Chapter 2
Fuzzy Set Theory

This chapter aims to present the main concepts and mathematical notions of the fuzzy set theory (also called fuzzy logic or fuzzy logic theory[1]) which are necessary for the understanding of this work. It also gives a non exhaustive overview of the application domains where the fuzziness has been successfully applied. For this purpose, Sect. 2.1 introduces the fuzzy logic with respect to the human beings then Sect. 2.2 describes the concept of fuzzy sets, Sects. 2.3 and 2.4 define the main fuzzy sets' properties and operations, finally Sect. 2.5 presents some of the main application areas of the fuzzy set theory.

2.1 Human Beings and Fuzziness

The fuzzy set theory has been proposed in 1965 by Lofti A. Zadeh from the University of Berkeley [131]. This theory is based on the intuitive reasoning by taking into account the human subjectivity and imprecision. It is not an imprecise theory but a rigorous mathematical theory which deals with subjectivity and/or uncertainty which are common in the natural language. The natural language is a very complicated structure which is fundamental, not only in the human communication, but also in the way human beings think and perceive the surrounding world [94]. The main idea of the fuzzy logic is to capture the vagueness of the human thinking and to express it with appropriate mathematical tools [44]. More precisely, "the fuzzy logic provides a mathematical power for the emulation of the higher order cognitive functions, the thought and perception" [46].

Unlike computers, the human reasoning is not binary where everything is either yes (true) or no (false) but deals with imprecise concepts like 'a tall man', 'a moderate temperature' or 'a large profit'. These concepts are ambiguous in the sense that they cannot be sharply defined. For instance, the question whether a person is tall cannot

[1] Although fuzzy logic is an application of the fuzzy set theory extending the boolean logic, it is often used as a generic term encompassing the fuzzy set theory and its applications.

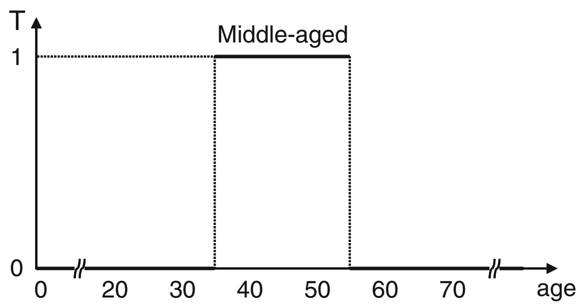
N. Werro, *Fuzzy Classification of Online Customers*, Fuzzy Management Methods,
DOI 10.1007/978-3-319-15970-6_2

be universally answered as some people will agree and others won't. Despite the fact that the definition of the word 'tall' is clear, it is not possible to sharply state if a person is tall because the answer may depend on the individual perception. Even for one person it may not be possible to give a clear and precise answer as the belonging to a concept (e.g. tall person) is often not sharp but fuzzy, involving a partial matching expressed in the natural language by the expressions 'very', 'slightly', 'more or less', etc.

Example 2.1 In order to illustrate this ambiguity, consider the concept 'middle-aged' for a person. This concept is clear in the people's mind however it is difficult to explicitly determine the precise beginning and ending years for a middle-aged person. Once again every individual might give a different definition of that concept. Let's assume that a survey states that a middle-aged person is between 35 and 55 years old. The concept 'middle-aged' can then be represented as a set illustrated in Fig. 2.1. This set represents the truth function of the concept 'middle-aged' according to the survey where the X-axis represents the age and the Y-axis contains the truth value. A truth value of 1 means that the age corresponds to the concept and a value of 0 indicates the age does not belong to that concept. This definition of the concept 'middle-aged' implies that a person who is 34 years old would suddenly become middle-aged on his next birthday. Similarly, just after his 56th birthday this person would no longer be middle-aged. This definition of the concept 'middle-aged' seems therefore unnatural as it does not match the human perception due to the sharply fixed boundaries.

A way of better modeling the imprecision of the human thinking is to introduce the notion of partial membership which allows a continuous transition between the different concepts. The notion of partial belonging can be represented by a fuzzy set (see Fig. 2.2). Fuzzy sets are the foundation of the fuzzy logic theory and are presented in Sect. 2.2. With this new definition a person enters the concept 'middle-aged' at the age of 20 with a continuous increment till the full belonging at the age of 45 and then progressively quits the concept. This way, there are no more extreme steps (sharp boundaries) such that just within a year somebody jumps into or out of the concept 'middle-aged'.

Fig. 2.1 Sharply defined concept of a middle-aged person



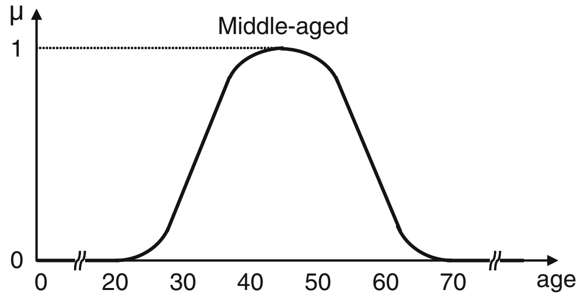


Fig. 2.2 Concept of a middle-aged person defined with a fuzzy set

The ambiguity is part of the human thinking and is ubiquitous in the natural language. Different aspects of ambiguity can be distinguished [88]:

- *Incompleteness*: The ambiguity is caused by a lack of information or knowledge. For example, a sentence in a foreign language cannot be understood if the given language is unknown.
- *Homonymy*: A word with several possible meanings might be ambiguous if the correct interpretation is not clear. The word 'trailer' for instance has two very different definitions; it can be a motion picture preview or a vehicle depending on the context in which it is mentioned.
- *Randomness*: The ambiguity comes from the fact that the result of an event is not known since it will happen in the future. An example is when a dice is thrown and it is not yet known which side will show up. This aspect of ambiguity is covered in mathematics by the probability theory.
- *Imprecision*: An information can be ambiguous due to its imprecision, i.e. the information contains errors or noise and is not exact.
- *Fuzziness*: This aspect covers the ambiguity with respect to words, that is to say the ambiguity of semantics. For instance, it is ambiguous whether a person is tall.

The fuzzy logic theory deals with the last kind of ambiguity, the fuzziness [88]. It proposes mathematical notions to model the imprecision of the human thinking. Considering that the fuzziness is ubiquitous and essential for the human beings, the fuzzy logic theory offers new perspectives for improving the human-machine interactions. One important aspect of this thesis is the ability of processing intuitive and human-oriented queries based on linguistic terms or expressions.

2.2 Concept of Fuzzy Sets

The fuzzy logic theory is based on fuzzy sets which are a natural extension of the classical set theory. A sharp set (also called crisp set) is defined by a bivalent truth function which only accepts the values 0 and 1 meaning that an element fully belongs to a set or does not at all, whereas a fuzzy set is determined by a membership function which accepts all the intermediate values between 0 and 1 (see Example 2.1). The values of a membership function, called membership degrees or grades of membership, precisely specify to what extent an element belongs to a fuzzy set, i.e. to the concept it represents.

Definition 2.1 A fuzzy set is built from a reference set called *universe of discourse*. The reference set is never fuzzy. Assume that $U = \{x_1, x_2, \ldots, x_n\}$ is the universe of discourse, then a *fuzzy set* A in U $(A \subset U)$ is defined as a set of ordered pairs

$$\{(x_i, \mu_A(x_i))\}$$

where $x_i \in U$, $\mu_A : U \to [0, 1]$ is the *membership function* of A and $\mu_A(x) \in [0, 1]$ is the *degree of membership* of x in A.

Example 2.2 Consider the universe of discourse $U = \{1, 2, 3, 4, 5, 6\}$. Then a fuzzy set A holding the concept 'large number' can be represented as

$$A = \{(1, 0), (2, 0), (3, 0.2), (4, 0.5), (5, 0.8), (6, 1)\}$$

With the considered universe, the numbers 1 and 2 are not 'large numbers', i.e. the membership degrees equal 0. Numbers 3–5 partially belong to the concept 'large number' with a membership degree of 0.2, 0.5 and 0.8. Finally number 6 is a large number with a full membership degree.

It is important to note that the definition of the membership degrees is subjective and context dependent, meaning that each person has his own perception of the concept 'large number' and that the interpretation is dependent on the universe of discourse and the context in which the fuzzy set is used. In Example 2.2 for instance, the membership degrees of the elements would be quite different if the universe of discourse contained numbers up to 100 or even 1000. In a similar manner, the concept 'large profit' would have a distinct signification for a small and a large enterprise.

Fuzzy sets are commonly represented by a membership function. Depending on the reference set, the membership functions are either discrete or continuous. Figure 2.3 shows the truth function of a sharp set in comparison to the membership functions of a discrete and a continuous fuzzy set.

Usually, several fuzzy sets are defined on the same reference set forming a fuzzy partition of the universe. A linguistic expression from the natural language can label the fuzzy sets in order to express their semantics. In the case of Example 2.1, the reference set can hold the concepts 'young', 'middle-aged' and 'old' at the same time allowing a continuous transition between them (see Fig. 2.4). This construct is essential in the fuzzy logic theory and is called a linguistic variable. A linguistic variable is a variable whose values are words or sentences instead of numerical values [134, 135, 136]. These values are called terms (also linguistic or verbal terms) and are represented by fuzzy sets.

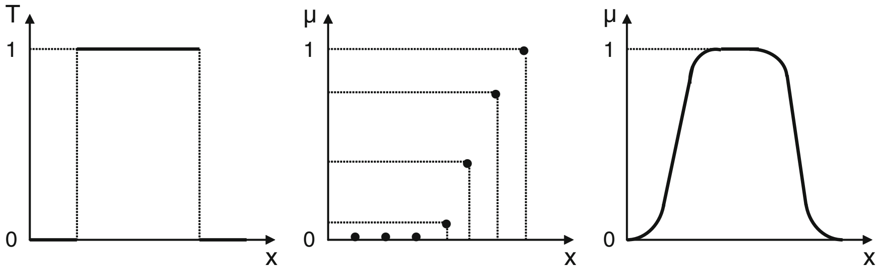
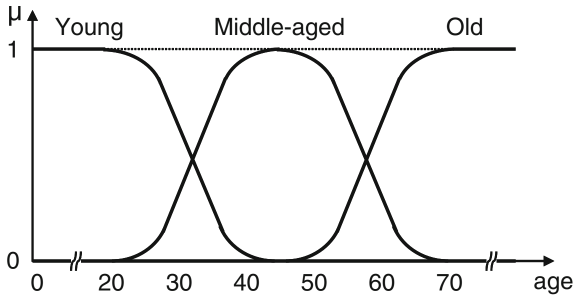


Fig. 2.3 Truth function of a sharp set and membership functions of a discrete and a continuous fuzzy set

Fig. 2.4 Fuzzy partition of the reference set with labelled fuzzy sets



Definition 2.2 A *linguistic variable* is characterized by a quintuple

$$(X, T, U, G, M)$$

where X is the name of the variable, T is the set of terms of X, U is the universe of discourse, G is a syntactic rule for generating the name of the terms and M is a semantic rule for associating each term with its meaning, i.e. a fuzzy set defined on U [134, 135, 136].

Example 2.3 The linguistic variable represented in Fig. 2.4 is defined by the quintuple (X, T, U, G, M) where X is 'age', T is the set {young, middle-aged, old} generated by G and M specifies for each term a corresponding fuzzy set on the universe $U = [0, 100]$.

The ability of giving a partial belonging to the elements allows a continuous transition between the fuzzy sets instead of having sharply fixed boundaries. This way, it is possible to better reflect the reality where everything is not black or white but often differentiated by grey values. The definition of a fuzzy set can therefore adequately express the subjectivity and the imprecision of the human thinking. Furthermore, the concept of linguistic variable is the basis for representing the human knowledge within human oriented rules or queries which can be processed by computers.

2.3 Properties of Fuzzy Sets

As the fuzzy set theory is an extension of the classical set theory, crisp sets are specific cases of the fuzzy sets. For this reason, the existing properties of the classical sets have to be extended and some new properties are introduced. Among the extended properties of the classical sets are the definitions of emptiness, equality, inclusion and cardinality. In order to take the wider scope of the fuzzy sets into account, the definitions of convexity, support, α-cut, kernel, width, height and normalization have been introduced.

A fuzzy set is considered to be empty if the membership degrees of all the elements of the universe are equal to zero.

Definition 2.3 A fuzzy set A, defined over a reference set U, is *empty* if

$$A = \varnothing \Leftrightarrow \mu_A(x) = 0, \ \forall x \in U$$

Two fuzzy sets are equal if their membership degrees are equal for all the elements of the reference set, i.e. if the two fuzzy sets have the same membership function.

Definition 2.4 Two fuzzy sets A and B, defined over a reference set U, are *equal* if

$$A = B \Leftrightarrow \mu_A(x) = \mu_B(x), \ \forall x \in U$$

A fuzzy set A is included in a fuzzy set B if the degrees of membership of A are smaller or equal to the membership degrees of B for all the elements of the universe (see Fig. 2.5).

Definition 2.5 Let A and B be two fuzzy sets defined over a reference set U, A is *included* in B if

$$A \subseteq B \Leftrightarrow \mu_A(x) \leq \mu_B(x), \ \forall x \in U$$

The cardinality of a crisp set equals the number of elements it contains. In a fuzzy set the elements can have a partial belonging, therefore the cardinality is the sum of the membership degrees of the reference set elements. If the reference set is infinite,

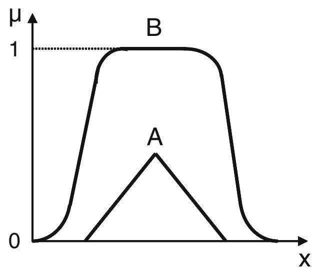


Fig. 2.5 Inclusion of the fuzzy set A in the fuzzy set B

an integral over the universe is used instead of the addition. It is also possible to derive the relative cardinality of a fuzzy set by dividing the cardinality of the fuzzy set by the cardinality of the universe. The relative cardinality allows fuzzy sets to be compared if they are based on the same universe.

Definition 2.6 The *cardinality* and the *relative cardinality* of a fuzzy set A, defined over a finite universe U, are defined as

$$Card(A) = |A| = \sum_{x \in U} \mu_A(x)$$

$$RelCard(A) = ||A|| = \frac{|A|}{|U|}$$

Example 2.4 Consider the fuzzy set A of Example 2.2.

$$A = \{(1, 0), (2, 0), (3, 0.2), (4, 0.5), (5, 0.8), (6, 1)\}$$

Then the cardinality and the relative cardinality of A are

$$Card(A) = 0 + 0 + 0.2 + 0.5 + 0.8 + 1 = 2.5$$

$$RelCard(A) = \frac{2.5}{6} \approx 0.417$$

Generally, linguistic notions are represented by a convex fuzzy set (see Fig. 2.6). A fuzzy set is convex if any point located between two other points has a higher membership degree than the minimum membership degree of these points.

Definition 2.7 A fuzzy set A defined over a reference set U is *convex* if

$$\forall x, y \in U, \forall \lambda \in [0, 1] : \mu_A(\lambda x + (1 - \lambda)y) \geq min(\mu_A(x), \mu_A(y))$$

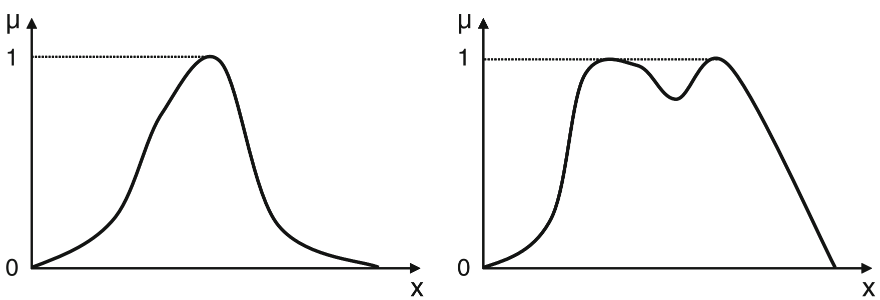


Fig. 2.6 Convex and non-convex fuzzy sets

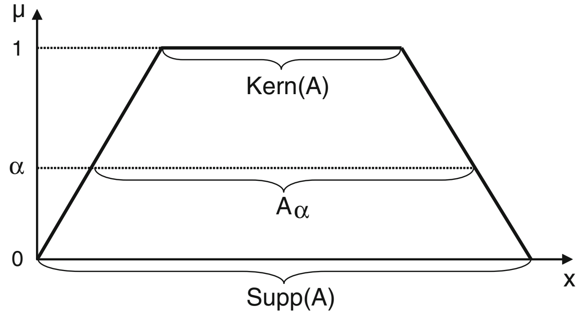


Fig. 2.7 Support, α-cut and kernel of a fuzzy set

The support of a fuzzy set is the sharp subset of the universe where the membership degrees are greater than zero (see Fig. 2.7).

Definition 2.8 The *support* of a fuzzy set A defined over a reference set U is a crisp subset of U that complies with

$$Supp(A) = \{x \in U,\ \mu_A(x) > 0\}$$

The α-cut, resp. the strong α-cut, of a fuzzy set is the crisp subset of the universe where the membership degrees are greater or equal, resp. greater, than the specified α value (see Fig. 2.7).

Definition 2.9 The *α-cut* and the *strong α-cut* of a fuzzy set A, defined over a reference set U, are a crisp subset of U that complies with

$$A_\alpha = \{x \in U,\ \mu_A(x) \geq \alpha,\ \alpha \in [0, 1]\}$$

$$A'_\alpha = \{x \in U,\ \mu_A(x) > \alpha,\ \alpha \in [0, 1]\}$$

The kernel of a fuzzy set is the crisp subset of the universe where the membership degrees are equal to 1 (see Fig. 2.7).

Definition 2.10 The *kernel* of a fuzzy set A defined over a reference set U is a crisp subset of U that complies with

$$Kern(A) = \{x \in U,\ \mu_A(x) = 1\}$$

The width of a convex fuzzy set is the length of the support, which in the case of a convex fuzzy set is an interval.

Definition 2.11 The *width* of a convex fuzzy set A with support $Supp(A)$, defined on a bounded reference set, is defined as

$$Width(A) = max(Supp(A)) - min(Supp(A))$$

The height of a fuzzy set is the maximum membership degree of all the elements of the universe.

Definition 2.12 The *height* of a fuzzy set A defined on a bounded reference set U is defined as

$$Hgt(A) = max_{x \in U}(\mu_A(x))$$

A fuzzy set is said to be normalized if at least one element of the universe has a membership degree equal to 1.

Definition 2.13 A fuzzy set A defined over a reference set U is *normalized* if and only if

$$\exists x \in U,\ \mu_A(x) = Hgt(A) = 1$$

2.4 Operations on Fuzzy Sets

The operations of complement, intersection and union of the classical set theory can also be generalized for the fuzzy sets. For these operations, several definitions with different implications exist. This section only presents the most common operators from the Zadeh's original proposition [131]. Further operators can be found in Appendix A.

The complement of a fuzzy set is 1 minus the membership degrees of the elements of the universe. This definition respects the notion of strong negation [41].

Definition 2.14 The *complement* of a fuzzy set A defined over a reference set U is defined as

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

For the intersection (resp. the union), Zadeh proposes to use the minimum operator (resp. the maximum operator). These operators have the advantages of being easily understandable and very fast to compute. The intersection (resp. the union) of two fuzzy sets is the minimum (resp. the maximum) value of the membership degrees of the two fuzzy sets for all the elements of the reference set (see Fig. 2.8).

Definition 2.15 The *intersection* of two fuzzy sets A and B defined over a reference set U is defined as

$$A \cap B = \mu_{A \cap B}(x) = \mu_A(x) \wedge \mu_B(x) = min(\mu_A(x), \mu_B(x)),\ x \in U$$

Definition 2.16 The *union* of two fuzzy sets A and B defined over a reference set U is defined as

$$A \cup B = \mu_{A \cup B}(x) = \mu_A(x) \vee \mu_B(x) = max(\mu_A(x), \mu_B(x)),\ x \in U$$

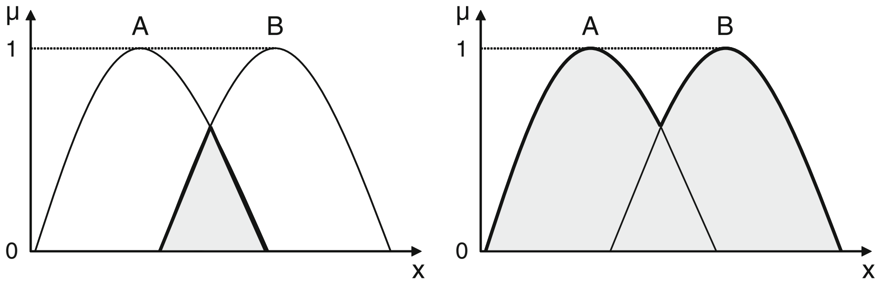


Fig. 2.8 Intersection and union of two fuzzy sets

Based on the intersection definition of two fuzzy sets, it is possible to introduce the notion of possibility (also called consistency or consensus) which is the fundament of the possibility theory which is briefly treated in Sect. 2.5.1. The possibility of two fuzzy sets, which determines the agreement degree between the concepts represented by the fuzzy sets, measures to what extent the fuzzy sets superpose each other and is defined as the highest membership degree of the intersection of these two fuzzy sets [137]. In the case of Zadeh's definition of the intersection this operation is called the max-min operation since it considers the maximum of the minimum values of the fuzzy sets.

Definition 2.17 The *possibility* of two fuzzy sets A and B defined over a reference set U is defined as

$$Poss(A, B) = max_{x \in U}(\mu_{A \cap B}(x)) = max_{x \in U}(min(\mu_A(x), \mu_B(x)))$$

Example 2.5 Consider the universe U and the fuzzy sets A and B.

$$U = \{a, b, c, d, e\}$$

$$A = \{(a, 0.4), (b, 0.8), (c, 1.0), (d, 0.8), (e, 0.2)\}$$

$$B = \{(a, 0.0), (b, 0.5), (c, 0.3), (d, 0.9), (e, 1.0)\}$$

Then the complement of A, the intersection, the union and the possibility of A and B are

$$\neg A = \{(a, 0.6), (b, 0.2), (c, 0.0), (d, 0.2), (e, 0.8)\}$$

$$A \cap B = \{(a, 0.0), (b, 0.5), (c, 0.3), (d, 0.8), (e, 0.2)\}$$

$$A \cup B = \{(a, 0.4), (b, 0.8), (c, 1.0), (d, 0.9), (e, 1.0)\}$$

$$Poss(A, B) = 0.8$$

It has to be noted that in contrast to the classical set theory, the intersection (resp. the union) of a fuzzy set and its complement does not result in the empty set (resp. in the universe).

$$A \cap \neg A = \{(a, 0.4), (b, 0.2), (c, 0.0), (d, 0.2), (e, 0.2)\} \neq \varnothing$$

$$A \cup \neg A = \{(a, 0.6), (b, 0.8), (c, 1.0), (d, 0.8), (e, 0.8)\} \neq U$$

More generally, the family of operators implementing the intersection is called *Triangular Norm* (abbreviated t-norm) and *Triangular Conorm* (abbreviated t-conorm or s-norm) for the union. These families comply with the properties showed in Table 2.1 [92].

As noted in Example 2.5, the t-norm and t-conorm operators do not comply with the Aristotle's laws of non-contradiction[2] and excluded middle[3] in order to express the vagueness of the human thinking [44]. The t-norm and t-conorm families are related by a general relation expressed in Definition 2.18 [1].

Definition 2.18 Let A and B be two fuzzy sets over the universe U, t a t-norm operator and s a t-conorm operator, then t and s are connected by the relation:

$$\mu_A(x)\ t\ \mu_B(x) = 1 - ((1 - \mu_A(x))\ s\ (1 - \mu_B(x))),\ \ x \in U$$

The t-norms and t-conorms are also known as *non compensatory* operators meaning that there is no compensation effect between the elements. For instance, the results of a t-norm (resp. t-conorm) operator has an upper (resp. lower) limit defined by the minimum (resp. maximum) operator (see Fig. 2.9). The notion of compensation has an important significance for human beings who instinctively weigh up elements, especially in the context of decision making. For this reason, there exist so-called *averaging* and *compensatory* operators which do not comply with all the properties of t-norms and t-conorms [29].

Table 2.1 Properties of t-norm and t-conorm operators

Property	T-norm	T-conorm
Identity	$1 \wedge x = x$	$0 \vee x = x$
Commutativity	$x \wedge y = y \wedge x$	$x \vee y = y \vee x$
Associativity	$x \wedge (y \wedge z) = (x \wedge y) \wedge z$	$x \vee (y \vee z) = (x \vee y) \vee z$
Monotonicity	if $v \leq w$ and $x \leq y$ then $v \wedge x \leq w \wedge y$	$v \vee x \leq w \vee y$

[2] The law of non-contradiction states that the same thing cannot at the same time belong and not belong to the same object and in the same respect [2].

[3] The law of excluded middle states that of any subject, one thing must be either asserted or denied [2].

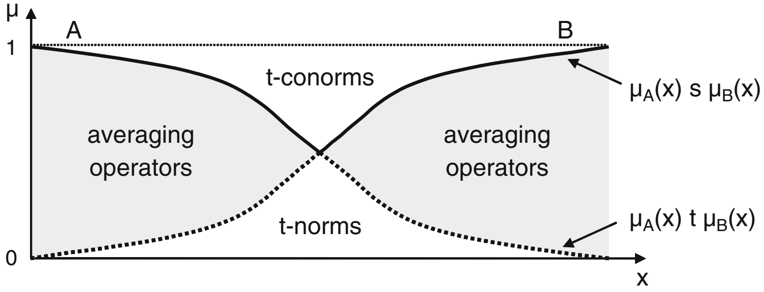


Fig. 2.9 t-norms, t-conorms and averaging operators

Averaging operators for intersection (resp. union) are more optimistic (resp. pessimistic) than t-norms (resp. t-conorms); their results are bounded by the chosen t-norm t and t-conorm s operators as shown in Fig. 2.9. In other words, "averaging operators realize the idea of trade-offs between conflicting goals when compensation is allowed" [139]. These operators have been empirically tested and proved to be well suited for modeling aggregation in a human decision environment [118].

Examples of averaging operators are the 'fuzzy and' and 'fuzzy or' proposed by Werners [122]. They are a combination of the minimum or maximum operator and the arithmetical mean weighted by a γ-argument. For $\gamma = 1$, the results equal the minimum, respectively the maximum, operator. For $\gamma = 0$, the results equal the arithmetical mean for both the 'fuzzy and' and the 'fuzzy or'.

Definition 2.19 The *fuzzy and* of two fuzzy sets A and B defined over a reference set U is defined as

$$\mu_{A \cap B}(x) = \gamma\ min(\mu_A(x), \mu_B(x)) + \frac{(1-\gamma)}{2}\ (\mu_A(x) + \mu_B(x))$$

where $\gamma \in [0, 1]$ and $x \in U$.

Definition 2.20 The *fuzzy or* of two fuzzy sets A and B defined over a universe U is defined as

$$\mu_{A \cup B}(x) = \gamma\ max(\mu_A(x), \mu_B(x)) + \frac{(1-\gamma)}{2}\ (\mu_A(x) + \mu_B(x))$$

where $\gamma \in [0, 1]$ and $x \in U$.

More interesting are the compensatory operators which have a compensation mechanism to reflect the human reasoning. Compensatory operators are located somewhere in between the intersection and union operators. An important compensatory operator is the γ-operator (also called 'compensatory and') which has been suggested as 'compensatory' and empirically tested by Zimmermann and Zysno [140]. It is composed by the algebraic product operator, a t-norm, and its counterpart the algebraic sum, a t-conorm following Definition 2.18 (see Appendix A). This operator has a γ-argument ranging from 0 to 1 which specifies whether the results should go in the direction of the algebraic product (t-norm) or the algebraic sum (t-conorm). The γ-argument therefore determines the strength the compensation mechanism.

Definition 2.21 The *γ-operator* of m fuzzy sets $A_1, \ldots, A_m$ defined over a reference set U with membership functions $\mu_1, \ldots, \mu_m$ is defined as

$$\mu_{A_i,comp}(x) = \left(\prod_{i=1}^{m} \mu_i(x)\right)^{(1-\gamma)} \left(1 - \prod_{i=1}^{m}(1 - \mu_i(x))\right)^{\gamma}, \ \gamma \in [0, 1] \text{ and } x \in U$$

Another interesting kind of operators are the *linguistic modifiers* (also called linguistic hedges) which expresses linguistic notions like ‘a little’, ‘slightly’, ‘very’, ‘extremely’, ‘more or less’, etc. [132]. These operators are called fuzzy sets modifiers as they slightly modify the shape of a membership function according to the expressed notion. For instance, if a fuzzy set expresses the notion of ‘young’, a new fuzzy set with semantics ‘very young’ can be created by applying the linguistic modifier ‘very’ (see Fig. 2.10).

Once again, several operators can be used for the different linguistic modifiers. The most common operator for the modifier ‘very’ is the concentration which is the square value of the original membership degrees. Other linguistic modifiers can be found in Appendix A.

Definition 2.22 The *concentration* of a fuzzy set A defined on a universe U is defined as

$$\mu_{CON(A)}(x) = \mu_A^2(x), \ \ x \in U$$

There exist many other concepts based on fuzzy sets which cannot be presented here, namely the notion of distance between fuzzy sets [96], the notion of necessity and compatibility measures [137], the notion of fuzzy numbers, relations and similarities [131], the extension principle [134, 135, 136], etc. A general introduction to these concepts can be found in [41, 88, 139]. Note also that a detailed discussion of the properties of the operators listed in Appendix A can be found in [31, 66, 68, 122, 129, 139, 140].

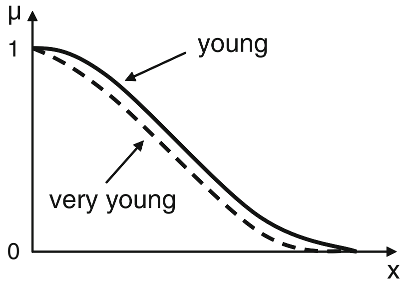


Fig. 2.10 Fuzzy set with the linguistic modifier ‘very’

2.5 Application Fields

The fuzzy set theory has been successfully applied in various domains. The most important application areas are the fuzzy control, the fuzzy diagnosis, the fuzzy data analysis and the fuzzy classification [85]. This section aims to explicit the implication of the fuzzy set theory in some of these domains. First, Sect. 2.5.1 introduces the possibility theory as a basis for the approximate reasoning which allows the integration of the natural language into the reasoning process. Based on the approximate reasoning, Sect. 2.5.2 presents the fuzzy control theory in comparison to the modern control theory. Examples of fuzzy diagnosis and fuzzy data analysis areas are fuzzy expert systems depicted in Sect. 2.5.3 and the fuzzy classification approach presented in Sect. 2.5.4. Last but not least, Sect. 2.5.5 presents different approaches which enable the representation and the storage of the imprecision, i.e. fuzzy databases systems.

Nowadays a large number of real-world applications take advantages of the approximate reasoning [26]. Many other applications fields could have been discussed like the neural networks, the genetic algorithms, the evolutionary programming, the chaos theory, etc., but their presentation is beyond the scope of this thesis.

2.5.1 Possibility Theory

A membership value of a fuzzy set has been defined as the degree to which an element belongs to this fuzzy set. It is possible to give other interpretations to the membership degree like a certainty factor, a degree of truth, a degree of satisfaction and a degree of possibility [88]. In 1978 Zadeh extended the fuzzy set theory to a possibility theory where the membership values are considered as degrees of possibility. Zadeh justifies the possibility theory by the fact that "the imprecision that is intrinsic in natural languages is, in the main, possibilistic rather than probabilistic in nature" [137]. In contrast to the statistical perspective of the information which is involved in the coding, the transmission and the reception of the data [110], the theory of possibility focuses on the meaning of the information.

One central concept in the possibility theory is the possibility distribution which is the counterpart of the probability distribution in the probability theory. A possibility distribution is a fuzzy set called fuzzy restriction, which acts as an elastic constraint, whose membership function determines the compatibility or the possibility with the concept of the fuzzy set. Given a possibility distribution it is possible to compute the possibility of another fuzzy set defined on the same universe (see Definition 2.17). Consider for instance the possibility distribution 'young' of a linguistic variable 'age' defined on the universe U and the fuzzy set 'around 35' also defined on U. By knowing that 'Mary is young' it is then possible to calculate the possibility that 'Mary is around 35'. Note that the possibility represents a degree of feasibility whereas the probability is related to a degree of likelihood implying that what is possible might not be probable and, conversely, what is improbable might not be impossible [69].

The possibility theory opens the door to the fuzzy reasoning which can represent and manipulate the natural language. Almost all human related problems are so complex and so vague that only approximate linguistic expression can be used [69]. The fuzzy approximate reasoning is based on different fuzzy inference patterns which deal with different implication interpretations and also determine the way the uncertainties are propagated. The fuzzy inference can then compute or deduct elastic constraints (fuzzy sets) determined by membership functions via the possibility concept.

2.5.2 *Fuzzy Control Theory*

One of the reasons the scientific community took an interest in the fuzzy logic theory is the financial success of fuzzy control in home appliances in the Japanese industry. In 1990, the consumer products market using fuzzy controllers was estimated to 2 billion dollars [43]. Interestingly enough L.A. Zadeh is a major contributor of the modern control theory. The control theory is a very precise and strict approach in order to model systems or phenomena. As all the aspects of the model have to be specified, modeling a complicated system is an extensive operation. For example, an application could be used to predict the path of a hurricane but if it has to be developed from scratch, the hurricane will be gone before the application is ready to use. In the control theory, the number of processes to be implemented grows exponentially relatively to the number of variables defining the system [88]. For this reason some systems cannot be modeled even by high speed computers. A solution to this problematic is to roughly define systems with the help of the fuzzy logic theory. The fuzzy control is based on the approximate reasoning which offers a more realistic framework for human reasoning than the two-valued logic [134, 135, 136]. The main advantages of fuzzy control over the classical control theory is its ability of implementing human expert knowledge, its methods for modeling non-linear systems and a shorter time to market development [29, 43].

In the control theory systems are characterized by input and output variables as well as a set of rules. These rules define the behavior of the system. The output variables are then calculated by inference based on the input variables and the given rules. An inference is the construct 'A implies B, B implies C then A implies C'. When a premise 'X implies Y' holds then Y is true if X is true and, conversely, X is false if Y is false. This is called a syllogism and a famous example is:

Implication:	All men are mortal
Premise:	Socrates is a man
Conclusion:	Socrates is mortal

In the control theory the premises are defined by rules in the form 'If X is F Then Y is G' where X (resp. Y) is an input (resp. output) variable and F (resp. G) is a condition on X (resp. Y). Zadeh introduced in 1973 the compositional rule of inference [133] which extends the inference mechanism in order to take the fuzziness into account. In 1993, Fullér and Zimmermann demonstrated the stability property of the conclusion using the compositional rule of inference which states that a conclusion depends continuously on the premise when the t-norm defining the composition and the membership function of the premise are continuous [38]. This property guarantees that small changes in the membership function of the premise, eventually due to errors, can imply only a small deviation in the conclusion.

The fuzzy rules can then be expressed in the natural language by the use of linguistic variables [134, 135, 136]. Zadeh's fuzzy inference example where the conditions are expressed by the means of words is:

Implication:	If a tomato is red then it is ripe
Premise:	This tomato is very red
Conclusion:	This tomato is very ripe

These words allow the fuzzy rules to integrate the semantics of the human knowledge and can be represented as fuzzy sets. The evaluation process of the fuzzy inference also differs from the classical control theory in the sense that all the rules involving a given output variable are computed simultaneously and their results are then merged in order to derive the value of the output variable. This is a major advantage over the classical control theory as it implies a compensation mechanism between the involved rules. As a result, a much smaller set of rules is required to model a system as the intermediate values of the input variables are dynamically interpolated from the existing rules. It also implies an inherent fault tolerance; consider that a rule has been erroneously implemented or that a hardware defect returns wrong results, the value of the output variable can be compensated by other rules defining this variable.

Many concrete applications using fuzzy control can be found. The most famous one is the opening in 1988 of a subway system in Sendai City (Japan) using the fuzzy control to accelerate and brake the trains more smoothly than a human driver. Compared to conventional control, this new approach achieved significant improvements in the fields of safety, riding comfort, accuracy of stop gap, running time and energy consumption [130]. Other concrete applications can be found in domestic appliances like washing machines and vacuum cleaners, in visual systems like camera auto focus and photocopiers, in embedded car systems like anti-lock braking systems, transmission systems, cruise control and air conditioning, etc. A review of fuzzy control applications can be found in [29].

2.5.3 Fuzzy Expert Systems

Expert systems are a successful example from the broad field of artificial intelligence. Expert systems are knowledge-based systems which can derive decision or conclusion based on an extensive knowledge on a particular domain. More precisely, "an expert system is a program that can provide expertise for solving problems in a defined application area in the way the experts do" [65]. This knowledge is represented in a set of 'If-Then' rules. By applying inferences on the specified rules, expert systems are able to derive optimal decisions.

A major problematic, however, is to convert the experts' knowledge into a set of 'If-Then' rules which are exact given that the human representation of the knowledge cannot be sharply determined. This drawback can be overcome by introducing the fuzziness. This is done by allowing the definition of fuzzy rules, i.e. rules with words determined by a membership function, and by applying the previously defined fuzzy inference. Just like in the fuzzy control, the fuzzy inference allows a dynamic compensation between the different fuzzy rules which results in the definition of a smaller set of rules. Fuzzy expert systems are usually involved when processes cannot be described by exact algorithms or when these processes are difficult to model with conventional mathematical models [49].

Although the rules definition and the inference mechanism of fuzzy expert systems are similar to those in fuzzy control, fuzzy expert systems do not come under the category of fuzzy control [85]. Fuzzy control applications (often called fuzzy controllers) work in a closed loop schema where the output variables, which are derived from the input variables, directly act on the considered object. The rules are then executed in cycles in order to maintain a system. In the case of fuzzy expert systems and, more generally, for fuzzy diagnosis, fuzzy data analysis and fuzzy classification systems, the output information of a fuzzy system is dedicated to a human user or a monitoring device and hasn't any impact on the object itself.

Earl Cox [26] has implemented different fuzzy expert systems which have been successfully applied to the following domains: transportation, managed health care, financial services, insurance risk assessment, database information mining, company stability analysis, multi-resource and multi-project management, fraud detection, acquisition suitability studies, new product marketing and sales analysis. By comparing fuzzy expert systems with conventional expert systems Cox stated that "generally, the final models were less complex, smaller, and easier to build, implement, maintain, and extend than similar systems built using conventional symbolic expert systems" [26].

2.5.4 Fuzzy Classification

The fuzzy classification is a natural extension of the traditional classification, the same way that the fuzzy sets extend the classical sets (see Sect. 2.2). In a sharp classification, each object is assigned to exactly one class, meaning that the membership degree of the object is 1 in this class and 0 in all the others. The belonging of the objects in the classes is therefore mutually exclusive. In contrast, a fuzzy classification allows the objects to belong to several classes at the same time; furthermore, each object has membership degrees which express to what extent this object belongs to the different classes.

Definition 2.23 Let O be an object characterized by a t-dimensional feature vector $\underline{x}_O$ of a universe of discourse U. Often U is the space R^t. Let $C_1, \ldots, C_n$ be a set of classes which is given a priori or has to be discovered. A *fuzzy classification* calculates a membership vector $M = \{m_1, \ldots, m_n\}$ for the object O. The vector element $m_i \in [0, 1]$ is the degree of membership of O in the class C_i [85].

In many real applications, a dichotomous assignment of an object in one class is often not possible as no unique conclusion can be derived from the object features and/or the object features cannot be exactly observed [85]. This is particularly true for problems related to the human evaluation, intuition, perception and decision making where the problem structure is not dichotomous [139]. The definition of the classes can be determined by using the knowledge of experts of the domain or can be automatically found by the use of data mining techniques like cluster analysis [85].

The fuzzy classification approach can be used for instance for diagnosis and for decision making support. In the case of a diagnosis system for ill persons, the classification procedure can derive the illness based on the symptoms of the patient or find a suitable therapy considering the illness of the patient [85, 88]. In a decision making process, the classification (also called segmentation depending on the context) is used to derive management decisions based on several characteristics of the objects. A major issue in this field is the complexity of the data, i.e. the abundance of information. This complexity is a source of uncertainty due to the limited capability of human beings to observe and handle large amounts of data simultaneously [91]. As in the management field a large number of objects described by many features is usually considered, the classification approach, by grouping similar objects into classes, results in a complexity reduction which enables a better situation analysis [102]. Furthermore, the fuzzy classification, in contrast to the classical one, by allowing objects to belong to several classes at the same time, reduces the complexity of the data and also provides a much more precise information about the classified elements.

The fuzzy classification concept given in Definition 2.23 specifies that a membership vector for each object in the different classes has to be calculated. There exist many approaches to derive the membership vector of the classified objects. In this thesis, a context based approach which extends the relational database schema to a fuzzy relational database schema has been chosen. Chapter 3 presents this approach which has been originally proposed by Schindler [102].

2.5.5 Fuzzy Database Systems & ER Models

Many contributions in the fields of database systems which study the representation and the processing of imprecise information can be found in the literature. In this context, the imprecision can be given several interpretations [41]: an information is *uncertain* if it is incomplete or fuzzy, *unknown* if no information is available or *undefined* if it is inapplicable to the predefined domain. Note that this definition of imprecision in data management overlaps the general aspects of ambiguity discussed in Sect. 2.1.

Different approaches besides fuzzy logic can be used to model the imprecision. In the field of relational databases, Codd first introduced the notion of *NULL* values which allows the modeling of unknown information [23], then extended it to distinguish between missing values which may be applicable (i.e. unknown) called A-mark or inapplicable (i.e. undefined) called I-mark [24, 25]. In order to encompass the uncertainty of an information, Grant proposed an extension of the relational model to accept not only null values but also interval values as tuple components [45].

Another approach to model the uncertainty of information is the development of statistical and probabilistic databases. For instance, Wong proposed a framework allowing the extraction of information from imprecise databases by statistical means [128]. Barbará et al. extended the relational model by a probabilistic model which allows probabilities to be associated with attributes values. This probabilistic data model also includes missing probabilities and can therefore capture uncertainty in data values [3]. Similarly, Cavallo and Pittarelli proposed a theory of probabilistic databases which is a generalization of the relational model to accommodate both probabilistic and relational data. By assigning all tuples of a relation a probability distribution, it is possible to know the probability that a tuple belongs to the relation [18].

In order to better model the imprecision in databases, the use of the fuzzy logic theory can be considered. For this purpose, two different architectural concepts can be differentiated, depending at which level the fuzziness is applied [87]. A first category of systems, the fuzzy database systems, directly integrates the fuzziness at the data level. In contrast, a second category of systems still works with relational database systems and specifies an additional upper layer where fuzzy queries can be formulated on top of the RDBMS. In this subsection fuzzy database systems are considered since fuzzy query languages for relational databases are treated in Chap. 3.

Following Galindo et al. [41], the basic fuzzy relational database model is an extension of the relational model by adding to each tuple a grade in the interval [0, 1]. This extension is very similar to the one of the probabilistic databases previously mentioned, however both models cannot be compared. As discussed in Sect. 2.5.1, the probability focuses on the likelihood of events and has specific constraints, e.g. the probability distribution has to sum up to 1, which is not the case in fuzzy databases where the grades can be assigned several meanings. The tuples' grade and their associated meaning can then be used in the querying process. The grades may represent the membership degree of the tuples in the relation, the dependence strength between two attributes as well as the fulfillment of a condition or the importance degree of the tuples of a relation [41]. This basic fuzzy relational model suffers however from not being able to represent imprecise information at the attribute level since grades are assigned to the tuples directly [41].

In 1986 Zvieli and Chen [143] proposed an extension of the entity-relationship model (ER) [21] by allowing the definition of fuzzy attributes, entities and relationships. In this approach, three levels of fuzziness in the ER model can be distinguished:

- On a first level, entities, relationships and attributes include a membership degree to the model. This degree expresses the importance of the element to the model such that given an importance threshold, some elements can be eliminated.
- On a second level, instances of entities and relationships can be fuzzy. In this case, a degree (whose meaning has to be specified) depicts the belonging of a tuple (resp. a relationship instance) in the entity (resp. the relationship).
- On a third level, entity and relationship attributes might contain fuzzy values. This allows the imprecision to be integrated in the attributes values.

In order to explicit the potential of fuzzy database systems, the Fuzzy Enhanced Entity-Relationship model (FuzzyEER) proposed by Elmasri and Navathe [34] is shortly summarized. FuzzyEER is an extension of the Enhanced Entity-Relationship model (EER) with fuzzy semantics and notations [41]. This model firstly allows the representation of fuzzy values by means of fuzzy attributes and fuzzy degrees. Fuzzy attributes, besides the support of unknown and undefined values, enable the representation of uncertain values by the use of possibility distributions and/or similarity relations over the domain of the attribute, e.g. the unknown age of a person can be characterized by 'about 35 years old'. Fuzzy degrees, in contrast, specify an additional grade for one or several attributes (fuzzy or not) whose meaning can be a fulfillment degree, an uncertainty degree, a possibility degree, an importance degree or a membership degree. Extreme cases of fuzzy degrees occur when a grade isn't assigned to any attribute and when a grade is assigned to the whole tuple, i.e. defining a fuzzy entity, just like in the basic fuzzy relational database model. In a similar manner, an entity, a relationship or an attribute can be associated with a fuzzy degree in respect to the model. The implementation of an importance degree in a model allows, for instance, a user to view the most important part of the model according to a given importance threshold. Other constructs of the FuzzyEER model are the fuzzy weak entities, the fuzzy aggregations, the fuzzy specializations and the fuzzy constraints.

Despite the flexibility offered by fuzzy databases systems, the fuzzy classification approach proposed in this thesis does not rely on such systems. As Chap. 3 relates, information systems are mostly stored in relational databases [7] and the use of fuzzy database systems would imply the migration of the existing data. Furthermore, as the management perspectives are changing over time, the fuzzy information contained in the fuzzy databases would have to be regularly adapted. For these reasons, the chosen fuzzy classification approach works with relational databases and implements a fuzzy classification database schema (see Chap. 3). The extension of the relational database schema, which consists of meta-tables added to the system catalog, enables the definition of fuzzy classifications (see Chap. 7). Since the meta-tables are independent from the business data, changes in the fuzzy classifications' definition do not impact the original data collections.

http://www.springer.com/978-3-319-15969-0

Fuzzy Classification of Online Customers
Werro, N.
2015, XIII, 141 p. 65 illus., Hardcover
ISBN: 978-3-319-15969-0