19}$$

Assuming a square compatibility function for both *much_more _than(x)* and *much_less_than(x)* their membership functions are given by Eqs. (20) and (21). The plot of the membership function for *much_less_than (5)* is shown in Figure 5.

$$\mu_{>>x}(y) = \begin{cases} 0, & \text{if } \mathbf{x} \le x \\ \left(\frac{\mathbf{y} - x}{\mathbf{x} - x}\right)^2, & \text{if } \mathbf{x} > \mathbf{y} > x \\ 1, & \text{else where} \end{cases} \tag{20}$$

$$\mu_{<<x}(y) = \begin{cases} \left(\frac{\mathbf{y} - x}{\mathbf{x} - x}\right)^2, & \text{if } \mathbf{x} < \mathbf{y} < x \\ 0, & \text{if } \mathbf{y} \ge x \\ 1, & \text{elsewhere} \end{cases} \tag{21}$$



$\mathbf{d} = \mathbf{y} - x$

$\mathbf{D} = \mathbf{x} - x; \ \mathbf{x} = R(x)$

Figure 4: Construction of membership function from fuzzy partitions.

After substituting the expression for $\mathbf{x}$ in Eq. (20) using Eq. (16) we get

$$\mu_{>>x}(y) = \begin{cases} 0, & \text{if } \mathbf{x} \le x \\ \left(\frac{\mathbf{y} - x}{\frac{x}{\tan\left(\frac{\pi}{4} - \frac{\pi}{48}\left(8 + \frac{1}{4 + \log x}\right)\right)} - x}\right)^2, & \text{if } \mathbf{x} > \mathbf{y} > x \\ 1, & \text{else where} \end{cases} \tag{22}$$

as the membership function for *much_more_than(x)*.

Using this approach membership functions can be constructed for all the fuzzy predicates in the PRED set. A Graphical representation of the fuzzy membership functions for *about(5)*, *more_or_less(5)*, and *roughly_equal_to(5)* are shown in Figures 7 to 8. It is clear by looking at the shapes of the

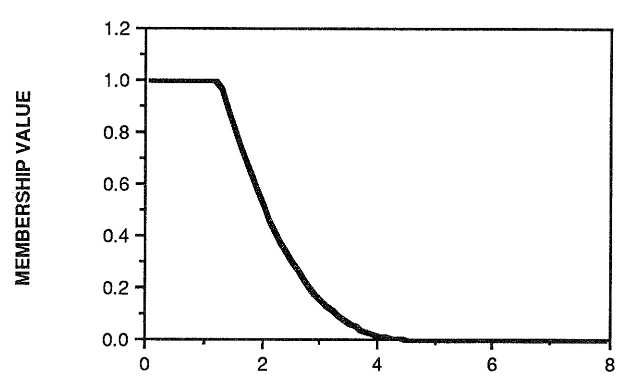


Figure 5: Membership Function for **much less than 5**

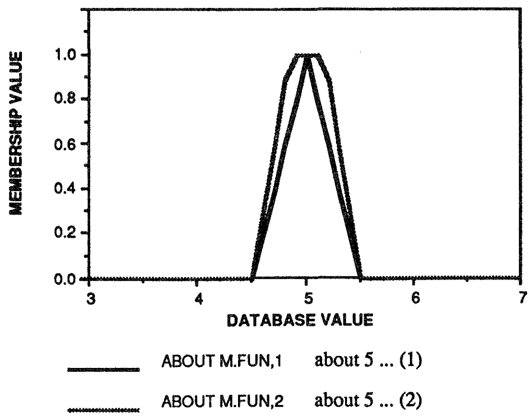


Figure 6: Membership Function for **about 5**; Two interpretations.

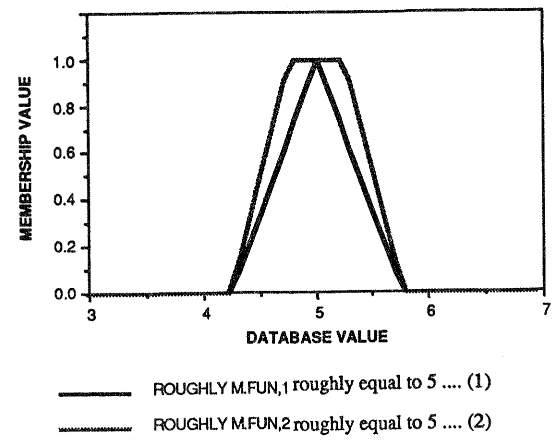


Figure 8: Membership function for **roughly equal to 5**; Two interpretations.

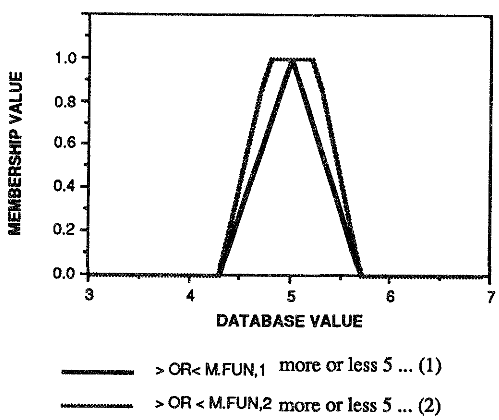


Figure 7: Membership function for **more or less 5**; Two interpretations.

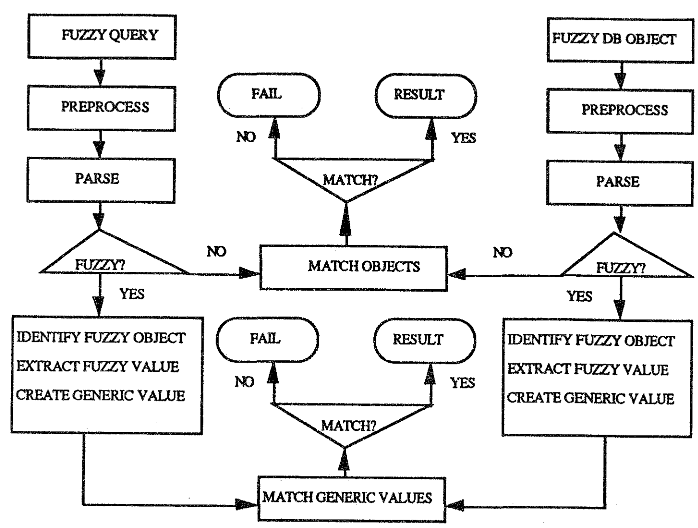


Figure 9: Design of the FUZZ subsystem for processing fuzzy query.

membership functions that the fuzzy partitions based representation of fuzzy numeric expressions produces membership functions similar to those produced by using standard membership functions. For comparison purposes reference can be made to Zadeh et al. (1975), Mizumoto and Tanaka (1975), Dubois and Prade (1980), Chen (1985), Kandel (1986) and Dubois and Prade (1988).

3.3 Fuzzy Database Query Using Generic Values Derived From the Fuzzy Geometric Partitions.

Fuzzy database retrieval requires the solution of two different problems. The first problem arises when a vague, or fuzzy, query is placed to a database containing precise, well defined, data or facts. The second problem concerns retrieval of precise queries placed to a fuzzy database. The concept of the generic value of a fuzzy numeric expression, introduced at the beginning of section 3, has been used to design a fuzzy comparison operator capable of solving the two problems, and therefore, capable of fuzzy database retrieval.

The design of the operator, called FUZZ(Mtalo, 1990), is shown in Figure 9. Using this operator both the fuzzy query and fuzzy database object are parsed into atomic fuzzy components which are then transformed into generic values for database matching purposes. If the query and database object currently being examined are both non-fuzzy the operator uses a simple database matching procedure, otherwise the operator generates and compares upper and lower bounding generic values to determine the matching object(s).

To ensure that borderline cases are not rejected offhand, the crisp interval represented by the upper and lower bounding generic values is "fuzzified" by allowing values *slightly greater* or *slightly less* than the computed generic values into the set of possible database objects. This allows queries not precisely matching the specifications of database objects to be processed.

The retrieval of fuzzy queries based on the concept of generic values must be regarded as an approximate method since it resorts to interval comparison. However vague user queries reflect a degree of uncertainty about what the user wishes to retrieve. Similarly fuzzy data reflects uncertainty about the information stored in the database. Under these circumstances the proposed method provides a simple and convenient way to represent and manipulate the uncertainty inherent in vague queries and fuzzy database information. Where more accuracy in the representation of uncertainty is needed fuzzy operators, based on the fuzzy set theory, must be used as elucidated in Dubois and Prade (1988), Kandel (1986) and other literature.

3.4 Fuzzy Database Query Using Membership Functions Based on the Fuzzy Geometrical Partitions.

In order to show the usefulness of the constructed membership functions in fuzzy query processing the new method was tested on a simulated database to determine value of the complex query objects "much greater than 5 and much less than 40" (Figure 10) and "slightly less than 5 or slightly more than 5 and roughly equal to 5" (Figure 11).

It must be noted that these kinds of queries cannot, in general, be solved by conventional database management systems because they require the interpretation of the fuzzy expressions "much greater than 5", "much less than 5", "slightly less than 5", "slightly more than 5", and "roughly equal to 5". However using a fuzzy query processor the membership functions for these fuzzy expressions can be constructed and used as the basis for selecting valid database objects.

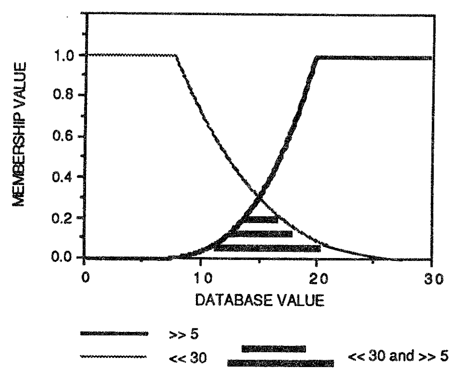


Figure 10: Selection of database values for the fuzzy query **much more than 5 and much less than 40** based on non-zero membership values.

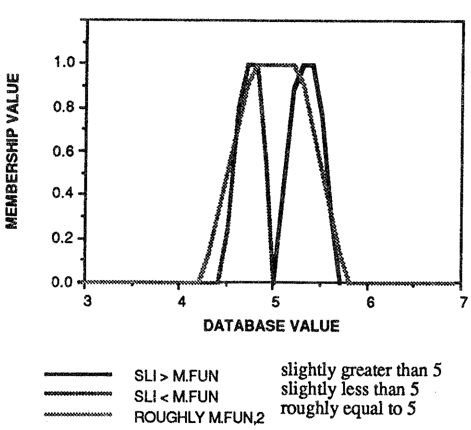


Figure 11: Fuzzy query selection for **slightly less than 5 or slightly more than and roughly equal to 5** using the intersection of the membership functions of the fuzzy query components.

The selection of database objects satisfying complex fuzzy queries corresponds to the intersection of the fuzzy membership functions as shown graphically in Figures 10 and 11 for the fuzzy queries discussed above. In fuzzy sets literature the process of determining the (crisp) object satisfying the condition set out by the fuzzy sets intersection is referred to as *defuzzification*. One defuzzification strategy involves selection of the object corresponding to the centroid of the intersected region(s).

In General once the fuzzy membership functions are available standard fuzzy set theoretic operators can be used to solve the fuzzy query (see Kandel, 1986; Magrez and Smets, 1989; Zadeh, 1989).

4. CONCLUDING REMARKS.

A new method for representing and manipulating fuzzy information has been formulated and tested. It has been amply illustrated in the paper that this method differs substantially from the traditional approach of representing fuzzy linguistic variables. In particular this method makes it possible to define the meaning of fuzzy numerical restrictions in the domain of discourse by fuzzy geometric partitions of the search space.

By means of geometric fuzzy partitions generic values of fuzzy numerical expressions can be constructed and used to facilitate comparison of fuzzy objects for database retrieval purposes. Alternatively it has been shown that geometric fuzzy partitions can form the basis for constructing reasonable membership functions for characterizing fuzzy numeric expressions.

The usefulness of fuzzy membership functions generated from fuzzy geometric partitions in fuzzy query processing was illustrated for a model database. In addition a fuzzy comparison operator based on the fuzzy geometrical partitioning of the search space was designed and implemented in an experimental knowledge based system.

REFERENCES

Chen, S. (1985). "Ranking fuzzy numbers with maximizing set and minimizing set," *Fuzzy Sets and Systems*, Vol.17, No.2, pp. 113-129.

Civanlar, M.R.; H.J. Trusell(1986). "Constructing membership functions using statistical data," *Fuzzy Sets and Systems*, 18 (1986) 1-13.

Dowsing, R.; V.J. Rayward-Smith; C.D. Walter (1986). *A First Course in Formal Logic and Its Applications in Computer Science*. Blackwell Scientific Publications, Oxford, London, Edinburg.

Dubois D.; H. Prade (1980). Fuzzy Sets and Systems: *Theory and Applications, Academic Press, N.Y., London.* Academic Press, N.Y., London.

Dubois D.; H. Prade (1988). *Possibility theory: An Approach to Computerized Processing of Uncertainty*, Trans. by E.F. Harding, Plenum Press, N.Y. London.

Dubois D.; H. Prade (1989). "Processing fuzzy temporal knowledge," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol.19, No.4, July/Aug., pp.729-743.

Kandel, A. (1986). *Fuzzy Mathematical Techniques With Applications*, Adison-Wesley Publishing Company, Reading, Massachusets.

Kaufmann, A. (1975). "Fuzzy graphs and fuzzy relations," In: A. Kaufmann *Introduction to the Theory of Fuzzy Subsets: Fundamental Theoretical Elements*, Vol. I., Academic Press, N.Y. San Fransisco, London.

Kaufmann, A.; M.M. Gupta (1988). *Fuzzy Mathematical Models in Engineering and Management Science*, Elsevier Science Publishers, North-Holland, Amsterdam, N.Y.

Klir, G.J.; T.A. Folger (1988). *Fuzzy Sets, Uncertainty and Information*, Prentice Hall, Engelwood Cliffs, N.J.

Magrez, P.; P. Smelt (1989). "Fuzzy modus ponens: A new model suitable for applications in knowledge-based systems," *International Journal of Intelligent Systems*, Vol. 4 pp. 181-200.

de Michiel, L.G. (1989). "Resolving database incompatibility: An approach to performing relational operations over mismatched domains," *IEEE Transactions on Knowledge and Data Engineering,* Vol. 1. No. 4, December, pp. 485-493.

Mizumoto, M., K. Tanaka (1979). "Some properties of fuzzy numbers," In: M.M. Gupta; R.K. Ragade; R.R. Yager (Eds.) *Advances in Fuzzy Sets Theory and Applications*, North-Holland Publishing Company, Amsterdam, pp. 153-164.

Mtalo, E.G. (1990). *SLEMS: A Knowledge Based Approach to Soil Loss Estimation and Modelling*, Msc. Thesis, UNB, Fredericton.

Schmucker, K.J. (1984). *Fuzzy Sets Natural Language Computations, and Risk Analysis,* The Computer Science Press Inc.

Zadeh, L.A.; K-S. Fu; K. Tanaka; M. Shimura (1975). *Fuzzy Sets and Their Applications to Cognitive and Decision Processes*, Proceedings of the US-Japan Seminar on Fuzzy Sets and Their Applications, The University of California Berkeley, July 1-4 1974, Academic Press Inc., N.Y.

Zadeh, L. A. (1989). "Knowledge representation in fuzzy logic," *IEEE Transactions on Knowledge and Data Engineering*, Vol.1, No.1, March, pp. 3-18.