Chapter I
Introduction and Trends to Fuzzy Logic and Fuzzy Databases

José Galindo
University of Málaga, Spain

ABSTRACT

This chapter presents an introduction to fuzzy logic and to fuzzy databases. With regard to the first topic, we have introduced the main concepts in this field to facilitate the understanding of the rest of the chapters to novel readers in fuzzy subjects. With respect to the fuzzy databases, this chapter gives a list of six research topics in this fuzzy area. All these topics are briefly commented on, and we include references to books, papers, and even to other chapters of this handbook, where we can find some interesting reviews about different subjects and new approaches with different goals. Finally, we give a historic summary of some fuzzy models, and we conclude with some future trends in this scientific area.

INTRODUCTION

Fuzzy logic is only a mathematical tool. It is possibly the best tool for treating uncertain, vague, or subjective information. Just to give an idea about the importance of this soft computing tool, we can mention the big quantity of publications in this field, including two research journals of great quality: *Fuzzy Sets and Systems*[1] and *IEEE Transactions on Fuzzy Systems*.[2] Particularly, fuzzy logic has been applied to databases in many scientific papers and real applications. Undoubtedly, it is a modern research field and it has a long road ahead. This handbook is only one step. Perhaps, it is a big step.

For that reason, we will begin by introducing some basic concepts of the fuzzy sets theory. We include definitions, examples, and useful tables with reference data (for example, lists of t-norms, t-conorms, and fuzzy implications). We can find these and other concepts in other chapters of this book, possibly with different notation. The second part of this chapter studies basic concepts about fuzzy databases, including a list of six research topics on fuzzy databases. All these topics are briefly commented on, and we include references to books, papers, and even to other chapters of this handbook. Then, an overview about the basic fuzzy database models is included to give an introduction to these topics and also to the whole handbook.

FUZZY SETS

In written sources, we can find a large number of papers dealing with this theory, which was first introduced by Lotfi A. Zadeh[3] in 1965 (Zadeh, 1965). A compilation of some of the most interesting articles published by Zadeh on the theme can be found in Yager, Ovchinnikov, Tong, and Nguyen (1987). Dubois and Prade (1980, 1988) and Zimmerman (1991) bring together the most important aspects behind the theory of fuzzy sets and the theory of possibility. A more modern synthesis of fuzzy sets and their applications can be found in Buckley and Eslami (2002); Kruse, Gebhardt, and Klawonn (1994); Mohammd, Vadiee, and Ross (1993); Nguyen and Walker (2005); and Piegat (2001), and particularly in Pedrycz and Gomide (1998). Ross (2004) includes some engineering applications and Sivanandam, Sumathi, and Deepa (2006) present an introduction using MATLAB. A complete introduction in Spanish is given in Escobar (2003) and Galindo (2001).

The original interpretation of fuzzy sets arises from a generalization of the classic concept of a subset extended to embrace the description of "vague" and "imprecise" notions. This generalization is made considering that the membership of an element to a set becomes a "fuzzy" or "vague" concept. In the case of some elements, it may not be clear if they belong to a set or not. Then, their membership may be measured by a degree, commonly known as the "membership degree" of that element to the set, and it takes a value in the interval [0,1] by agreement.

Using classic logic, it is only possible to deal with information that is totally true or totally false; it is not possible to handle information inherent to a problem that is imprecise or incomplete, but this type of information contains data that would allow a better solution to the problem. In classic logic, the membership of an element to a set is represented by 0 if it does not belong and by 1 if it does, having the set {0,1}. On the other hand, in fuzzy logic, this set is extended to the interval [0,1]. Therefore, it could be said that fuzzy logic is an extension of the classic systems (Zadeh, 1992). Fuzzy logic is the logic behind approximate reasoning instead of exact reasoning. Its importance lies in the fact that many types of human reasoning, particularly the reasoning based on common sense, are by nature approximate. Note the great potential that the use of membership degrees represents by allowing something qualitative (fuzzy) to be expressed quantitatively by means of the membership degree. A fuzzy set can be defined more formally as:

Definition 1: ***Fuzzy set*** *A over a universe of discourse* X *(a finite or infinite interval within which the fuzzy set can take a value) is a set of pairs:*

$$A = \{\mu_A(x)/x : x \in \mathrm{X},\ \mu_A(x) \in [0,1] \in \Re\} \quad (1)$$

where $\mu_A(x)$ *is called the* ***membership degree*** *of the element x to the fuzzy set A. This degree ranges between the extremes* **0** *and* **1** *of the dominion of the real numbers:* $\mu_A(x) = 0$ *indicates that x in no way belongs to the fuzzy set A, and* $\mu_A(x) = 1$ *indicates that x completely belongs to the fuzzy set A. Note that* $\mu_A(x) = 0.5$ *is the greatest uncertainty point.*

Sometimes, instead of giving an exhaustive list of all the pairs that make up the set (discreet values), a definition is given for the function $\mu_A(x)$*, referring to it as characteristic function or* ***membership function****.*

The universe X may be called *underlying universe* or underlying domain, and in a more generic way, a fuzzy set *A* can be considered a function μ_A that matches each element of the universe of discourse X with its membership degree to the set *A*:

$$\mu_A(x): \quad \mathrm{X} \rightarrow [0,1] \quad (2)$$

The universe of discourse X, or the set of considered values, can be of these two types:

- Finite or discrete universe of discourse $\mathrm{X} = \{x_1, x_2, \ldots, x_n\}$, where a fuzzy set *A* can be represented by:

$$A = \mu_1 / x_1 + \mu_2 / x_2 + ... + \mu_n / x_n \qquad (3)$$

where μ_i with $i = 1, 2, ..., n$ represents the membership degree of the element x_i. Normally, the elements with a zero degree are not listed. Here, the + does not have the same significance as in an arithmetical sum, but rather, it has the meaning of aggregation, and the / does not signify division, but rather the association of both values.

- Infinite universe of discourse, where a fuzzy set A over X can be represented by:

$$A = \int \mu_A(x) / x \qquad (4)$$

Actually, the membership function $\mu_A(x)$ of a fuzzy set A expresses the degree in which x verifies the category specified by A.

A **linguistic label** is that word, in natural language, that expresses or identifies a fuzzy set, that may or may not be formally defined. With this definition, we can assure that in our every day life we use several linguistic labels for expressing abstract concepts such as "young," "old," "cold," "hot," "cheap," "expensive," and so forth. Another interesting concept, the linguistic variable (Zadeh, 1975), is defined in the chapter by Xexéo and Braga in this handbook. Basically, a linguistic variable is a variable that may have fuzzy values. A linguistic variable is characterized by the name of the variable, the underlying universe, a set of linguistic labels, or how to generate these names and their definitions. The intuitive definition of the labels not only varies from one to another person depending on the moment, but also it varies with the context in which it is applied. For example, a "high" person and a "high" building do not measure the same.

Example 1: The "Temperature" is a linguistic variable. We can define four linguistic labels, like "Very_Cold," "Cold," "Hot," and "Very_Hot," using the membership functions depicted in Figure 1.

The frame of cognition, or frame of knowledge, is the set of labels, usually associated to normalized fuzzy sets (Definition 11), used as reference points for fuzzy information processing.

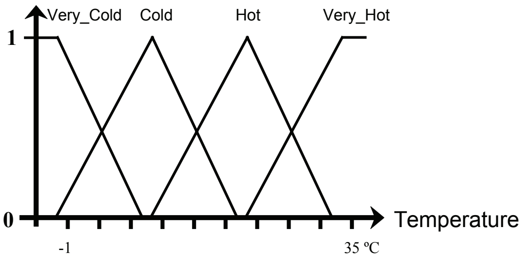
Characteristics and Applications

This logic is a multivalued logic, the main characteristics of which are (Zadeh, 1992):

- In fuzzy logic, exact reasoning is considered a specific case of approximate reasoning.
- Any logical system can be converted into terms of fuzzy logic.
- In fuzzy logic, knowledge is interpreted as a set of flexible or fuzzy restrictions over a set of variables (e.g., the variable Temperature is Cold).
- Inference is considered as a process of propagation of those restrictions. Inference is understood to be the process by which a result is reached, consequences are obtained, or one fact is deduced from another.
- In fuzzy logic, everything is a matter of degree.

From this simple concept, a complete mathematical and computing theory has been developed which facilitates the solution of certain problems (see the references in the beginning of this chapter). Fuzzy logic has been applied to a multitude of disciplines such as control systems, modeling, simulation, prediction, optimization, pattern rec-

Figure 1. A frame of cognition with four linguistic labels for temperature (Example 1)



ognition (e.g., word recognition), information or knowledge systems (databases, knowledge management systems, case-based reasoning systems, expert systems, etc.), computer vision, biomedicine, picture processing, artificial intelligence, artificial life, and so forth. Summarizing, fuzzy logic may be an interesting tool where hitherto known methods fail, highlighting complex processes, where we need to introduce the expert knowledge from experienced people, or where there are unknown magnitudes or ones that are difficult to measure in a reliably way. In general, fuzzy logic is used when we need to represent and operate with uncertain, vague, or subjective information. Many applications use fuzzy logic with other general or soft computing tools like genetic algorithms (GAs), neural networks (NNs), or rule based systems.

Membership Functions

Zadeh proposed a series of membership functions that could be classified into two groups: those made up of straight lines, or "linear," and Gaussian forms, or "curved." We will now go on to look at some types of membership functions. These types of fuzzy sets are those known as convex fuzzy sets in fuzzy set theory, with the exception of that known as extended trapezium that does not necessarily have to be convex, although for semantic reasons, this property is always desirable.

- **Triangular (Figure 2):** Defined by its lower limit a**,** its upper limit b, and the modal value m, so that $a<m<b$. We call the value $b - m$ *margin* when it is equal to the value $m - a$.

$$A(x) = \begin{cases} 0 & \text{if } x \le a \text{ or } x \ge b \\ (x-a)/(m-a) & \text{if } x \in (a, m] \\ (b-x)/(b-m) & \text{if } x \in (m, b) \end{cases} \tag{5}$$

- **Singleton (Figure 3):** It takes the value zero in all the universe of discourse except in the point m where it takes the value 1. It is the representation of a nonfuzzy (crisp) value.

$$SG(x) = \begin{cases} 0 & \text{if } x \ne m \\ 1 & \text{if } x = m \end{cases} \tag{6}$$

- **L Function (Figure 4):** This function is defined by two parameters, a and b, in the following way, using linear shape:

$$L(x) = \begin{cases} 1 & \text{if } x \le a \\ \dfrac{b-x}{b-a} & \text{if } a < x < b \\ 0 & \text{if } x \ge b \end{cases} \tag{7}$$

- **Gamma Function (Figure 5):** It is defined by its lower limit a and the value $k > 0$. Two definitions:

$$\Gamma(x) = \begin{cases} 0 & \text{if } x \le a \\ 1 - e^{-k(x-a)^2} & \text{if } x > a \end{cases} \tag{8}$$

$$\Gamma(x) = \begin{cases} 0 & \text{if } x \le a \\ \dfrac{k(x-a)^2}{1 + k(x-a)^2} & \text{if } x > a \end{cases} \tag{9}$$

  - This function is characterized by rapid growth starting from a.
  - The greater the value of k**,** the greater the rate of growth.
  - The growth rate is greater in the first definition than in the second.
  - Horizontal asymptote in 1.
  - The gamma function is also expressed in a linear way (Figure 5b):

$$\Gamma(x) = \begin{cases} 0 & \text{if } x \le a \\ \dfrac{x-a}{b-a} & \text{if } a < x < b \\ 1 & \text{if } x \ge b \end{cases} \tag{10}$$

- **S Function (Figure 6):** Defined by its lower limit a, its upper limit b, and the value m or point of inflection so that $a<m<b$. A typical value is: $m = (a+b)/2$. Growth is slower when the distance $a–b$ increases.

Figure 2. Triangular fuzzy sets: (a) General, (b) symmetrical

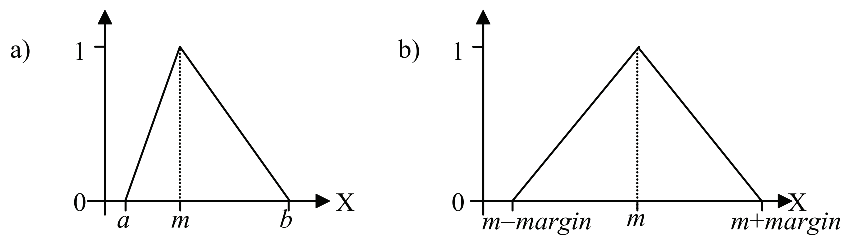


Figure 3. Singleton fuzzy set

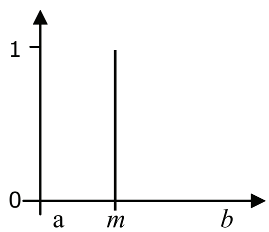


Figure 4. L fuzzy set

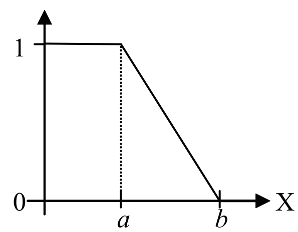
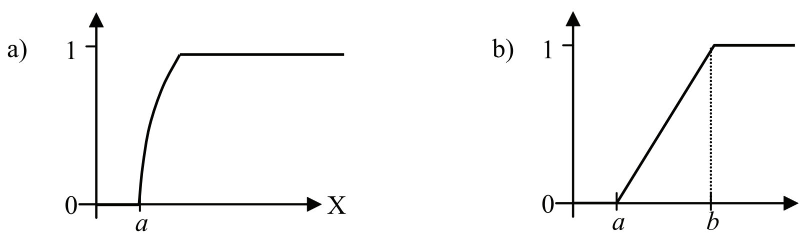


Figure 5. Gamma fuzzy sets: (a) General, (b) linear



$$S(x) = \begin{cases} 0 & \text{if } x \le a \\ 2\{(x-a)/(b-a)\}^2 & \text{if } x \in (a,m] \\ 1-2\{(x-b)/(b-a)\}^2 & \text{if } x \in (m,b) \\ 1 & \text{if } x \ge b \end{cases} \qquad (11)$$

- **Trapezoid Function (Figure 7):** Defined by its lower limit *a*, its upper limit *d*, and the lower and upper limits of its nucleus or kernel, *b* and *c*, respectively:

$$T(x) = \begin{cases} 0 & \text{if } (x \le a) \text{ or } (x \ge d) \\ (x-a)/(b-a) & \text{if } x \in (a,b) \\ 1 & \text{if } x \in [b,c] \\ (d-x)/(d-c) & \text{if } x \in (c,d) \end{cases} . \qquad (12)$$

Figure 6. S fuzzy set

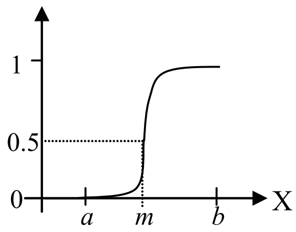


Figure 7. Trapezoidal fuzzy set

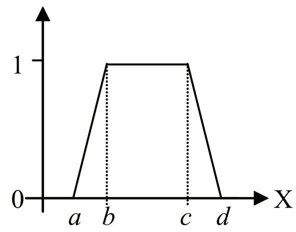


Figure 8. Gaussian fuzzy set

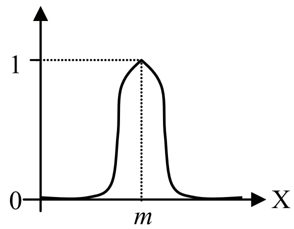
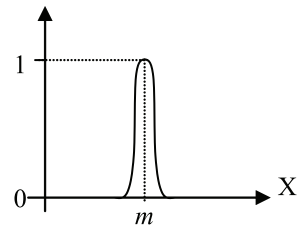


Figure 9. Pseudo-exponential fuzzy set



- **Gaussian Function (Figure 8):** This is the typical Gauss bell, defined by its midvalue *m* and the value of *k*>0. The greater *k* is, the narrower the bell.

$$G(x) = e^{-k(x-m)^2} \quad (13)$$

- **Pseudo-Exponential Function (Figure 9):** Defined by its midvalue *m* and the value *k*>1. As the value of *k* increases, the growth rate increases and the bell becomes narrower.

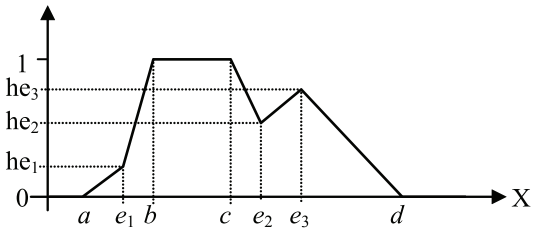
$$P(x) = \frac{1}{1 + k(x-m)^2} \quad (14)$$

- **Extended Trapezoid Function (Figure 10):** Defined by the four values of a trapezoid (*a*,*b*,*c*,*d*), and a list of points between a and b, and/or between c and d, with their membership values (height) associated to each of these points (e_i, he_i).

Comments:

- In general, the **trapezoid** function adapts quite well to the definition of any concept in human contexts, with the advantage that it is easy to define, easy to represent, and simple to calculate.
- In specific cases, the **extended trapezoid** is very useful. This allows greater expressiveness through increased complexity.
- In general, the use of a **more complex function** is usually difficult to define with precision and probably it **does not give increased precision**, as we must keep in mind that we are defining a **fuzzy** concept.
- Concepts that require a **nonconvex function** can be defined. In general, a nonconvex function expresses **the union**

Figure 10. Extended trapezoidal fuzzy set



of two or more concepts, the representation of which is convex.

In fuzzy control, for example, the aim is to express the notions of "increase," "decrease," and "approximation," and in order to do this, the types of membership functions previously mentioned are used. The membership functions Gamma and S would be used to represent linguistic labels such as "tall" or "hot" in the dominion of height and temperature. Linguistic labels, such as "small" and "cold," would be expressed by means of the L function. On the other hand, approximate notions are sometimes difficult to express with one word. In the dominion of temperature, it would be "comfortable" or "approximately 20ºC," which would be expressed by means of the triangle, trapezoid, or gaussian function.

Concepts about Fuzzy Sets

In this section, the most important concepts about fuzzy sets are defined. This series of concepts regarding fuzzy sets allow us to deal with fuzzy sets, measure and compare them, and so on.

Definition 2: *Let A and B be two fuzzy sets over* X. *Then, A is* ***equal*** *to B if:*

$$A = B \Leftrightarrow \forall x \in \mathrm{X},\ \mu_A(x) = \mu_B(x) \tag{15}$$

Definition 3*: Taking two fuzzy sets A and B over* X, A *is said to be* ***included*** *in B if:*

$$A \subseteq B \Leftrightarrow \forall x \in \mathrm{X},\ \mu_A(x) \le \mu_B(x) \tag{16}$$

A fuzzy inclusion may be defined using a degree of subsethood. For example, when both fuzzy sets are defined in a finite universe, this degree may be computed as (Kosko, 1992):

$$S(A,B) = \frac{1}{\mathrm{Card}(A)}\left\{\mathrm{Card}(A) - \sum_{x \in X} \max\{0, A(x) - B(x)\}\right\} \tag{17}$$

Definition 4: *The* ***support*** *of a fuzzy set A defined over* X *is a subset of that universe that complies with:*

$$\mathrm{Supp}(A) = \{x : x \in \mathrm{X},\ \mu_A(x) > 0\} \tag{18}$$

Definition 5*: The* α***-cut*** *of a fuzzy set* A, *denoted by* A_α *is a classic subset of elements in* X, *whose membership function takes a greater or equal value to any specific* α *value of that universe of discourse that complies with:*

$$A_\alpha = \{x : x \in \mathrm{X},\ \mu_A(x) \ge \alpha, \alpha \in [0,1]\} \tag{19}$$

The **Representation Theorem** allows the representation of any fuzzy set A by means of the union of its α-cuts.

Definition 6: *The* ***Representation Theorem*** *states that any fuzzy set A can be obtained from the union of its* α*-cuts.*

$$A = \bigcup_{\alpha \in [0,1]} \alpha A_\alpha \tag{20}$$

Definition 7: *By using the Representation Theorem, the concept of* ***convex fuzzy set*** *can be established as that in which all the* α*-cuts are convex:*

$$\forall x, y \in \mathrm{X}, \forall \lambda \in [0,1]:$$
$$\mu_A(\lambda \cdot x + (1-\lambda) \cdot y) \ge \min(\mu_A(x), \mu_A(y)) \tag{21}$$

This definition means that any point situated between another two will have a higher membership degree than the minimum of these two points. Figures 7, 8, or 9 are typical examples of convex fuzzy sets, whereas Figure 10 represents a non-convex fuzzy set.

Definition 8: *A **concave fuzzy set** complies with:*

$$\forall x, y \in \mathrm{X}, \forall \lambda \in [0,1]: \quad \mu_A(\lambda \cdot x + (1-\lambda) \cdot y) \leq \min(\mu_A(x), \mu_A(y)) \tag{22}$$

Definition 9: *The **kernel** of a fuzzy set A, defined over* X, *is a subset of that universe that complies with:*

$$\mathrm{Kern}(A) = \{x : x \in \mathrm{X},\ \mu_A(x) = 1\} \tag{23}$$

Definition 10: *The **height** of a fuzzy set A defined over* X *is:*

$$\mathrm{Hgt}(A) = \sup_{x \in \mathrm{X}} \mu_A(x) \tag{24}$$

Definition 11: *A fuzzy set A is **normalized** if and only if:*

$$\exists x \in \mathrm{X}, \mu_A(x) = \mathrm{Hgt}(A) = 1 \tag{25}$$

Definition 12: *The **cardinality** of a fuzzy set A with finite universe* X *is defined as:*

$$\mathrm{Card}(A) = \sum_{x \in \mathrm{X}} \mu_A(x) \tag{26}$$

If the universe is infinite, the addition must be changed for an integral defined within the universe.

Membership Function Determination

If the system uses badly defined membership functions the system will not work well, and these functions must therefore be carefully defined. The membership functions can be calculated in several ways. The chosen method will depend on the concrete application, the manner in which the uncertainty is to be represented, and how this one is to be measured during the experiments. The following points give a brief summary of some of these methods (Pedrycz & Gomide, 1998).

1. **Horizontal method:** It is based on the answers of a group of **N** "experts."
   - The question takes the following form. Can ***x*** be considered compatible with the concept ***A***?"
   - Only "Yes" and "No" answers are acceptable, so:

   $A(x)$ = (Affirmative Answers) / N (27)

2. **Vertical method:** The aim is to build several α-cuts (Definition 5), for which several values are selected for α.
   - Now, the question that is formulated for these predetermined α values is as follows: Can the elements of **X** that belong to ***A*** with a degree that is not inferior to α be identified?"
   - From these α-cuts, the fuzzy set ***A*** can be identified, using the so-called identity principle or representation theorem (Definition 6).

3. **Pair comparison method** (Saaty, 1980): Supposing that we already have the fuzzy set *A* over the universe of discourse X of *n* values ($x_1, x_2, ..., x_n$), we could calculate the reciprocal matrix $M=[a_{hi}]$, a square matrix $n \times n$ with the following format:

$$M = \begin{bmatrix} 1 & \frac{A(x_1)}{A(x_2)} & \cdots & \frac{A(x_1)}{A(x_n)} \\ \frac{A(x_2)}{A(x_1)} & 1 & \cdots & \frac{A(x_2)}{A(x_n)} \\ \cdots & \cdots & \frac{A(x_i)}{A(x_j)} & \cdots \\ \frac{A(x_n)}{A(x_1)} & \frac{A(x_n)}{A(x_2)} & \cdots & 1 \end{bmatrix} \quad (28)$$

- This matrix has the following properties: The principal diagonal is always one, $a_{hi}\, a_{ih}=1$ (property of reciprocity), and $a_{hi}\, a_{ik} = a_{hk}$ (transitive property), $\forall\ i, j, k = 1, 2, \ldots, n$.
- If we want to calculate the fuzzy set A, the process is reversed: The matrix M is calculated, and then A is calculated from M.
- In order to calculate M, the level of priority or the highest membership degree of a pair of values is numerically quantified: x_i with respect to x_j.
  - The number of comparisons is: $n\ (n - 1) / 2$.
  - Transitivity is difficult to achieve (the eigenvalue of the matrix is used to measure the consistency of the data, so that if it is very low, the experiments should be repeated).

4. **Method based on problem specification:** This method requires a numerical function that should be approximate. The error is defined as a fuzzy set that measures the quality of the approximation.
5. **Method based on the optimization of parameters:** The shape of a fuzzy set A depends on some parameters, indicated by the vector p, which is represented by $A(x;p)$.

- Some experimental results in the form of pairs (element, membership degree) are needed: (E_k, G_k) with $k = 1, 2...., N$.
- The problem consists of optimizing the vector p, for example, minimizing the squared error:

$$\min_p \sum_{k=1}^{N} \left[G_k - A(E_k; p)\right]^2 \quad (29)$$

6. **Method based on fuzzy clustering:** This is based on clustering together the objects of the universe in overlapping groups where levels of membership to each group are considered as fuzzy degrees. There are several fuzzy clustering algorithms, but the most widely used is the algorithm of "fuzzy isodata" (Bezdek, 1981). In this handbook, there is a chapter by Feil and Abonyi explaining some data mining techniques, including fuzzy clustering.

Fuzzy Set Operations

Fuzzy sets theory generalizes the classic sets theory. It means that fuzzy sets allow operations of union, intersection, and complement. These and other operations can be found in Pedrycz and Gomide (1998) and Petry (1996) such as concentration (the square of the membership function), dilatation (the square root of the membership function), contrast intensification (concentration in values below 0.5 and dilatation in the rest of values), and fuzzification (the inverse operation). These operations can be used when linguistic hedges, such as "very" or "not very," are used.

Union and Intersection: T-conorms and T-norms

Definition 13: *If A and B are two fuzzy sets over a universe of discourse* X*, the membership function of the* ***union*** *of the two sets A∪B is expressed by:*

$$\mu_{A \cup B}(x) = f(\mu_A(x), \mu_B(x)), x \in \mathrm{X} \quad (30)$$

where f is a t-conorm (Schweizer & Sklar, 1983).

Definition 14: *If A and B are two fuzzy sets over a universe of discourse* X*, the membership function of the* **intersection** *of the two sets A∩B, is expressed by:*

$$\mu_{A\cap B}(x) = g(\mu_A(x), \mu_B(x)), x \in X \qquad (31)$$

where g is a t-norm (Schweizer & Sklar, 1983).

Both t-conorms (s-norms) and t-norms establish **generic models** respectively for the operations of **union** and **intersection,** which must comply with certain basic properties (commutative, associative, monotonicity, and border conditions). They are concepts derived from Menger (1942) and Schweizer and Sklar (1983), and that have been studied in-depth more recently (Butnario & Klement, 1993).

Definition 15: *Triangular Norm,* ***t-norm:*** *binary operation, t:* $[0,1]^2 \to [0,1]$ *that complies with the following properties:*

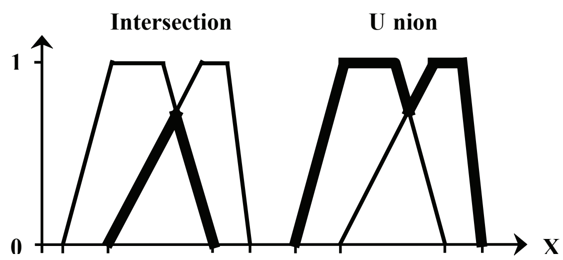
1. Commutativity: $x\ t\ y = y\ t\ x$.
2. Associativity: $x\ t\ (y\ t\ z) = (x\ t\ y)\ t\ z$.
3. Monotonicity: If $x \le y$, and $w \le z$ then $x\ t\ w \le y\ t\ z$.
4. Boundary conditions**:** $x\ t\ 0 = 0$, and $x\ t\ 1 = x$.

Definition 16: *Triangular Conorm,* ***t-conorm,*** *or* ***s-norm****: Binary operation, s:* $[0,1]^2 \to [0,1]$ *that complies with the following properties:*

1. Commutativity: $x\ s\ y = y\ s\ x$.
2. Associativity: $x\ s\ (y\ s\ z) = (x\ s\ y)\ s\ z$.
3. Monotonicity: If $x \le y$, and $w \le z$ then $x\ s\ w \le y\ s\ z$.
4. Boundary conditions: $x\ s\ 0 = x$, and $x\ s\ 1 = 1$.

The most widely used of this type of function is the **t-norm of the minimum** and the **t-conorm** or **s-norm of the maximum** as they have retained a large number of the properties of the boolean operators, such as the property of *idempotency* ($x\ t\ x = x$; $x\ s\ x = x$). In Figure 11, we can see the intersection and union, using respectively the minimum and maximum, of two trapezoid fuzzy sets.

Figure 11. Intersection (minimum) and union (maximum)



There is an extensive set of operators, called **t-norms (triangular norms)** and **t-conorms (triangular conorms)**, that can be used as connectors for modeling the intersection and union respectively (Dubois & Prade, 1980; Piegat, 2001; Predycz & Gomide, 1998; Yager, 1980). The most important are shown in Tables 1 and 2.

A relationship exists between t-norms (*t*) and t-conorms (*s*). It is an extension of De Morgan's Law:

$$\begin{aligned} x\ s\ y &= 1-(1-x)\ t\ (1-y) \\ x\ t\ y &= 1-(1-x)\ s\ (1-y) \end{aligned} \qquad (32)$$

When a t-norm or a t-conorm comply with this property they are said to be **conjugated** or **dual.**

T-norms and t-conorms cannot be ordered from larger to smaller. However, it is easy to identify the largest and the smallest t-norm and t-conorm: the largest and smallest t-norm are respectively the minimum and the drastic product, and the largest and smallest t-conorm are respectively the drastic sum and the maximum function. Note that if two fuzzy sets are convex, their intersection will also be (but not necessarily their union).

Negations or Complements

The notion of the complement can be constructed using the concept of *strong negation* (Trillas, 1979).

Definition 17: *A function* N: $[0,1] \rightarrow [0,1]$ *is a strong* ***negation*** *if it fulfills the following conditions:*

1. *Boundary conditions:* $N(0) = 1$ *and* $N(1) = 0$.
2. *Involution:* $N(N(x)) = x$.
3. *Monotonicity: N is nonincreasing.*
4. *Continuity: N is continuous.*

Although there are several types of operators which satisfy such properties or relaxed versions of them, Zadeh's version of the complement (Zadeh, 1965) is mainly used: $N(x) = 1 - x$. Thus, for a fuzzy set A in the universe of discourse X, the membership function of the complement, denoted by $\neg A$, or by $\overline{A}$, is shown as:

$$\mu_{\neg A}(x) = 1 - \mu_A(x), x \in X \qquad (33)$$

Implication Operators

A fuzzy implication (Dubois & Prade, 1984; Zadeh, 1975) is a function to compute the fulfillment degree of a rule expressed by IF X THEN Y, where the antecedent or premise and the consequent or conclusion are fuzzy.

Definition 18: *A function* f: $[0, 1] \times [0, 1] \rightarrow [0, 1]$ *is a* ***fuzzy implication***, $f(x,y) \in [0,1]$, *also denoted by* $x \Rightarrow_f y$, *if it fulfils the following conditions:*

1. $0 \Rightarrow_f a = 1, \forall a \in [0,1]$
2. $a \Rightarrow_f 1 = 1, \forall a \in [0,1]$
3. $1 \Rightarrow_f a = a, \forall a \in [0,1]$
4. *Decreasing (respectively increasing) monotonicity with respect to the first (respectively second) argument.*

Sometimes, another condition is added: $(x \Rightarrow_f (y \Rightarrow_f z)) = (y \Rightarrow_f (x \Rightarrow_f z))$. The most important implication functions are shown in Table 3. Note that Kleene-Dienes implication is based on the classical implication definition ($x \Rightarrow y = \neg x \vee y$); that is, it is a strong implication, using the Zadeh's negation and the maximum s-norm.

In standard fuzzy sets theory, there are, basically, four models for implication operations (Trillas & Alsina, 2002; Trillas, Alsina, & Pradera, 2004; Trillas, Cubillo, & del Campo, 2000; Ying, 2002): (1) Strong or S-implications ($x \Rightarrow_f y = N(x)\ s\ y$), (2) Residuated or R-implications ($x \Rightarrow_f y = \sup_c \{x \in [0,1], x\ s\ c \leq y\}$), (3) Quantum logic, Q-implications, or QM-implications ($x \Rightarrow_f y = N(x)\ s\ (x\ t\ y)$), and (4) Mamdani–Larsen or ML-implications ($x \Rightarrow_f y = \varphi_1(x)\ t\ \varphi(y)$, where φ_1 is an order automorphism on [0,1] and φ_2: $[0,1] \rightarrow [0,1]$ is a non-null contractive mapping, that is, $\varphi_2(w) \leq w, \forall w \in [0,1]$). Some of these types of fuzzy implications are overlapping (for example, Łukasiewicz implication is an S-implication and an R-implication).

Some applications utilize implication functions, which do not fulfill all the conditions in the previous definition, like the modified Łukasiewicz implication. Besides, it is very usual to use t-norms as implications functions (Gupta & Qi, 1991) obtaining very good results, especially the minimum (Mandani implication) and the product t-norms.

Comparison Operations on Fuzzy Sets

The fuzzy sets, defined using a membership function, can be compared in different ways. We will now list several methods used to compare fuzzy sets (Pedrycz & Gomide, 1998).

Distance Measures: A distance measure considers a distance function between the membership functions of two fuzzy sets in the same universe. In such a way, it tries to indicate the proximity between the two fuzzy sets. In general, the distance between A and B, defined in the same universe of discourse, can be defined using the Minkowski distance:

$$d(A,B) = \left[\int_X |A(x) - B(x)|^p \, dx \right]^{1/p} \qquad (34)$$

where $p \geq 1$ and we assume that the integral exists. Several specific cases are typically used:

1. Hamming Distance ($p = 1$):

$$d(A,B) = \int_X |A(x) - B(x)| \, dx \qquad (35)$$

Table 1. t-norms functions: f(x,y) = x t y

t-norms	Expression
Minimum	$f(x,y)=\min(x,y)$
Product (Algebraic)	$f(x,y)=xy$
Drastic Product	$f(x,y)=\begin{cases} xy, & \text{if } y=1 \\ y, & \text{if } x=1 \\ 0, & \text{otherwise} \end{cases}$
Bounded Product (bounded difference)	$f(x,y)=\max[0,(1+p)(x+y-1)-pxy], \quad \text{where } p\geq -1$
Hamacher Product	$f(x,y)=\frac{xy}{p+(1-p)(x+y-xy)}, \quad \text{where } p\geq 0$
Yager Family	$f(x,y)=1-\min(1,[(1-x)^p+(1-y)^p]^{/p}), \quad \text{where } p>0$
Dubois-Prade Family	$f(x,y)=\frac{xy}{\max(x,y,p)}, \quad \text{where } 0\leq p\leq 1$
Frank Family	$f(x,y)=\log_p\left(1+\frac{(p^x-1)(p^y-1)}{p-1}\right), \quad \text{where } p>0;\ p\neq 1$
Einstein Product	$f(x,y)=\frac{xy}{1+(1-x)+(1-y)}$
Others	$f(x,y)=\frac{1}{1+[((1-x)/x)^p+((1-y)/y)^p]^{/p}}, \quad \text{where } p>0$ $f(x,y)=\frac{1}{1/x^p+1/y^p-1}$ $f(x,y)=[\max(0,x^p+y^p-1)]^{/p}$

2. Euclidean Distance ($p = 2$):

$$d(A,B)=\left[\int_X |A(x)-B(x)|^2\, dx\right]^{1/2} \quad (36)$$

For discrete universe of discourses, integration is replaced with sum. The more similar are the fuzzy sets, smaller is the distance between them. Therefore, it is convenient to normalize the function of distance, denoted by $d_n(A,B)$, and use this form to express the similarity as a direct complementation: $1 - d_n(A,B)$.

Equality Indexes: This is based on the logical expression of equality, i.e., two sets A and B are equals if $A\subset B$ and $B\subset A$. In fuzzy sets, a certain degree of equality can be found. With that, the following expression is defined:

$$(A\equiv B)(x)=$$

$$\frac{[A(x)\varphi B(x)]\wedge[B(x)\varphi A(x)]+[\bar{A}(x)\varphi\bar{B}(x)]\wedge[\bar{B}(x)\varphi\bar{A}(x)]}{2} \quad (37)$$

Table 2. s-norms functions: f(x,y) = x s y

t-conorms or s-norms	Expression
Maximum	$f(x,y)=\max(x,y)$
Sum-Product (Algebraic sum)	$f(x,y)=x+y-xy$
Drastic sum	$f(x,y)=\begin{cases} x, & \text{if } y=0 \\ y, & \text{if } x=0 \\ 1, & \text{otherwise} \end{cases}$
Bounded sum	$f(x,y)=\min(1,x+y+pxy)$, where $p\geq 0$
Einstein sum	$f(x,y)=\dfrac{x+y}{1+xy}$
Sugeno Family	$f(x,y)=\min(1,x+y+p-xy)$, where $p\geq 0$
Yager Family	$f(x,y)=\min(1,[x^p+y^p]^{1/p})$, where $p>0$
Dubois-Prade Family	$f(x,y)=\dfrac{(1-x)(1-y)}{\max(1-x,\ 1-y,\ p)}$, where $p\in[0,1]$
Frank Family	$f(x,y)=\log_p\left(1+\dfrac{(p^{1-x}-1)(p^{1-y}-1)}{p-1}\right)$ where $p>0;\ p\neq 1$
Others	$f(x,y)=\dfrac{x+y-xy-(1-p)xy}{1-(1-p)xy}$, where $p\geq 0$ $f(x,y)=1-\max(0,[(1-x)^p+(1-y)^p-1]^{1/p})$, where $p>0$ $f(x,y)=\dfrac{1}{1-[x/(1-x)^p+y/(1-y)^p]^{1/p}}$, where $p>0$ $f(x,y)=\dfrac{1}{1-[1/(1-x)^p+1/(1-y)^p-1]^{1/p}}$, where $p>0$

where the conjunction ($\wedge$) is modeled on the minimum operation, and the inclusion is represented by the operator φ (phi), induced by a continuous t-norm *t*:

$$A(x)\,\varphi\,B(x)=\sup_{c\in[0,1]}\left[A(x)\,t\,c\leq B(x)\right] \qquad (38)$$

Taking the t-norm of bounded product with *p*=0 (Table 1) as an example:

$$A(x)\varphi B(x)=\begin{cases} 1 & \text{if } A(x)<B(x) \\ B(x)-A(x)+1 & \text{if } A(x)\geq B(x) \end{cases}\Rightarrow$$

$$(A\equiv B)\,(x)=\begin{cases} A(x)-B(x)+1 & \text{if } A(x)<B(x) \\ B(x)-A(x)+1 & \text{if } A(x)\geq B(x) \end{cases} \qquad (39)$$

Three basic methods can be used to obtain a single value ($\forall x\in \mathrm{X}$):

Table 3. Implication functions: $f(x,y) = x \Rightarrow_f y$

Implication	Expression
Kleene-Dienes	$f(x, y) = \max(1 - x, y)$
Reichenbach, M izumoto or Kleene-Dienes-Łukasiewicz	$f(x, y) = 1 - x + xy$
Klir-Yuan	$f(x, y) = 1 - x + x^2 y$
Gödel	$f(x,y) = \begin{cases} 1, & \text{if } x \le y \\ y, & \text{otherwise} \end{cases}$
Rescher-Gaines	$f(x,y) = \begin{cases} 1, & \text{if } x \le y \\ 0, & \text{otherwise} \end{cases}$
Goguen	$f(x,y) = \begin{cases} 1, & \text{if } x = 0 \\ \min(1, y/x), & \text{otherwise } (x \ne 0) \end{cases}$
Łukasiewicz	$f(x,y) = \begin{cases} 1, & \text{if } x \le y \\ 1 - x + y, & \text{otherwise} \end{cases}$
Modified Łukasiewicz	$f(x, y) = 1 - \lvert x - y \rvert$
Yager	$f(x, y) = x^y$
Zadeh	$f(x, y) = \max(1 - x, \min(x, y))$

- Optimistic Equality Index:

$$(A \equiv B)_{opt} = \sup_{x \in X} (A \equiv B)(x) \qquad (40)$$

- Pessimistic Equality Index:

$$(A \equiv B)_{pes} = \inf_{x \in X} (A \equiv B)(x) \qquad (41)$$

- Medium Equality Index:

$$(A \equiv B)_{avg} = \left(1 / Card(x) \right) \int_x (A \equiv B)(x) dx \qquad (42)$$

Thus, the following relationship is satisfied:

$$(A \equiv B)_{pes} \le (A \equiv B)_{avg} \le (A \equiv B)_{opt} \qquad (43)$$

Possibility and Necessity Measures: These concepts use the fuzzy sets as possibility distributions where $A(x)$ measures the possibility of being A for each value in X (Zadeh, 1978). Thus, the comparison, that is, the possibility of value A being equal to value B, measures the extent to which A and B superpose each other. It is denoted by Poss(A,B) and defined as:

$$\text{Poss } (A, B) = \sup_{x \in X} \left[\min \left(A(x), B(x) \right) \right] \qquad (44)$$

The necessity measure describes the degree to which B is included in A, and it is denoted by Nec(A,B):

$$\text{Nec}(A, B) = \inf_{x \in X} \left[\max \left(A(x), 1 - B(x) \right) \right] \qquad (45)$$

In Figures 12 and 13, we can see graphically how these measurements, for two concrete fuzzy sets, are calculated. It can be stated that: $\text{Poss } (A, B) = \text{Poss } (B, A)$. On the other hand, the measurement of necessity is asymmetrical, $\text{Nec } (A, B) \ne \text{Nec } (B, A)$. However, the following relation is fulfilled:

Figure 12. General illustration of the Poss(A,B) concept using the minimum t-norm

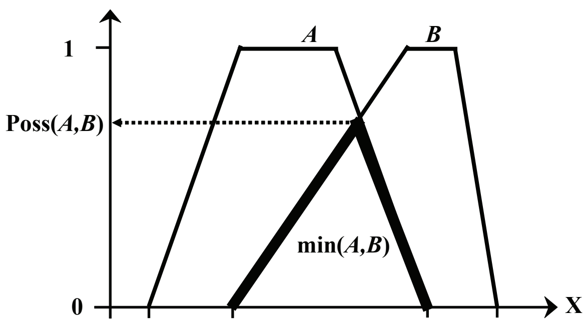
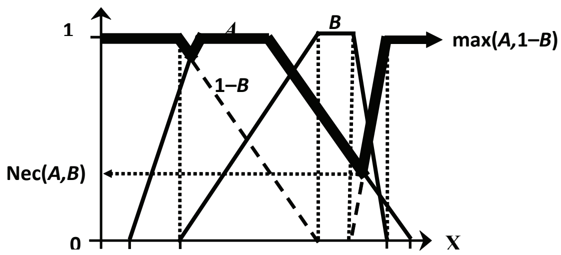


Figure 13. General illustration of the Nec(A,B) concept using the maximum t-conorm



$$\text{Nec}\ (A, B)\ +\ \text{Poss}\ (\overline{A}, B)\ = 1 \qquad (46)$$

Other equivalences are:

$$\text{Poss}\ (A \cup B, C)\ =\ \max\left\{\text{Poss}\,(A, C), \text{Poss}\,(B, C)\right\} \qquad (47)$$

$$\text{Poss}\ (A \cap B, C)\ =\ \min\left\{\text{Nec}\,(A, C), \text{Nec}\,(B, C)\right\} \qquad (48)$$

The generalization of the possibility and necessity measurements use triangular t-norms or t-conorms instead of min and max functions, respectively. If the concept is extended, the possibility of a fuzzy set *A* (or a possibility distribution) in the universe X can be defined as:

$$\Pi\,(A) = \text{Poss}\,(A, \text{X}) = \sup_{x \in \text{X}} \left[\min\left(A(x), 1\right)\right] = \sup_{x \in \text{X}}\ A(x) \qquad (49)$$

This possibility measures whether or not a determined event (the fuzzy set *A*) is possible in universe X. It would not measure uncertainty, because if $\Pi\,(A) = 1$, we know that event *A* is possible, but:

- if $\Pi\,(\overline{A}) = 1$, then the certainty is indeterminate.
- if $\Pi\,(\overline{A}) = 0$, then the occurrence of *A* is certain.

Therefore, the following two equalities are always satisfied:

- $\Pi\,(\text{X}) = 1$ (possibility of an element of the universe).
- $\Pi\,(\phi) = 0$ (possibility of an element *not* in the universe).

Similarly, the necessity of a fuzzy set *N*(*A*) in X can be defined, and then we can set some equivalences of possibility and necessity (see Equation 50 and 51).

These equivalences explain why the necessity complements the information about the certainty of event *A*:

- The greater *N*(*A*), the smaller possibility of opposite event (¬*A*).
- The greater Π (*A*), the smaller necessity of the opposite event (¬*A*).
- *N*(*A*) = 1 ⇔ ¬*A* is totally impossible (if an event is totally necessary, then the opposite event is totally impossible).
- Π(A) = 1 ⇔ ¬*A* is not necessary at all *N*(¬*A*) = 0 (if an event is totally possible, then the opposite event cannot be necessary in any way).
- *N*(*A*) = 1 ⇒ Π(*A*) = 1 (if *A* is a totally necessary event, then must be totally possible). Note that the opposite is not satisfied.

Equation 50.

$$N(A) = \inf_{x\in X} \{A(x)\} = 1-\sup_{x\in X} \{1 - A(x)\} = 1 - \Pi(\neg A)\text{: } N(A) = 1 - \Pi (\neg A)$$

Equation 51.

$$\Pi (A) = \sup_{x\in X} \{A(x)\} = 1 - \inf_{x\in X} \{1 - A(x)\} = 1 - N(\neg A)\text{: } \Pi (A) = 1 - N(\neg A)$$

- $A \subseteq B \Rightarrow N(A) \leq N(B)$ and $\Pi(A) \leq \Pi (B)$.

Compatibility Measures: This comparison operation measures the extent to which a certain fuzzy set is compatible with another (defined in the same space). The result is not a single number but a fuzzy set defined in the unit interval, [0,1], known as fuzzy set of compatibility. The compatibility of B with A can be defined as:

$$Comp(B, A)(u) = \sup_{u=A(x)} \{B(x)\},\ u \in [0,1] \tag{52}$$

Set B can be seen as a "fuzzy value" and set A as a "fuzzy concept." Therefore, *Comp*(B,A) measures the compatibility with which B is A.

Example 2: Let B be the value "approx. 70 years" and A be the concept "very old." Then, the fuzzy set *Comp*(B,A) is represented in Figure 14 and the fuzzy set *Comp*(A,B) in Figure 15.

The compatibility measurement has the following properties:

- It measures the degree to which B can fulfill concept A. That degree will be greater, the more similar the fuzzy set *Comp* (B, A) is to the singleton "1" value (maximum compatibility).
- Supposing A is a normalized fuzzy set: $Comp(A, A)(u) = u$ (Linear membership function).
- If A is not normalized, the function will be the same between 0 and the height of set A: If u > Height(A), $Comp(A, A)(u)$= indeterminate (0).
- If B is a number x ("singleton" fuzzy set), the result will also be another "singleton" in the $A(x)$ value:

$$Comp\ (B, A)(u) = \begin{cases} 1, & \text{if } u = A(x) \\ 0, & \text{otherwise} \end{cases} \tag{53}$$

- If B is not normalized, the result will not be either; its height is the same as that of set B.
- If Support (A) ∩ Support(B) = ∅, then:

Figure 14. Example 2: Illustration of Comp(B,A)

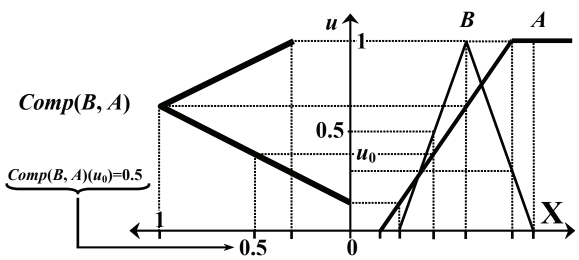
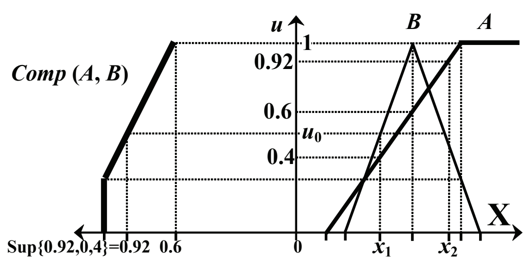


Figure 15. Example 2: Illustration of Comp(A,B)



$$Comp\,(B,A)(u) = Comp\,(A,B)(u) = \begin{cases} 1, & \text{if } u = 0 \text{ (minimum compatibility)} \\ 0, & \text{otherwise} \end{cases} \tag{54}$$

- The possibility and necessity measurements between *A* and *B* are included in the support of *Comp*(*B*, *A*).

In order to have a clearer vision of what this measurement means, we can look at the examples shown in Figure 16. We can conclude that fuzzy set *B* is more compatible with another *A*, the closer *Comp*(*B*, *A*) is to 1 and the further it is from 0 (the less area it has).

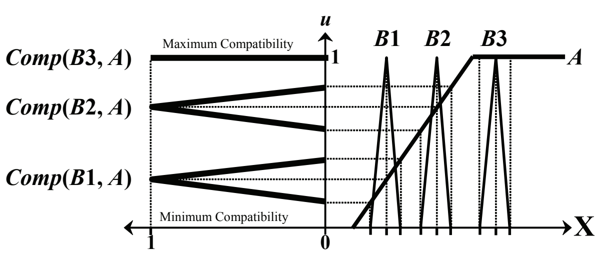
Fuzzy Relations

A classic relation between two universes X and Y is a subset of the Cartesian product X×Y. Like the classic sets, the classic relation can be described using a characteristic function. In the same way, a fuzzy relation *R* is a fuzzy set of tuples. In the event of a binary relation, the tuple has two values.

Definition 19: *Let* U *and* V *be two infinite (continuous) universes and* $\mu_R : \mathrm{U} \times \mathrm{V} \to [0,1]$*. Then, a* ***binary fuzzy relation*** *R is defined as:*

$$R = \int_{\mathrm{U}\times\mathrm{V}} \mu_R(u,v)/(u,v) \tag{55}$$

Figure 16. Three sets (B1, B2, and B3) with the same shape placed in different positions and compared to A



The function μ_R may be used as a similarity or proximity function. It is important to stress that not all functions are relations and not all relations are functions. Fuzzy relations generalize the concept of relation by allowing the notion of partial belonging (association) between points in the universe of discourse.

Example 3: Take as an example the fuzzy relation in $\Re^2$ (binary relation), "approximately equal," with the following membership function in $X \subset \Re$, with $X^2 = \{1,2,3\}^2$: 1/(1,1) + 1/(2,2) + 1(3,3) + 0.8/(1,2) + 0.8/(2,3) + 0.8/(2,1) + 0.8/(3,2) + 0.3/(1,3) + 0.3/(3,1). This fuzzy relation may be defined as:

$$\text{x approximately equal to y}: \quad R(\mathrm{x},\mathrm{y}) = \begin{Bmatrix} 1, & \text{if } \mathrm{x} = \mathrm{y} \\ 0.8, & \text{if } |\mathrm{x} - \mathrm{y}| = 1 \\ 0.3, & \text{if } |\mathrm{x} - \mathrm{y}| = 2 \end{Bmatrix}$$

where $\mathrm{x}, \mathrm{y} \in \Re$. When the universe of discourse is finite, a matrix notation can be quite useful to represent the relation. This example would be shown as:

X^2	1	2	3
1	1	0.8	0.3
2	0.8	1	0.8
3	0.3	0.8	1

Definitions of basic operations with fuzzy relations are closely linked to operations of fuzzy sets. Let *R* and *W* be two fuzzy relations defined in X × Y:

- Union: $(R \cup W)(x,y) = R(x,y)\ s\ W(x,y)$, using a s-norm *s*.
- Intersection: $(R \cup W)(x,y) = R(x,y)\ t\ W(x,y)$, using a t-norm *t*.
- Complement: $(\neg R)(x,y) = 1 - R(x,y)$.
- Inclusion: $R \subseteq W \Leftrightarrow R(x,y) \leq W(x,y)$.
- Equality: $R = W \Leftrightarrow R(x,y) = W(x,y)$.

Fuzzy Numbers

The concept of fuzzy number was first introduced in Zadeh (1975) with the purpose of analyzing and manipulating approximate numeric values, for example, "near 0," "almost 5," and so forth. The concept has been refined (Dubois & Prade, 1980, 1985), and several definitions exist.

Definition 20: *Let A be a fuzzy set in* X *and* $\mu_A(x)$ *be its membership function with* $x \in X$. *A is a **fuzzy number** if its membership function satisfies that:*

1. $\mu_A(x)$ is convex.
2. $\mu_A(x)$ is upper semicontinuity.
3. Support of A is bounded.

These requirements can be relaxed. The general form of the membership function of a fuzzy number *A* with support (*a,d*) and kernel or modal interval (*b,c*) can be defined as:

$$\mu_A(x) = \begin{cases} r_A(x) & \text{if } x \in (a,b) \\ h & \text{if } x \in [b,c] \\ s_A(x) & \text{if } x \in (c,d) \\ 0 & \text{otherwise} \end{cases} \quad (56)$$

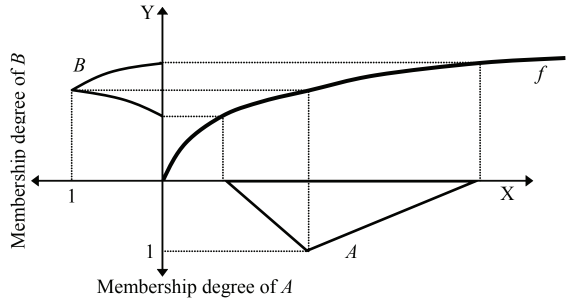
where r_A, s_A:X → [0,1], r_A is not decreasing, s_A is not increasing

$$r_A(a) = s_A(d) = 0 \quad \text{and} \quad r_A(b) = s_A(c) = h \quad (57)$$

with $h \in (0,1]$ and $a, b, c, d \in X$. The number *h* is called the height of the fuzzy number, and some authors include the necessity of normalized fuzzy numbers, that is, with $h = 1$. The numbers $b - a$ and $d - c$ are the left and right spaces, respectively.

Throughout this study, we will often use a particular case of fuzzy numbers that is obtained when we consider the functions r_A and s_A as linear functions. We will call this type of fuzzy number triangular or trapezoidal, and it takes the form shown in Figure 7. Many applications usually work with normalized trapezoidal fuzzy numbers (*h*=1) because these fuzzy numbers are easily characterized using the four really necessary numbers: $A \equiv (a, b, c, d)$.

Figure 17. Graphic representation of the extension principle, where f carries out its transformation from X to Y



The Extension Principle

One of the most important notions in the fuzzy sets theory is the extension principle, proposed by Zadeh (1975). It provides a general method that allows nonfuzzy mathematical concepts to be extended to the treatment of fuzzy quantities. It is used to transform fuzzy quantities, which have the same or different universes, according to a transformation function between those universes.

Let *A* be a fuzzy set, defined in universe of discourse X and *f* a nonfuzzy transformation function between universes X and Y, so that *f*: X → Y. The purpose is to extend *f* so that it can also operate on the fuzzy sets in X. The result must be a fuzzy set *B* in Y: $B = f(A)$. This transformation is represented in Figure 17. It is achieved with the use of the Sup-Min composition, which will now be described in a general way in the case of the Cartesian product in *n* universes.

Definition 21: *Let* X *be a Cartesian product of n universes such as* $X = X_1 \times X_2 \times ... \times X_n$, *and* $A_1, A_2, ..., A_n$ *are n fuzzy sets in those n universes respectively. Moreover, we have a function f from* X *to the universe* Y, *so a fuzzy set B from* Y *is* *defined by the* ***extension principle*** *as B = f(A$_1$, A$_2$,..., A$_n$) defined as:*

$$\mu_B(y) = \sup_{x \in X, y=f(x)} \min(\mu_{A_1}(x_1),..., \mu_{A_n}(x_n)) \quad (58)$$

Example 4: Let both X and Y be the universe of natural numbers.

- **Sum 4** function**:** ***y = f(x) = x + 4;***
- ✓ *A = 0.1/2 + 0.4/3 + 1/4 + 0.6/5;*
- ✓ *B = f(A) = 0.1/6 + 0.4/7 + 1/8 + 0.6/9;*

- **Sum: y =** ***f*****(x$_1$,x$_2$) = x$_1$ + x$_2$:**
- ✓ *A_1 =0.1/2 + 0.4/3 + 1/4 + 0.6/5;*
- ✓ *A_2 =0.4/5 + 1/6;*
- ✓ *B = f(A_1, A_2) = 0.1/7 + 0.4/8 + 0.4/9 + 1/10 + 0.6/11;*

We can conclude that the extension principle allows us to extend any function (for example arithmetic) to the field of fuzzy sets, making possible the fuzzy arithmetic.

Fuzzy Arithmetic

Thanks to the extension principle (Definition 21), it is possible to extend the classic arithmetical operations to the treatment of fuzzy numbers (see Example 4). In this way, the four main operations are extended in:

1. **Extended sum:** Given two fuzzy quantities A_1 and A_2 in X, the membership function of the sum $A_1 + A_2$ is found using the expression:

$$\mu_{A1+A2}(y) = \sup\{\min(\mu_{A1}(y-x), \mu_{A2}(x)) / x \in X\} \quad (59)$$

   In this way, the sum is expressed in terms of the supreme operation. The extended sum is a commutative and associative operation and the concept of the symmetrical number does not exist.

2. **Extended difference:** Given two fuzzy quantities A_1 and A_2, in X, the membership function of the difference $A_1 - A_2$ is found using the expression:

$$\mu_{A1-A2}(y) = \sup\{\min(\mu_{A1}(y+x), \mu_{A2}(x)) / x \in X\} \quad (60)$$

3. **Extended product:** The product of two fuzzy quantities $A_1 * A_2$ is obtained as follows:

$$\mu_{A_1 * A_2}(y) = \begin{cases} \sup\{\min(\mu_{A_1}(y/x), \mu_{A_2}(x)), \quad x \in X - \{0\}\} & \text{if } x \neq 0 \\ \max(\mu_{A_1}(0), \mu_{A_2}(0)) & \text{if } x = 0 \end{cases} \quad (61)$$

4. **Extended division:** The division of two fuzzy quantities $A_1 \div A_2$ is defined as follows:

$$\mu_{A1 \div A2}(y) = \sup\{\min(\mu_{A1}(xy), \mu_{A2}(x)), \quad x \in X\} \quad (62)$$

From these definitions, we can easily conclude that if A_1 and A_2 have a discrete universe (with finite terms) and they have *n* and *m* terms respectively, then the number of terms of A_1+A_2 and of A_1–A_2 is (*n*–1)+(*m*–1)+1, that is, *n*+*m*–1. Based on a particular expression from the uncertainty principle, adapted to the use of α-cuts and in a type of numbers similar to those previously described, called LR fuzzy numbers (Dubois & Prade, 1980), rapid calculus formulae for the previous arithmetical operations are described.

It is important to point out that if we have two fuzzy numbers, the sum or remainder of both fuzzy numbers will be fuzzier (it will have greater cardinality) than the most fuzzy of the two (that which has greatest cardinality). This is logical, since if we add two "approximate" values, the exact value of which we do not know, the result can be as varied as the initial values are. The same thing happens with division and multiplication but on a larger scale.

Possibility Theory

This theory is based on the idea of linguistic variables and how these are related to fuzzy sets (Dubois & Prade, 1988; Zadeh, 1978). In this way, we can evaluate the possibility of a determinate variable X being (or belonging to) a determinate set *A*, like the membership degree of the X elements in *A*.

Definition 22: *Let there be a fuzzy set A defined in* X *with membership function* $\mu_A(x)$ *and a variable x in* X *(whose value we do not know). So, the proposition "x is A" defines a* ***possibility distribution****, in such a way that it is said that the possibility of* $x = u$ *is* $\mu_A(u)$, $\forall u \in X$.

The concepts of fuzzy sets and membership functions are now interpreted as linguistic labels and possibility distributions. Instead of membership degrees, we have possibility degrees, but all the tools and properties defined for fuzzy sets are also applicable to possibility distributions.

Fuzzy Quantifiers

Fuzzy or linguistic quantifiers (Liu & Kerre, 1998a, 1998b; Yager, 1983; Zadeh, 1983) have been widely applied to many applications, including database applications (Galindo, 1999; Galindo, Medina, Cubero, & García, 2001). Fuzzy quantifiers allow us to express fuzzy quantities or proportions in order to provide an approximate idea of the number of elements of a subset fulfilling a certain condition or the proportion of this number in relation to the total number of possible elements. Fuzzy quantifiers can be *absolute* or *relative*:

- **Absolute quantifiers** express quantities over the total number of elements of a particular set, stating whether this number is, for example, "much more than 10," "close to 100," "a great number of,"and so forth. Generalizing this concept, we can consider fuzzy numbers as absolute fuzzy quantifiers, in order to use expressions like "approximately between 5 and 10," "approximately −8,"and so on. Note that the expressed value may be positive or negative. In this case, we can see that the truth of the quantifier depends on a single quantity. For this reason, the definition of absolute fuzzy quantifiers is, as we shall see, very similar to that of fuzzy numbers.
- **Relative quantifiers** express measurements over the total number of elements, which fulfill a certain condition depending on the total number of possible elements (the proportion of elements). Consequently, the truth of the quantifier depends on two quantities. This type of quantifier is used in expressions such as "the majority" or "most," "the minority," "little of," "about half of,"and so forth. In this case, in order to evaluate the truth of the quantifier, we need to find the total number of elements fulfilling the condition and consider this value with respect to the total number of elements which could fulfill it (including those which fulfill it and those which do not fulfill it).

Some quantifiers such as "many" and "few" can be used in either sense, depending on the context (Liu & Kerre, 1998a). In Zadeh (1983), absolute fuzzy quantifiers are defined as fuzzy sets in positive real numbers and relative quantifiers as fuzzy sets in the interval [0,1]. We have extended the definition of absolute fuzzy quantifiers to all real numbers.

Definition 23: *A* ***fuzzy quantifier*** *named* Q *is represented as a function Q, the domain of which depends on whether it is absolute or relative:*

$$Q_{abs} : \Re \to [0,1] \tag{63}$$

$$Q_{rel} : [0,1] \to [0,1] \tag{64}$$

where the domain of Q_{rel} *is* [0,1] *because the division* $a/b \in [0,1]$*, where a is the number of elements fulfilling a certain condition, and b is the total number of existing elements.*

In order to know the fulfillment degree of the quantifier over the elements that fulfill a certain condition, we can apply the function Q of the quantifier to the value of quantification Φ, with Φ = a if Q is absolute and Φ = a/b if Q is relative.

There are two very important classic quantifiers: The universal quantifier (for all, ∀), and the existential quantifier (exist, ∃). The first of them is relative and the second one is absolute. They are discretely defined as:

$$Q_{\forall}(x) = \begin{cases} 1 & \text{if } x = 1 \\ 0 & \text{otherwise} \end{cases} \tag{65}$$

$$Q_{\exists}(x) = \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{otherwise} \end{cases} \tag{66}$$

Some quantifiers (absolute or relative) may have arguments, and in these cases, the function is defined using the arguments (Galindo, Urrutia, & Piattini, 2006). A survey of methods for evaluating quantified sentences and some new methods are shown in the literature (Delgado, Sánchez, & Vila, 1999, 2000), and in this volume, see the chapter by Liétard and Rocacher.

FUZZY DATABASES

In the Foreword of this handbook, Vila and Delgado refer to a series of five reports, which are a guideline for the development of research in this area. They state that "one of the lines that appears with more continuity and insistence is the treatment of the imprecise and uncertain information in databases." Imprecision has been studied in order to elaborate systems, databases, and consequently applications which support this kind of information. Most works which studied the imprecision in information have used possibility, similarity, and fuzzy techniques.

If a regular or classical database is a structured collection of records or data stored in a computer, a fuzzy database is a database which is able to deal with uncertain or incomplete information using fuzzy logic. Basically, a fuzzy database is a database with fuzzy attributes, which may be defined as attributes of an item, row, or object in a database, which allows storing fuzzy information (Bosc, 1999; De Caluwe & De Tré, 2007; Galindo et al., 2006; Petry, 1996).

There are many forms of adding flexibility in fuzzy databases. The simplest technique is to add a fuzzy membership degree to each record, that is, an attribute in the range [0,1]. However, there are other kinds of databases allowing fuzzy values to be stored in fuzzy attributes using fuzzy sets, possibility distributions, or fuzzy degrees associated to some attributes and with different meanings (membership degree, importance degree, fulfillment degree, etc.).

Of course, fuzzy databases should allow fuzzy queries using fuzzy or nonfuzzy data, and there are some languages based on SQL (ANSI, 1992; Date & Darwen, 1997) that allow these kind of queries like FSQL (Galindo, 2007; Galindo et al., 2006) or SQLf (Bosc & Pivert, 1995; Goncalves & Tineo, 2006). The research on fuzzy databases has been developed for about 20 years and concentrated mainly on the following areas:

1. Fuzzy querying in classical databases,
2. Fuzzy queries on fuzzy databases,
3. Extending classical data models in order to achieve fuzzy databases (fuzzy relational databases, fuzzy object-oriented databases, etc.),
4. Fuzzy conceptual modeling tools,
5. Fuzzy data mining techniques, and
6. Applications of these advances in real databases.

All of these different issues have been studied in different chapters of this volume, except the fourth item because, in general, there is little interest in fuzzy conceptual issues and, besides, these subjects have been studied in some other works in a very exhaustive manner (Chen, 1998; Galindo et al., 2006; Kerre & Chen, 2000; Ma, 2005; Yazici & George, 1999).

The first research area, fuzzy queries in classical databases, is very useful because currently there are many classical databases. The second item includes the first one, but we prefer to separate them because item 2 finds new problems that must be studied and because it must be framed in a concrete fuzzy database model (third item). The querying with imprecision, contrary to classical querying, allows the users to use fuzzy linguistic labels (also named linguistic terms) and express their preferences to better qualify the data they wish to get. An example of a flexible query, also named in this context fuzzy query, would be "list of the *young* employees, working in department with *big budget*." This query contains the fuzzy linguistic labels "*young*" and "*big budget*." These labels are words, in natural language, that express or identify a fuzzy set.

In fact, the flexibility of a query reflects the preferences of the end user. This is manifested by using a fuzzy set representation to express a flexible selection criterion. The extent to which an object in the database satisfies a request then becomes a matter of degree. The end user provides a set of attribute values (fuzzy labels) which are fully acceptable to the user, and a list of minimum thresholds for each of these attributes. With these elements, a fuzzy condition is built and the fuzzy querying system ranks the answered items according to their fulfillment degree. Some approaches, the so-called bipolar queries, need both the fuzzy condition (or fuzzy constraint) and the less-compulsory positive preferences or wishes. A very interesting work about bipolar queries may be found in this volume in the chapter by Dubois and Prade. In another chapter, Urrutia, Tineo and Gonzalez study the two most known fuzzy querying languages, FSQL and SQLf. Of course, we must reference the interesting and general review about the fuzzy querying proposals, written by Zadrożny, de Tré, de Caluwe, and Kacprzyk. Other chapters about fuzzy queries study different aspects, such as evaluation strategies or quantified statements, for example.

About fuzzy data mining issues, this handbook includes a complete review chapter by Feil and Abonyi, studying the main fuzzy data mining methods. Perhaps the more interesting and useful tools are the fuzzy clustering and the fuzzy dependencies, and both of them are also studied in different other chapters of this handbook. The last item, applications, is also studied in some chapters. These chapters mix different theoretical issues like data mining to real contexts with different goals.

About the third item, extending classical data models in order to achieve fuzzy databases, this handbook also includes interesting chapters. Ben Hassine et al. study in their chapter how to achieve fuzzy relational databases, giving different methods and addressing their explanations, mainly to database administrators and enterprises, in order to facilitate the migration to fuzzy databases. Takači and Škrbić present their fuzzy relational model and propose a fuzzy query language with the possibility to specify priorities for fuzzy statements. Barranco et al. give a good approach of a fuzzy object-relational database model, whereas some other interesting fuzzy object-oriented database models are presented and summarized respectively (De Caluwe, 1997; Galindo et al., 2006). This last book includes fuzzy time data types and in the book you have in your hands, Schneider defines fuzzy spatial datatypes. In another chapter, Belohlavek presents an overview of foundations of formal concept analysis of data with graded attributes, which provides elaborated mathematical foundations for relational data in some fuzzy databases.

In this section, we want to give a wide historical point of view summarizing the main published models aiming at solving the problem of representation and treatment of imprecise information in relational databases. This problem is not trivial because it requires relation structure modification, and actually, the operations on these relations also need to be modified. To allow the storage of imprecise information and the making of an inaccurate query of such information, a wide variety of case studies is required, which do not occur in the classic model, without imprecision.

The first approaches, which do not utilize the fuzzy logic, were proposed by Codd (1979, 1986, 1987, 1990). Then, some basic models were proposed, like the Buckles-Petry model (1982a, 1982b, 1984), the Prade-Testemale model (1984, 1987a, 1987b; Prade, 1984), the Umano-Fukami model (Umano, 1982, 1983; Umano & Fukami, 1994), and the GEFRED model of Medina-Pons-Vila (1994; Galindo, Medina, & Aranda, 1999; Galindo et al., 2001; Medina, 1994).

Imprecision without Fuzzy Logic

In this section, some ideas allowing for imprecise information treatment will be summarized, *without* utilizing either the fuzzy set theory or possibility theory. In the bibliography, these models are dealt with globally in the section on imprecision in conventional databases, although some of the ideas discussed here have not been implemented in any of the models. The first attempt to represent imprecise information on databases was the introduction of NULL values by Codd (1979), which was further expanded (Codd, 1986, 1987, 1990). This model did not use the fuzzy set theory. A NULL value in an attribute indicates that such a value is any value included in the domain of such an attribute.

Any comparison with a NULL value originates an outcome that is neither True (T) nor False (F) called "*maybe*" (m) (or unknown, in the SQL of Oracle). The truth tables of the classical comparators NOT, AND, and OR can be seen in Table 4.

Later on, another nuance was added, differentiating the NULL value in two marks: The "A-mark" representing an absent or unknown value, although it was applicable, and the "I-mark" representing the absence of the value because it is not applicable (undefined). An I-mark may be situated, for instance, in the car plate attribute of someone who does not have a car. This is a tetravalued logic where the A value, having a similar meaning to that of the m in the trivalued logic mentioned above, is generated by comparing any value containing an A-mark, and a new I value is added as a result of the comparison of any value containing an I-mark. The tetravalued logic is shown in Table 5.

In Galindo et al. (2006), some other approaches are summarized, like the "default values" approach by Date (1986), similar to the DEFAULT clause in SQL, the "interval values" approach by Grant (1980), who expands the relational model in order to allow that a possible value range/interval be stored in one attribute, and statistical and probabilistic databases.

Basic Model of Fuzzy Databases

The simplest model of fuzzy relational databases consists of adding a *grade*, normally in the [0,1] interval, to each instance (or tuple). This keeps database data homogeneity. Nevertheless, the semantic assigned to this grade will determine its usefulness, and this meaning will be utilized in the query processes. This grade may have the meaning of *membership degree* of each tuple to the relation (Giardina, 1979; Mouaddib, 1994), but it may mean something different, like the *dependence strength level* between two attributes, thus representing the relation between them (Baldwin, 1983), the *fulfillment degree* of a condition or the *importance degree* (Bosc, Dubois, Pivert, & Prade, 1997) of each tuple in the relation, among others.

The main problem with these fuzzy models is that they do not allow the representation of imprecise information about a certain attribute of

Table 4. Truth tables for the trivalued logic: True, false, and maybe

NOT	
T	F
m	m
F	T

AND	T	m	F
T	T	m	F
m	m	m	F
F	F	F	F

OR	T	m	F
T	T	T	T
M	T	m	m
F	T	m	F

Table 5. Truth tables for the tetravalued logic

NOT	
T	F
A	A
I	I
F	T

AND	T	A	I	F
T	T	A	I	F
A	A	A	I	F
I	I	I	I	F
F	F	F	F	F

OR	T	A	I	F
T	T	T	T	T
A	T	A	A	A
I	T	A	I	F
F	T	A	F	F

a specific entity (like the "tall" or "short" values for a "height" attribute). Besides, the fuzzy character is assigned globally to each instance (tuple) it making impossible to determine the specific fuzzy contribution from each constituting attribute. These problems are solved in the model presented in Galindo et al. (2006), and you can learn about this model in the chapter by Ben Hassine et al. in this handbook.

Similarity Relations Model: Buckles-Petry Model

This is the first model that utilizes similarity relations (Zadeh, 1971) in the relational model. It was proposed by Buckles and Petry (1982a, 1982b, 1984). In this model, a *fuzzy relation* is defined as a subset of the following Cartesian product: $P(D_1)\times \ldots \times P(D_m)$, where $P(D_i)$ represents the parts set of a D_i domain, including all the subsets that could be considered within the D_i domain (having any number of elements). The data types permitted by this model are finite set of scalars (labels), finite set of numbers, and fuzzy number set. The meaning of these sets is disjunctive, that is, the real value is one belonging to the set.

The equivalence types on a domain are constructed from a *similarity function or relation*, in which the values taken by such a relation are provided by the user. Typically, these similarity values are standardized in the [0,1] interval, where 0 corresponds to "totally different" and 1 to "totally similar." A *similarity threshold* can be established with a value between 0 and 1 in order to get the values whose similarity is greater than the threshold, or to consider those values indistinguishable.

Possibilistic Models

Under this denomination, models using the possibility theory to represent imprecision are included. The most important models in this group are Prade-Testemale model, Umano-Fukami model, and GEFRED model. Another important model is the the Zemankova-Kaendel model (1984, 1985), which is briefly summarized in Galindo et al. (2006).

Prade-Testemale Model

Prade and Testemale published a fuzzy relational database (FRDB) model that allows the integration of what they call *incomplete* or *uncertain* data in the possibility theory sphere (Prade, 1984; Prade & Testemale, 1984, 1987a, 1987b). An attribute A, having a D domain, is considered. All the available knowledge about the value taken by A for an x object can be represented by a *possibility distribution* $\pi_{A(x)}$ about $D \cup \{e\}$, where e is a special element denoting the case in which A is not applied to x. In other words, $\pi_{A(x)}$ is an application that goes from $D \cup \{e\}$ to the [0,1] interval. From this formulation, all value types adopted by this model can be represented.

In every possibilistic model one must take into account that, for a value $d \in D$, if $\pi_{A(x)}(d) = 1$, then this just indicates that the d value is totally possible for A(x), and not that the d value is true for A(x),

Table 6. Representation of information in two possibilistic models

Information	Prade-Testemale Model	Umano-Fukami Model
The precise data are known and this is *crisp*: c	$\pi_{A(x)}(e) = 0$ $\pi_{A(x)}(c) = 1$ $\pi_{A(x)}(d) = 0, \quad \forall\, d \in D, d \neq c$	$\pi_{A(x)}(d) = \{1 / c\}$
Unknown but applicable	$\pi_{A(x)}(e) = 0$ $\pi_{A(x)}(d) = 1, \quad \forall\, d \in D$	Unknown (Equation 67)
Not applicable or nonsense	$\pi_{A(x)}(e) = 1$ $\pi_{A(x)}(d) = 0, \quad \forall\, d \in D$	Undefined (Equation 68)
Total ignorance	$\pi_{A(x)}(d) = 1, \quad \forall\, d \in D \cup \{e\}$	Null (Equation 69)
Range [m, n]	$\pi_{A(x)}(e) = 0$ $\pi_{A(x)}(d) = 1, \quad \text{if } d \in [m, n] \subseteq D$ $\pi_{A(x)}(d) = 0, \quad \text{in other case}$	$\pi_{A(x)}(d) = 1, \text{ if } d \in [m, n] \subseteq D$ $\pi_{A(x)}(d) = 0, \quad \text{in other case}$
The information available is a possibility distribution μ_a	$\pi_{A(x)}(e) = 0$ $\pi_{A(x)}(d) = \mu_a(d), \quad \forall\, d \in D$	$\pi_{A(x)}(d) = \mu_a(d), \quad \forall\, d \in D$
The possibility that it may not be applicable is λ and, in case it is applicable, the data are μ_a	$\pi_{A(x)}(e) = \lambda$ $\pi_{A(x)}(d) = \mu_a(d), \quad \forall\, d \in D$	Without representation

unless this is the only possible value, that is, $\pi_{A(x)}(d') = 0, \forall d' \neq d$. Both the information and representation of this model is shown in Table 6. How two possibility distributions can be compared was discussed in the Comparison Operations on Fuzzy Sets section earlier in this chapter. In general, the most commonly used measurements are possibility and necessity.

Umano-Fukami Model

This proposal (Umano, 1982, 1983; Umano & Fukami, 1994) also utilizes the possibility distributions in order to model information knowledge. In this model, if D is the discourse universe of A(x), $\pi_{A(x)}(d)$ represents the possibility that A(x) takes the value $d \in D$. The following kind of knowledge may be modeled: unknown and applicable information, the non-applicable information (undefined), and the total ignorance (we do not know if it is applicable or non-applicable):

$$\text{Unknown} = \pi_{A(x)}(d) = 1, \quad \forall\, d \in D \tag{67}$$

$$\text{Undefined} = \pi_{A(x)}(d) = 0, \quad \forall\, d \in D \tag{68}$$

$$\text{Null} = \{1/\text{Unknown},\ 1/\text{Undefined}\} \tag{69}$$

For the remaining cases of imprecise information, a similar model to the one above is adopted. The kind of fuzzy information and the representation of this model are shown in Table 6. Besides, every instance of a relation in this model has a possibility distribution associated with it in the [0,1] interval, thus indicating the membership degree of that particular instance to such a relation. In other words, a fuzzy relation R, with m attributes, is defined as the following membership function:

$$\mu_R: P(U_1) \times P(U_2) \times \ldots \times P(U_m) \rightarrow P([0,1]) \tag{70}$$

where the × symbol denotes the Cartesian product, $P(U_j)$ with j=1, 2, ..., m is the collection of all the possibility distributions in the discourse universe U_j of the j-th R attribute. The function μ_R associates a P([0,1]) value to every instance of the relation R, which corresponds to all the possibility distributions in the [0,1] interval; this shall be considered an R membership degree of such an instance. Finally, in the query process, expressed either in fuzzy or precise terms, the model solves the query problem by dividing the set of instances involved in the relation into three subsets, where the first subset contains the instances completely satisfying the query; the second subset groups those instances that might satisfy the query; and the third subset consists of those instances which do not satisfy the query.

The GEFRED Model by Medina-Pons-Vila

The GEFRED model dates back to 1994, and it has experienced subsequent expansions (Medina, 1994; Medina et al., 1994; Galindo et al., 1999, 2001, 2006). This model is an eclectic synthesis of some of the previously discussed models. One of the major advantages of this model is that it consists of a general abstraction that allows for the use of various approaches, regardless of how different they might look. As a possibilistic model, it refers particularly to generalized fuzzy domains, thus admitting the possibility distribution in the domains, but it also includes the case where the underlying domain is not numeric but scalars of any type. It includes UNKNOWN, UNDEFINED, and NULL values as well, having the same sense as that in Umano-Fukami model.

The GEFRED model is based on the definition which is called Generalized Fuzzy Domain (D) and Generalized Fuzzy Relation (R), which include classic domains and classic relations, respectively. Basically, the Generalized Fuzzy Domain is the basic domain, with possibility distributions defined for this domain and the NULL value. All data types that can be represented are shown in Table 1 in the chapter by Ben Hassine et al.

On the other hand, the Generalized Fuzzy Relations of GEFRED model are relations whose attributes have a Generalized Fuzzy Domain, and each attribute may be associated to a "compatibility attribute" where we can store a compatibility degree. The compatibility degree for an attribute value is obtained by manipulation processes (such as queries) performed on that relation, and it indicates the degree to which that value has satisfied or met the operation performed on it.

The GEFRED model defines fuzzy comparators that are general comparators based on any existing classical comparator (>, <, =, etc.), but it does not consolidate the definition of each one. The only requirement established is that the fuzzy comparator should respect the classical comparators outcomes when comparing possibility distributions expressing nonfuzzy values (*crisp*). For example, the "approximately equal" comparator, "possibly equal" or "fuzzy equal" (FEQ), may be defined using the possibility measure (see the Comparison Operations on Fuzzy Sets section in this chapter).

On these definitions, GEFRED redefines the relational algebraic operators in the so-called Generalized Fuzzy Relational Algebra: union, intersection, difference, Cartesian product, projection, selection, join, and division. These operators are defined giving a generalized fuzzy relation, which is the result of the operation. All these operators are defined in the definition of GEFRED, but the fuzzy division is defined in Galindo et al. (2001). Fuzzy Relational Calculus is defined in Galindo et al. (1999). These theoretical concepts have been applied in an extended fuzzy version of SQL for fuzzy queries, the so-called FSQL (Galindo et al., 2006). Some of these definitions have been implemented in a free FSQL server (Galindo, 2007). The characteristics of FSQL are summarized for example in the aforementioned chapters by Urrutia et al. and by Ben Hassine et al. Applications in the data mining fields are proposed in some works (Carrasco, Vila, & Galindo, 2003; Galindo et al., 2006), and it is shown in the chapter by Carrasco et al. in this handbook. An extension for Fuzzy

Deductive Relational Databases was presented by Blanco, Cubero, Pons, and Vila (2000).

CONCLUSION AND FUTURE TRENDS

This chapter presents an introduction to fuzzy logic and to fuzzy databases. With regard to the first topic, we have introduced concepts like fuzzy sets and fuzzy numbers, fuzzy logic and its main characteristics, linguistic labels, membership functions and their determination methods, support and kernel, height and cardinality, α-cut, the representation theorem and the extension principle, fuzzy set operations like union and intersection (t-norms and t-conorms), negations, fuzzy implications, different comparison operations, fuzzy relations, fuzzy quantifiers, and the possibility theory.

With respect to the fuzzy databases, this chapter gives a list of six research topics in this *fuzzy* area. All these topics are briefly commented on, and we include references to books, papers, and even to other chapters of this handbook, where we can find some interesting reviews about different subjects and new approaches with different goals. Finally, we summarize the main published models approaching the problem of representation and treatment of imprecise information in relational databases, including methods with and without fuzzy logic.

Perhaps the main difficulty in fuzzy database technology is to solve the subjectivity problem in the definition and usage of fuzzy concepts, and the dependency between these concepts and the context. Bordogna and Psaila study in their chapter, in this handbook, some ideas about this context dependency, but there are other problems that should be studied, like fuzzy interfaces or that, in many applications, possibly it is better to use fuzzy intervals or approximate values instead of predefined fuzzy labels. Another interesting future line is to introduce fuzzy data types in current database management systems (relational, object-oriented, etc.) and/or introduce fuzzy comparators and some other tools of fuzzy queries in the standard SQL. In this line, some researches of this volume may be very useful, as for example, the chapters by Urrutia et al., Ben Hassine et al., Barranco et al., Scheneider, or Takači and Škrbić. We can see that this is not only a theoretical research field and this book includes some interesting applications like those in the chapters by Veryha et al., Chen et al., Carrasco et al., or Xexéo and Braga. Surely, in some years, what now is a research proposal, will be a running application, and in this sense, this book includes many interesting works as, for example, those about fuzzy queries and fuzzy data mining.

ACKNOWLEDGMENT

This work has been partially supported by the "Ministry of Education and Science" of Spain (projects TIN2006-14285 and TIN2006-07262) and the Spanish "Consejería de Innovación Ciencia y Empresa de Andalucía" under research project TIC-1570.

REFERENCES

ANSI (American National Standard Institute). (1992). *Database language SQL* (Tech. Rep. No. ANXI X3, 135-1992). New York.

Baldwin, J. (1983). Knowledge engineering using a fuzzy relational inference language. In *Proceedings of the IFAC Symposium on Fuzzy Information Knowledge Representation and Decision Analysis* (pp. 15-21).

Bezdek, J. (1981). *Pattern recognition with fuzzy objective function algorithms.* New York: Plenum Press.

Blanco, I., Cubero, J.C., Pons, O., & Vila, M.A. (2000). An implementation for fuzzy deductive relational databases. In G. Bordogna and G. Pasi (Eds.), *Recent issues on fuzzy databases* (pp. 183-

207). Physica-Verlag (Studies in Fuzziness and Soft Computing).

Bosc, P. (1999). Fuzzy databases. In J. Bezdek (Ed.), *Fuzzy sets in approximate reasoning and information systems* (pp. 403-468). Boston: Kluwer Academic Publishers.

Bosc, P., Dubois, D., Pivert, O., & Prade, H. (1997). Flexible queries in relational databases: The example of the division operator. *Theoretical Computer Science, 171*, 281-302.

Bosc, P., & Pivert, O. (1995). SQLf: A relational database language for fuzzy querying. *IEEE Transactions on Fuzzy Systems, 3*, 1-17.

Buckles, B.P., & Petry, F.E. (1982a). A fuzzy representation of data for relational databases. *Fuzzy Sets Systems, 7*, 213-226.

Buckles, B.P., & Petry, F.E. (1982b). Fuzzy databases and their applications. In M. Gupta & E. Sanchez (Eds.), Fuzzy information and decision processes (vol. 2, pp. 361-371). Amsterdam: North-Holland.

Buckles, B.P., & Petry, F.E. (1984). Extending the fuzzy database with fuzzy numbers. *Information Sciences, 34*, 45-55.

Buckley, J.J., & Eslami, E. (2002). *An introduction to fuzzy logic and fuzzy sets* (advances in soft computing). Physica-Verlag Heidelberg.

Butnario, D., & Klement, E.P. (1993). *Triangular norm-based measures and games with fuzzy coalitions.* Dordrecht: Kluwer Academic Publishers.

Carrasco, R., Vila, M.A., & Galindo, J. (2003). FSQL: A flexible query language for data mining. *Enterprise Information Systems, IV*, 68-74. Hingham, MA: Kluwer Academic Publishers.

Chen, G. (1998). *Fuzzy logic in data modeling: Semantics, constraints and database design*. London: Kluwer Academic Publishers.

Codd, E.F. (1979). Extending the database relational model to capture more meaning. *ACM Transactions on Database Systems, 4*, 262-296.

Codd, E.F. (1986). Missing information (applicable and inapplicable) in relational databases. *ACM SIGMOD Record, 15*(4).

Codd, E.F. (1987). More commentary on missing information in relational databases. *ACM SIGMOD Record, 16*(1).

Codd, E.F. (1990). *The relational model for database management* (version 2). Reading, MA: Addison-Wesley.

Date, C.J. (1986). Null values in database management. In C.J. Date (Ed.), *Relational database: Selected writings*. Reading MA: Addison-Wesley.

Date, C.J., & Darwen, H. (1997). A guide to SQL standard (4th ed.). Addison-Wesley. ISBN 0-201-96426-0.

De Caluwe, R. (1997). Fuzzy and uncertain object-oriented databases: Concepts and models. In R. De Caluwe (Ed.), *Advances in fuzzy system: Application and theory* (vol. 13). World Scientific.

De Caluwe, R., & De Tré, G. (Eds.). (2007). Special issue on advances in fuzzy database technology. *International Journal of Intelligent Systems, 22*.

Delgado, M., Sánchez, D., & Vila, M.A. (1999). A survey of methods for evaluating quantified sentences. In *Proceedings of the EUSFLAT-ESTYLF Joint Conference* (pp. 279-282). Palma de Mallorca (Spain).

Delgado, M., Sánchez, D., & Vila, M.A. (2000). Fuzzy cardinality based evaluation of quantified sentences. *International Journal of Approximate Reasoning, 23*, 23-66.

Dubois, D., & Prade, H. (1980). Fuzzy sets and systems: Theory and applications. New York: Academic Press.

Dubois, D., & Prade, H. (1984). A theorem on implication functions defined from triangular norms. *Stochastica, 8*, 267-279. Also in (1993), D. Dubois, H. Prade, & R.R. Yager (Eds.), *Readings in fuzzy sets for intelligent systems* (pp. 105-112). Morgan & Kaufmann.

Dubois, D., & Prade, H. (1985). Fuzzy number: An overview. In J.C. Bezdek (Ed.), *The analysis of fuzzy information*. Boca Raton, FL: CRS Press.

Dubois, D., & Prade, H. (1988). *Possibility theory: An approach to computerized processing of uncertainty.* New York: Plenum Press.

Escobar, C. (2003). Software para Control Difuso de Todo Tipo de Sistemas (SCD): Aplicación al Control de Invernaderos Industriales. Proyecto Fin de Carrera de Ingeniería Técnica Industrial en Electrónica (Universidad de Málaga), directed by J. Galindo: www.lcc.uma.es/~ppgg/PFC

Galindo, J. (1999). Tratamiento de la Imprecisión en Bases de Datos Relacionales: Extensión del Modelo y Adaptación de los SGBD Actuales. Ph. Doctoral Thesis, University of Granada, Spain. Retrieved from: www.lcc.uma.es.

Galindo, J. (2001). Curso sobre Conjuntos y Sistemas Difusos (Lógica Difusa y Aplicaciones). Informe Técnico de Docencia. LCC-ITI-2001-11, del Dpto. de Lenguajes y Ciencias de la Computación de la Universidad de Málaga (www.lcc.uma.es). Retrieved from: www.lcc.uma.es/~ppgg/FSS

Galindo, J. (2007). *FSQL (fuzzy SQL): A fuzzy query language.* Retrieved January 12, 2008, from http://www.lcc.uma.es/~ppgg/FSQL

Galindo, J., Medina, J.M., & Aranda, M.C. (1999). Querying fuzzy relational databases through fuzzy domain calculus. *International Journal of Intelligent Systems, 14*(4), 375-411.

Galindo, J., Medina, J.M., Cubero, J.C., & García, M.T. (2001, June). Relaxing the universal quantifier of the division in fuzzy relational databases. *International Journal of Intelligent Systems, 16*(6), 713-742.

Galindo, J., Urrutia, A., & Piattini, M. (2006). *Fuzzy databases: Modeling design and implementation*. Hershey, PA: Idea Group.

Giardina, C. (1979). *Fuzzy databases and fuzzy relational associative processors* (Technical Report). Hoboken, NJ: Stevens Institute of Technology.

Goncalves, M., & Tineo, L. (2006). SQLf vs. Skyline: Expressivity and performance. In *Proceedings of the 15th IEEE International Conference on Fuzzy Systems Fuzz-IEEE 2006* (pp. 2062-2067), Vancouver, Canada.

Grant, J. (1980). Incomplete information in a relational database. *Fundamenta Informaticae, 3*, 363-378.

Gupta, M.M., & Qi, J. (1991). Theory of t-norms and fuzzy inference methods. *Fuzzy Sets and Systems, 40*, 431-450.

Kerre, E.E., & Chen, G.Q. (2000). Fuzzy data modeling at a conceptual level: Extending ER/EER concepts. In O. Pons (Ed.), *Knowledge management in fuzzy databases* (pp. 3-11). Heidelberg: Physica-Verlag.

Kosko, B. (1992). *Neural networks and fuzzy systems: A dynamical systems approach to machine intelligence.* Englewood Cliffs, NJ: Prentice Hall.

Kruse, R., Gebhardt, J., & Klawonn, F. (1994). *Foundations of fuzzy systems.* John Wiley & Sons.

Liu, Y., & Kerre, E.E. (1998a). An overview of fuzzy quantifiers. (I). Interpretations. *Fuzzy Sets and Systems, 95*(1), 1-21.

Liu, Y., & Kerre, E.E. (1998b). An overview of fuzzy quantifiers. (II). Reasoning and applications. *Fuzzy Sets and Systems, 95*(2), 135-146.

Ma, Z. (2005). *Fuzzy database modeling with XML*. Springer-Verlag.

Medina, J.M. (1994). Bases de datos Relacionales Difusas: Modelo Teórico y Aspectos de su Implementación. Ph. Doctoral Thesis, Universidad de Granada, España. Retrieved from: decsai.ugr.es.

Medina, J.M., Pons, O., & Vila, M.A. (1994). GEFRED: A generalized model of fuzzy relational databases. *Information Sciences, 76*(1-2), 87-109.

Menger, K. (1942). Statistical metric spaces. *Proceedings of the National Academy of Sciences, 37*, 535-537.

Mohammd, J., Vadiee, N., & Ross, T.J. (Eds.). (1993). *Fuzzy logic and control: Software and hardware applications.* Eaglewood Cliffs, NJ: Prentice Hall PTR.

Mouaddib, N. (1994). Fuzzy identification database: The nuanced relation division. *International Journal of Intelligent System, 9*, 461-473.

Nguyen, H.T., & Walker, E.A. (2005). *A first course in fuzzy logic* (3rd ed.). Chapman & Hall/CRC.

Pedrycz, W., & Gomide, F. (1998). *An introduction to fuzzy sets: Analysis and design* (A Bradford Book). The MIT Press.

Petry, F.E. (1996). *Fuzzy databases: Principles and applications (with chapter contribution by Patrick Bosc)* (International Series in Intelligent Technologies). Kluwer Academic Publishers.

Piegat, A. (2001). *Fuzzy modeling and control.* Physica-Verlag (Studies in Fuzziness and Soft Computing).

Prade, H. (1984). Lipski's approach to incomplete information databases restated and generalized in the setting of Zadeh's possibitity theory. *Information Systems, 9*, 27-42.

Prade, H., & Testemale, C. (1984). Generalizing database relational algebra for the treatment of incomplete/uncertain information and vague queries. *Information Sciences, 34*, 115-143.

Prade, H., & Testemale, C. (1987a). Fuzzy relational databases: Representational issues and reduction using similarity measures. *Journal of the American Society of Information Sciences, 38*(2), 118-126.

Prade, H., & Testemale, C. (1987b). Representation of soft constraints and fuzzy attribute values by means of possibility distributions in databases. In J. Bezdek (Ed.), Analysis of fuzzy information (vol. II: Artificial intelligence and decision systems, pp. 213-229). CRC Press.

Ross, T.J. (2004). *Fuzzy logic with engineering applications.* Wiley.

Saaty, T.L. (1980). *The analytic hierarchy processes.* New York: McGraw-Hill.

Schweizer, B., & Sklar, A. (1983). *Probabilistic metric spaces.* North-Holland.

Sivanandam, S.N., Sumathi, S., & Deepa, S.N. (2006). *Introduction to fuzzy logic using MATLAB.* Springer.

Trillas, E. (1979). Sobre funciones de negación en la Teoría de Conjuntos Difusos. *Stochastica, 3*(1), 47-59.

Trillas, E., & Alsina, C., (2002). On the law $[p \wedge q \rightarrow r] = [(p \rightarrow r) \vee (q \rightarrow r)]$ in fuzzy logic. *IEEE Transactions on Fuzzy Systems, 10*(1), 84-88.

Trillas, E., Alsina, C., & Pradera, A. (2004). On MPT-implication functions for fuzzy logic. Spanish "Rev. Real Academia de Ciencias", Serie A. Mathematics (RACSAM), vol. 98 (1), pp. 259–271.

Trillas, E., Cubillo, S., & del Campo, C. (2000). When QM-operators are implication functions and conditional fuzzy relations? *International Journal Intelligence Systems, 15*, 647-655.

Umano, M. (1982). Freedom-O: A fuzzy database system. In M. Gupta & E. Sanchez (Eds.), Fuzzy information and decision processes (pp. 339-347). Amsterdam: North-Holland.

Umano, M. (1983). Retrieval from fuzzy database by fuzzy relational algebra. In M. Gupta & E. Sanchez (Eds.), *Fuzzy information, knowledge representation and decision analysis* (pp. 1-6). New York: Pergamon Press.

Umano, M., & Fukami, S. (1994). Fuzzy relational algebra for possibility-distribution-ruzzy-relation model of fuzzy data. *Journal of Intelligent Information System, 3*, 7-28.

Yager, R.R. (1980). On a general class of fuzzy connectives. *Fuzzy Sets and Systems,* 235-242.

Yager, R.R. (1983). Quantified propositions of a linguistic logic. *International Journal of Man-Machine Studies, 19*, 195-227.

Yager, R.R., Ovchinnikov, S., Tong, R.M., & Nguyen, H.T. (Eds.). (1987). *Fuzzy sets and applications: Selected papers by L. A. Zadeh.* John Wiley & Sons.

Yazici, A., & George, R. (1999). *Fuzzy database modeling.* Physica-Verlag.

Ying, M. (2002). Implication operators in fuzzy logic. *IEEE Transactions on Fuzzy Systems, 10*(1), 88-91.

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

Zadeh, L.A. (1971). Similarity relations and fuzzy orderings. *Information Sciences, 3*, 177-200.

Zadeh, L.A. (1975). The concept of a linguistic variable and its application to approximate reasoning (parts I, II, and III). *Information Sciences, 8*, 199-251, 301-357; *9*, 43-80.

Zadeh, L.A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems, 1*, 3-28.

Zadeh, L.A. (1983). A computational approach to fuzzy quantifiers in natural languages. *Computer Mathematics with Applications, 9*, 149-183.

Zadeh, L.A. (1992). *Knowledge representation in fuzzy logic: An introduction to fuzzy logic applications in intelligent systems.* Kluwer Academic Publisher.

Zemankova-Leech, M., & Kandel, A. (1984). *Fuzzy relational databases: A key to expert systems.* Köln, Germany: Verlag TÜV Rheinland.

Zemankova-Leech, M., & Kandel, A. (1985). Implementing imprecision in information systems. *Information Sciences, 37*, 107-141.

Zimmermann, H.-J. (1991). *Fuzzy set theory and its applications* (2nd ed.). Kluwer Academic Publishers.

KEY TERMS

Fuzzy Attribute: In a database context, a fuzzy attribute is an attribute of a row or object in a database, with a fuzzy datatype, which allows storing fuzzy information. Sometimes, if a classic attribute allows fuzzy queries, then it is also called fuzzy attribute, because it has only some of the fuzzy attribute characteristics.

Fuzzy Database: If a regular or classical database is a structured collection of information (records or data) stored in a computer, a fuzzy database is a database which is able to deal with uncertain or incomplete information using fuzzy logic. There are many forms of adding flexibility in fuzzy databases. The simplest technique is to add a fuzzy membership degree to each record, that is, an attribute in the range [0,1]. However, there are other kinds of databases allowing fuzzy values to be stored in fuzzy attributes using fuzzy sets, possibility distributions, or fuzzy degrees associated to some attributes and with different meanings (membership degree, importance degree, fulfillment degree, etc.). Of course, fuzzy databases should allow fuzzy queries using fuzzy or nonfuzzy data and there are some languages that allow this kind of queries, like FSQL or SQLf. In synthesis, the research in fuzzy databases includes the following areas: flexible querying in classical or fuzzy databases, extending classical data models in order to achieve fuzzy databases (fuzzy relational databases, fuzzy object-oriented databases, etc.), fuzzy conceptual modeling, fuzzy data mining techniques, and applications of these advances in real databases.

Fuzzy Implication: Function computing the fulfillment degree of a rule expressed by IF X THEN Y, where the antecedent and the consequent are fuzzy. These functions must comply with certain basic properties and the most typical is the Kleene-Dienes implication, based on the classical implication definition ($x \Rightarrow y = \neg x \vee y$), using the Zadeh's negation and the maximum s-norm, but other fuzzy implication functions exist (Table 3).

Fuzzy Logic: Fuzzy logic is derived from fuzzy set theory by Zadeh (1965), dealing with reasoning that is approximate rather than precisely deduced from classical predicate logic. It can be thought of as the application side of fuzzy set theory dealing with well thought out real world expert values for a complex problem.

Fuzzy Quantifiers: Expressions allowing us to express fuzzy quantities or proportions in order to provide an approximate idea of the number of elements of a subset fulfilling a certain condition or of the proportion of this number in relation to the total number of possible elements. Fuzzy quantifiers can be absolute or relative. Absolute quantifiers express quantities over the total number of elements of a particular set, stating whether this number is, for example, "much more than 10," "close to 100," "a great number of," and so forth. Relative quantifiers express measurements over the total number of elements, which fulfill a certain condition depending on the total number of possible elements. This type of quantifier is used in expressions such as "the majority" or "most," "the minority," "little of," "about half of," and so on.

Possibility Theory: This theory is based on the idea that we can evaluate the possibility of a determinate variable X being (or belonging to) a determinate set or event A. Here, fuzzy sets are called possibility distributions and instead of measuring the membership degrees, they measure the possibility degrees. All the tools and properties defined for fuzzy sets are also applicable to possibility distributions.

Soft Computing: Computational techniques in computer science and some engineering disciplines, which attempt to study, model, and analyze very complex phenomena: those for which more conventional methods have not yielded low cost, analytic, and complete solutions. Earlier computational approaches could model and precisely analyze only relatively simple systems. More complex systems arising in biology, medicine, the humanities, management sciences, artificial intelligence, machine learning, and similar fields often remained intractable to conventional mathematical and analytical methods. Soft computing techniques include: fuzzy systems (FS), neural networks (NN), evolutionary computation (EC), probabilistic reasoning (PR), and other ideas (chaos theory, etc.). Soft computing techniques often complement each other.

T-conorm or S-norm: Function *s* establishing a generic model for the operation of union with fuzzy sets. These functions must comply with certain basic properties: commutative, associative, monotonicity, and border conditions ($x \, s \, 0 = x$, and $x \, s \, 1 = 1$). The most typical is the maximum function, but other widely accepted s-norms exist (Table 2).

T-norm: Function *t* establishing a generic model for the operation of intersection with fuzzy sets. These functions must comply with certain basic properties: commutative, associative, monotonicity, and border conditions ($x \, t \, 0 = 0$, and $x \, t \, 1 = x$). The most typical is the minimum function, but there exists other t-norms widely accepted (Table 1).

ENDNOTES

1. "Fuzzy Sets and Systems" is an international journal in information science and engineering, published by Elsevier (www.elsevier.com), Official Publication of the International Fuzzy Systems Association (IFSA). Currently, the co-editors-in-chief are B. De Baets and D. Dubois. Website: http://www.elsevier.com/locate/fss
2. "IEEE Transactions on Fuzzy Systems" is also an international journal, published by IEEE (Institute of Electrical and Electronics Engineers, www.ieee.org). Currently, the editor-in-chief is N.R. Pal. Website: http://ieee-cis.org/pubs/tfs
3. L.A. Zadeh is a professor in the University of California, Berkeley. He has received many awards all over the world and holds honorary doctorates in universities worldwide, one of

them being from the University of Granada, Spain, in 1996, in recognition of his important contribution in this scientific field. Website: http://www.cs.berkeley.edu/~zadeh