Chapter 2
Type-1 Fuzzy Sets and Fuzzy Logic

2.1 Crisp Sets

Recall that a *set* A in a universe of discourse X (which provides the set of allowable values for a variable) can be defined by listing all of its members or by identifying the elements $x \subset A$. One way to do the latter is to specify a condition or conditions for which $x \subset A$; thus, A can be defined as $A = \{x|x \text{ meets some condition(s)}\}$. Alternatively, one can introduce a zero-one *membership function* (MF) (also called a characteristic function, discrimination function, or indicator function) for A, denoted $\mu_A(x)$, such that

$$A \Rightarrow \mu_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} \tag{2.1}$$

Set A (which can also be treated as a subset of X) is mathematically equivalent to its MF $\mu_A(x)$ in the sense that knowing $\mu_A(x)$ is the same as knowing A itself. In order to distinguish between a set and a fuzzy set, the former will be referred to as a "crisp set."

Example 2.1 (Mendel 1995a) Consider the set of all automobiles in New York City; this is X. The elements of X are individual cars; but, there are many different types of subsets that can be established for X, including the three that are depicted in Fig. 2.1. Either a car has or does not have six cylinders. This is a very crisp requirement. Hence, if your car has four cylinders, its MF value (i.e., membership grade) for the subset of four cylinder cars is unity, whereas its membership grades for the subsets of six cylinder or eight cylinder cars are zero.

Example 2.2 (Mendel 2015) Suppose that the domain of x is partitioned into five regions, and one knows exactly where the dividing line is between each region,

J.M. Mendel, *Uncertain Rule-Based Fuzzy Systems*,
DOI 10.1007/978-3-319-51370-6_2

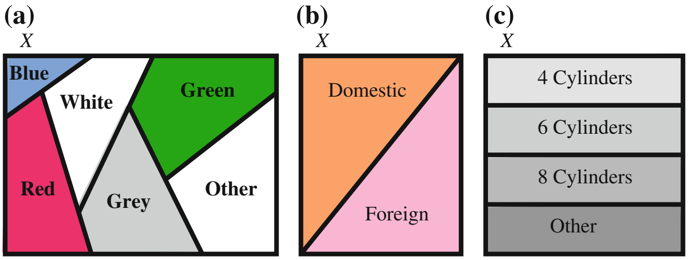


Fig. 2.1 Partitioning of the set of all automobiles in New York City into subsets by **a** color, **b** domestic or foreign, and **c** number of cylinders (Mendel 1995a © 1995, IEEE)

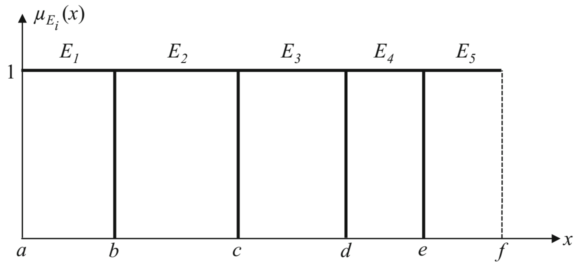


Fig. 2.2 Interpreting crisp sets as crisp partitions (Mendel 2015 © Springer 2015)

so one is in the situation that is depicted in Fig. 2.2, where no uncertainty exists about $x = b, c, d, e$. Each of the intervals $[a, b]$, $(b, c]$, $(c, d]$, $(d, e]$, $(e, f]$ is a *crisp partition* (Definition 1.1), i.e. x is either in it (with membership value of 1) or not in it (with membership value of 0), and x cannot simultaneously be in more than one of these intervals. Each interval is associated with a crisp set that is described by a linguistic term, E_1, or E_2, or $\ldots$, or E_5, such as a level of temperature or pressure, and there is always a sharp jump from one set to another at $x = b, c, d, e$. As mentioned in connection with Fig. 1.1a, this crisp model serves us well in many situations, but it does not allow any uncertainty about $x = b, c, d, e$. A fuzzy set will allow for this, as shall be seen.

2.2 Type-1 Fuzzy Sets and Associated Concepts

This section provides the background that is needed to read Chaps. 3 and 4. To begin, a short section about the father of fuzzy sets and logic, Professor Lotfi A. Zadeh, is provided.

2.2.1 *Lotfi A. Zadeh*

Fuzzy sets[1] were invented around 1965 by Prof. Lotfi A. Zadeh, but why? In Zadeh (1973), he states:

> Essentially our contention is that the conventional quantitative techniques of system analysis are intrinsically unsuited for dealing with humanistic systems or, for that matter, any system whose complexity is comparable to that of humanistic systems. The basis for this contention rests on what might be called the *principle of incompatibility*. Stated informally, the essence of this principle is that as the complexity of a system increases, our ability to make precise and yet significant statements about its behavior diminishes until a threshold is reached beyond which precision and significance (or relevance) become almost mutually exclusive characteristics (a corollary to this principle may be stated succinctly as, "The closer one looks at a real-world problem, the fuzzier becomes its solution."). It is in this sense that precise quantitative analyses of the behavior of humanistic systems are not likely to have much relevance to the real world societal, political, economic, and other types of problems which involve humans either as individuals or in groups.

Prof. Zadeh[2] (Fig. 2.3), born in Baku, Azerbaijan on February 4, 1921, and educated at Alborz College in Tehran, the University of Tehran, M.I.T. and Columbia University, spent most of his career at the University of California at Berkeley, after ten years at Columbia University. He was already a famous system theorist when in 1965 he published what has now become the seminal paper on fuzzy sets (Zadeh 1965). This paper, which, as of February 2017, has been cited in Google Scholar more than 69,500 times, and is the most highly cited paper in all of computer science, marked the beginning of a new direction; by introducing the concept of a fuzzy set, that is a class with un-sharp boundaries, he provided a basis for a qualitative approach to the analysis of complex systems in which linguistic rather than numerical variables are employed to describe system behavior and performance. In this way, a much better understanding of how to deal with uncertainty may be achieved, and better models of human reasoning may be constructed. Although his unorthodox ideas were initially met with some skepticism, they have gained wide acceptance in recent years and have found application in just about every field imaginable. He is now acknowledged to be the "Father of Fuzzy Sets and Fuzzy Logic."

[1]The English word "fuzzy" has a negative connotation when it used in a technical context. It may be okay to describe a soft teddy bear, a cuddly pet, or a peach but for it to be used for mathematics and its applications is a red flag. Prof. Zadeh was well aware of this but felt that in 1965 "fuzzy" was the best word for him to use for this kind of a set. I propose that, after more than 50 years, these sets be called *Zadehian* sets. I am not going to use my proposed replacement in this book, because, although I would like to do it, if I did almost no one would know what I was talking about.

[2]This short biographical sketch was taken mostly from Mendel (2007).

Fig. 2.3 Professsor Lotfi A. Zadeh, the Father of Fuzzy Sets and Fuzzy Logic. Photo taken at Mendel Symposium, at University of Southern California, May 2009

2.2.2 Type-1 Fuzzy Set Defined

Definition 2.1 A *type-1 fuzzy set*[3] A is (Aisbett et al. 2010) a set function on universe X (sometimes denoted D_A) into $[0, 1]$, possibly constrained to belong to a family such as continuous functions, i.e. $\mu_A{:}X \to [0, 1]$. The MF of A is denoted $\mu_A(x)$ and is called a *type-1 MF*, i.e.

$$A = \{(x, \mu_A(x)) | x \in X\} \tag{2.2}$$

in which $0 \le \mu_A(x) \le 1$. A can also be expressed in fuzzy set notation[4] for continuous universes X, as

[3]In order to distinguish among different fuzzy set models, what were originally called *fuzzy sets* are in this book called *type-1 fuzzy sets*. Beginning with Chap. 6, type-2 fuzzy sets are studied.

[4]*Fuzzy set notation* was introduced in Zadeh (1965) and has remained popular for more than 50 years, although many people find it somewhat strange and object to its use of symbols such as the integral and summation. Aisbett et al. (2010) distinguish between "fuzzy set notation" and "standard mathematical notation." In Definition 2.1, $\mu_A{:}X \to [0, 1]$ is the description of a type-1 fuzzy set in standard mathematical notation. My own preference is to use each notation where it is useful.

$$A = \int_{x \in X} \mu_A(x)/x \tag{2.3}$$

where $\int$ denotes union over all $x \in X$, or for discrete universes X_d, as

$$A = \sum_{x \in X_d} \mu_A(x)/x \tag{2.4}$$

where $\sum$ denotes union over all $x \in X_d$. The slash in (2.3) and (2.4) associates the elements in X with their membership grades, where $\mu_A(x) > 0$. The value of $\mu_A(x)$ is called the *degree of membership*, or *membership grade*, of x in A. If $\mu_A(x) = 1$ or $\mu_A(x) = 0$ for all $x \in X$, then the fuzzy set A reduces to a crisp set.

$\mu_A(x)$ is also said to provide a *measure of the degree of similarity* of an element in X to the fuzzy set. Note that A can also be treated as a subset of X. Unlike a crisp set, that can be described in different ways (as is explained in Sect. 2.1), a fuzzy set can only be described by its MF.

Example 2.1 (Continued) (Mendel 1995a) Referring to the middle of Fig. 2.1, observe that cars can also be partitioned into the two subsets, domestic and foreign. But, a car can be viewed as "domestic" or "foreign" from different perspectives. One perspective is that a car is domestic if it carries the name of a U.S. auto manufacturer; otherwise it is foreign. There is nothing fuzzy about this perspective. Many people today, however, feel that the distinction between a domestic and foreign automobile is not as crisp as it once was, because many of the components for what one considers to be domestic cars (e.g., Fords, GMs, and Chryslers) are produced outside of the United States. Additionally, some "foreign" cars are manufactured in the United States. Consequently, one could think of the MFs for domestic and foreign cars looking like $\mu_D(x)$ and $\mu_F(x)$ depicted in Fig. 2.4. Observe that a specific car (located along the horizontal axis by determining the percentage of its parts made in the United States) exists in both subsets simultaneously—domestic cars and foreign cars—but to different degrees of membership. For example, if a car has 75% of its parts made in the United States, then[5] $\mu_D(75\%) = 0.90$ and $\mu_F(75\%) = 0.25$. Ultimately, one would describe such a car as domestic. In fact, when one does this, the subset is decided upon by choosing it to be associated with the maximum of $\mu_D(75\%) = 0.90$ and $\mu_F(75\%) = 0.25$.

The main point of this example is to demonstrate that in a fuzzy set an element can reside in more than one set to different degrees of similarity. This cannot occur in a crisp set.

---

[5]For fuzzy sets, there is absolutely no requirement that $\mu_D(x) + \mu_F(x) = 1$, even though some authors impose this (e.g., Ruspini 1969; Bezdek 1981). When the constraint that the sum of the fuzzy set memberships must add to 1 for $x \in X$ is imposed, the result is called a *fuzzy partition*. Fuzzy partitions are not used in this book because, in the opinion of this author, they impose unnecessary constraints on fuzzy set MFs, especially when MF parameters are optimized, as is commonly done in rule-based fuzzy systems.

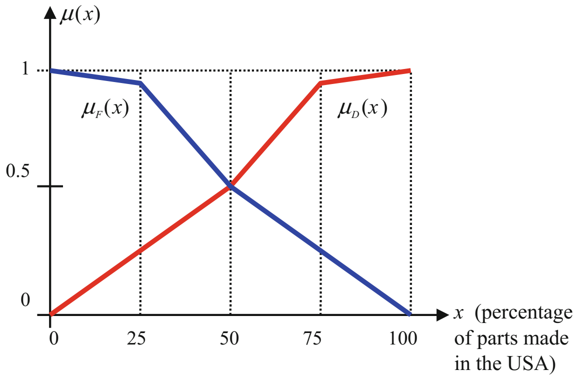


Fig. 2.4 MFs for domestic and foreign cars, based on the percentage of parts in the car made in the United States (Mendel 1995a © 1995, IEEE)

Note that describing a car by its color is also not a crisp description, because each color has different shades associated with it.

Definition 2.2 The *support* of a type-1 fuzzy set A is the crisp set of all points $x \in X$ such that $\mu_A(x) > 0$. A type-1 fuzzy set whose support is a single point in X with $\mu_A(x) = 1$ is called a (type-1) *fuzzy singleton*.

Definition 2.3 The *height* of a type-1 fuzzy set is the maximum MF value. A *normal* type-1 fuzzy set is one for which $\sup_{x \in X} \mu_A(x) = 1$, that is, its height equals 1.

Example 2.2 (Continued) (Mendel 2015) Referring to Fig. 2.2, suppose one now wants a model that allows for uncertainty about $x = b, c, d, e$, so that one is in the situation of Fig. 2.5, where in $[a, b_l]$ x resides only in E_1, whereas in $(b_l, b_r]$ x resides simultaneously in E_1 and E_2, but to different degrees, $\mu_{E_1}(x)$ and $\mu_{E_2}(x)$, respectively; in $(b_r, c_l]$ x resides only in E_2, whereas in $(c_l, c_r]$ x resides simultaneously in E_2 and E_3, but to different degrees, $\mu_{E_2}(x)$ and $\mu_{E_3}(x)$, respectively; etc. The MF $\mu_{E_i}(x)$ for E_i is no longer only 0 or 1, and MFs can overlap. So, a type-1 fuzzy set allows x to be partitioned using *overlapping partitions*, where one is absolutely certain about where the overlap begins and ends, i.e. as *first-order uncertainty partitions* (Definition 1.2), something that cannot be done by a crisp set. Overlapping partitions lead to smooth transitions from one set to another, which is very different from the sharp jumps that occur when crisp sets are used. As mentioned in connection with Fig. 1.1b, this fuzzy set model serves us well in many situations, but it does not allow for any uncertainty about the overlap. A type-2 fuzzy set will allow for this.

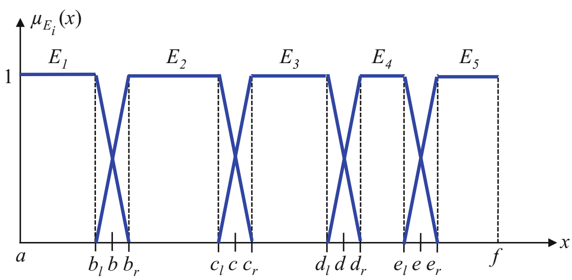


Fig. 2.5 Interpreting type-1 fuzzy sets as overlapping partitions (Mendel 2015 © Springer 2015)

Each of the five fuzzy sets in Fig. 2.5 is a normal type-1 fuzzy and the support of E_1 is $[a, b_r)$, the support of E_2 is $(b_l, c_r), \ldots$, and the support of E_5 is $(e_l, f]$.

Example 2.3 (Zimmerman 1991) Let F = integers close to 10; then, one choice for $\mu_F(x)$ is:

$$\mu_F(x) \equiv 0.1/7 + 0.5/8 + 0.8/9 + 1/10 + 0.8/11 + 0.5/12 + 0.1/13 \tag{2.5}$$

Three points to note from this MF are:

1. The integers for x not explicitly shown all have MFs equal to zero—by convention, such elements are not listed.
2. The values for the MFs were chosen by a specific individual; except for the unity membership value when $x = 10$, they can be modified based on one's own personal interpretation of the word "close," i.e. *words mean different things to different people*.
3. The MF is symmetric about $x = 10$, because there is no reason to believe that integers to the left of 10 are close to 10 in a different way than are integers to the right of 10; but again, other interpretations are possible.
4. F is a normal type-1 fuzzy set.
5. The fuzzy set F is an example of a *type-1 fuzzy number*, which will be defined formally in Sect. 2.2.3 (Definition 2.5).

Definition 2.4 A type-1 fuzzy set A is *convex* (Klir and Yuan 1995) if and only if

$$\mu_A(\lambda x_1 + (1 - \lambda)x_2) \geq \min[\mu_A(x_1), \mu_A(x_2)] \tag{2.6}$$

This can be interpreted as (Lin and Lee 1996): Take any two elements x_1 and x_2 in fuzzy set A; then the membership grade of all points between x_1 and x_2 must be greater than or equal to the minimum of $\mu_A(x_1)$ and $\mu_A(x_2)$. This will always occur when the MF of A is first monotonically non-decreasing and then monotonically non-increasing.[6] The five MFs in Fig. 2.5 are convex, whereas the two MFs in Fig. 2.4 are not.

Example 2.4 The MF of a convex type-1 fuzzy set A often satisfies the following structure:

$$\mu_A(x) = \begin{cases} g(x)|_{g(a)=0, g(b)=1}, & x \in [a, b] \\ 1, & x \in [b, c] \\ h(x)|_{h(c)=1, h(d)=0}, & x \in [c, d) \end{cases} \tag{2.7}$$

[6]In mathematics a real-value function $f(x)$ defined on an interval is called *convex* if the line segment between any two points on the graph of the function lies above or on the graph (e.g., a parabola). Why the fuzzy set A that satisfies (2.6) is called "convex" rather than "concave" is a bit mysterious. Maybe it is due to a concave function also being known in mathematics as a convex upwards, convex cap, or upper convex function.

where $g(x) \in [0,1]$ is monotonically non-decreasing and $h(x) \in [0,1]$ is monotonically non-increasing.

In rule-based applications of fuzzy logic, the MFs $\mu_A(x)$ are associated with linguistic terms that appear in the antecedents or consequents of rules, or in phrases (e.g., *foreign* cars).

Example 2.5 Some examples of rules and associated MFs (shown in brackets) are: (1) IF one is tracking a *large* target at one instant of time, THEN the target will not be *too far away* at the next instant of time $[\mu_{LARGE}(t), \mu_{TOO\text{-}FAR\text{-}AWAY}(x)]$; (2) IF the horizontal position is *medium positive* and the angular position is *small negative*, THEN the control angle is *large positive* $[\mu_{MEDIUM\text{-}POSITIVE}(x),\ \mu_{SMALL\text{-}NEGATIVE}(\theta),\ \mu_{LARGE\text{-}POSITIVE}(\phi)]$ and, (3) IF $y(t)$ is *close to 0.5*, THEN $f(y)$ is *close to zero* $[\mu_{CLOSE\text{-}TO\text{-}0.5}(y),\ \mu_{CLOSE\text{-}TO\text{-}ZERO}(f(y))]$.

The most commonly used shapes for MFs are triangular, trapezoidal, piecewise linear, Gaussian, and bell-shaped. MFs can either be chosen by the user arbitrarily, based on the user's experience (hence, the MFs for two users could be quite different depending upon their experiences, perspectives, cultures, etc.), or, they can be designed using optimization procedures, e.g., Horikawa et al. (1992), Jang (1992), Wang and Mendel (1992a, b).

The number of MFs is free to be chosen. Greater resolution is achieved by using more MFs at the price of greater computational complexity. MFs don't have to overlap; but one of the great strengths of fuzzy logic is that MFs can be made to overlap. This expresses the fact that "the glass can be partially full and partially empty at the same time." In this way (as will become clear in later chapters, e.g., Chap. 3) one is able to distribute decisions over more than one input class, which helps to make fuzzy logic systems robust.

The MF of a type-1 fuzzy set is specified exactly, which seems counter-intuitive for something that is supposed to be "fuzzy." This was one of the very early criticisms of a fuzzy set and is something that shall be returned to in Chap. 6 when type-2 fuzzy sets are studied.

2.2.3 Type-1 Fuzzy Numbers

When there is some uncertainty about a number (due, e.g., to measurement errors or linguistic uncertainty about it) it can be modeled as a fuzzy set, in which case it is called a *fuzzy number*. When the uncertainty is modeled using a type-1 fuzzy set it is called a *type-1 fuzzy number*. These numbers can be defined in different ways (e.g., Dubois and Prade 1980; Jang and Ralescu 2001; Klir and Yuan 1995; Wang 1997).

Definition 2.5 Let A be a fuzzy set in R. A is called a *type-1 fuzzy number* if: (i) A is normal, (ii) A is convex, and (iii) A has a bounded support.

It is tempting to do away with the requirement that A has a bounded support, but to do so makes no physical sense, since uncertainty about a real number should be finite. Regardless, it is not uncommon for an uncertain number to be modeled as a Gaussian type-1 fuzzy set that is centered about that number, and for this to be referred to as a "Gaussian fuzzy number." Strictly speaking, this designation is incorrect because the support for such a fuzzy set is unbounded. Occasionally, however, this designation is used even in this book, out of convenience, and because the author feels that its use will not confuse the reader.

Example 2.6 Formulas for triangle and trapezoidal type-1 fuzzy numbers are given in (2.8) and (2.9), respectively (see, also Table 2.3, in which alternate symbols are used for the parameters that define these fuzzy numbers):

$$\mu_A(x) = \mu_A(x; a, b, c) = \begin{cases} (x-a)/(b-a) & \text{if } a \le x < b \\ (c-x)/(c-b) & \text{if } b \le x \le c \\ 0 & \text{if } x > c \text{ or } x < a \end{cases} \tag{2.8}$$

$$\mu_A(x) = \mu_A(x; a, b, c, d) = \begin{cases} (x-a)/(b-a) & \text{if } a \le x < b \\ 1 & \text{if } b \le x \le c \\ (d-x)/(d-c) & \text{if } c < x \le d \\ 0 & \text{if } x > d \text{ or } x < a \end{cases} \tag{2.9}$$

Note, also, that when $b = c$ in (2.7), the resulting fuzzy set is often called an *LR fuzzy number* (Dubois and Prade 1980).

Type-1 fuzzy numbers are sometimes used in a type-1 rule-based fuzzy system during the front-end fuzzification process. More will be said about this in Sect. 2.2.3 (Definition 3.5). The extension of a type-1 fuzzy number to an interval type-2 fuzzy number is described in Sect. 6.5.

Definition 2.6 A *type-1 interval fuzzy number A* is a type-1 fuzzy number for which $\mu_A(x) = 1$, $x \in [l, r]$.

These kinds of type-1 fuzzy numbers play an important role in interval type-2 fuzzy sets and systems.

2.2.4 Linguistic Variables

Zadeh (1975, p. 201) states:

> In retreating from precision in the face of overpowering complexity, it is natural to explore the use of what might be called linguistic variables, that is, variables whose values are not numbers but words or sentences in a natural or artificial language. The motivation for the use of words or sentences rather than numbers is that linguistic characterizations are, in general, less specific than numerical ones.

Definition 2.7 If a variable can take words in natural languages as its values, it is called a *linguistic variable*, where the words are characterized by fuzzy sets defined in the universe of discourse in which the variable is defined (Wang 1997). Each

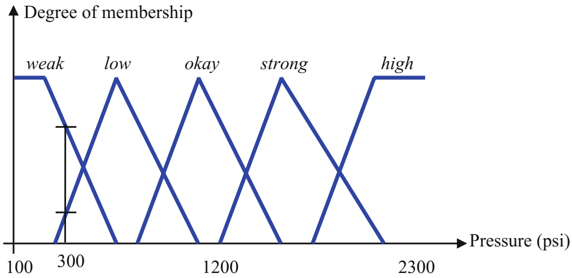


Fig. 2.6 MFs for T (pressure) = {*weak*, *low*, *okay*, *strong*, *high*}. The shapes of the MFs as well as their degree of overlap are quite arbitrary [©1992 IEEE. This figure has been taken from Cox (1992)]

linguistic variable (Klir and Yuan 1995; Zadeh 1973, 1975) is fully characterized by a quintuple (v, T, X, g, m) in which v is the name of the variable, T is the set of linguistic terms[7] of v that refer to a base variable whose values range over the universal set X, g is a syntactic rule for generating linguistic terms, and m is a semantic rule that assigns each linguistic term $t \in T$ its meaning, $m(t)$, which is a fuzzy set on X, that is, $m{:}T \to F(X)$, where $F(X)$ denotes the set of fuzzy sets of X, one fuzzy set for each $t \in T$. It is common to refer to v as the linguistic variable.

Example 2.7 Some examples of linguistic variables, v, are: Pressure, Horsepower, Acceleration, Production Rate, Developed Country, Industrial Country, Profitable Company, Institutional Veto Points, All-day School Systems, etc. Some examples of the set of linguistic terms, T, for these linguistic variables are[8]:

1. For Pressure, T = {*weak, low, okay, strong, high*}
2. For Horsepower, Acceleration and Production Rate, $T \triangleq$ {*very low, low, moderate, high, very high*}
3. For Developed or Industrial (Country) and Profitable (Company), $T \triangleq$ {*barely, hardly, somewhat, moderately, fully, extremely*}
4. For Institutional Veto Points and All-day School Systems, $T \triangleq$ {*none to very few, some, a moderate amount of, many, a large number of, a very large number of*}.

Observe that linguistic terms should make linguistic sense for its linguistic variable, which is where g in Definition 2.7 comes into play; so, for example, *somewhat acceleration* makes no linguistic sense nor does *very high all-day school systems*. Note, also, that *it is the elements of T that are treated as fuzzy sets,* and, of course, each of these fuzzy sets is described by a MF.

Figure 2.6 depicts the MFs for Pressure (Cox 1992) when its universe of discourse is X = [100 psi, 2300 psi]. One might interpret *weak* as a pressure below

[7]Although "term" means one or more words, it is quite common in the fuzzy set literature to see "word" used instead of "term," even when a term includes more than one word. In this book, "term" and "word" are also used interchangeably.

[8]Because some of the linguistic terms may be so similar to each other, it may not be necessary to use all of them. One usually chooses the linguistic terms so that their MFs overlap and cover X.

200 psi, *low* as a pressure close to 700 psi, *okay* as a pressure close to 1050 psi, *strong* as a pressure close to 1500 psi, and *high* as a pressure above 2200 psi. Measured values of pressure (x) lie along the pressure axis, and a vertical line from any value of pressure intersects, at most, two MFs. So, for example, $x = 300$ psi resides in the fuzzy sets *weak pressure* and *low pressure*, but to different degrees of similarity.

Zadeh (1999, p. 107) has used the word *perception* to describe the terms associated with linguistic variables. For example, he states:

> A fundamental difference between measurements and perceptions is that, in general, measurements are crisp numbers whereas perceptions are fuzzy numbers or, more generally, fuzzy granules, that is, clumps of objects in which the transition from membership to non-membership is gradual rather than abrupt.

Indeed, in Example 2.7, the terms *weak, low, okay, strong*, and *high* are perceptions about the level of pressure.

Example 2.8 Let X be the set of all men. The term "height" can mean different things to different people. Figure 2.7 depicts two sets of MFs for the set of terms {*short men, medium men, tall men*}. Clearly, the terms *short men*, *medium men*, and *tall men* will have a very different meaning to a professional basketball player than they will to most other people. This illustrates the fact that MFs can be quite context dependent.

The number of linguistic terms in T for a linguistic variable v will affect the calibration of the fuzzy sets, (i.e., the specification of its MFs). If, for example, only three linguistic terms are used to describe Height, namely {*short*, *medium*, *tall*}, then their fuzzy sets will look very different from their fuzzy sets when the following seven terms are used: {*very short*, *moderately short*, *short*, *medium*, *moderately tall*, *tall*, *very tall*}. This is because the terms *very short* and *moderately short* now appear before *short*, and *tall* is sandwiched between *moderately tall* and *very tall*. In many applications of rule-based fuzzy systems, the names that are given to the fuzzy sets are unimportant because interpretability of the rules is unimportant; however, there are other applications where interpretability of rules is very important. More is said about this in Sects. 3.9.5 and 9.13.6.

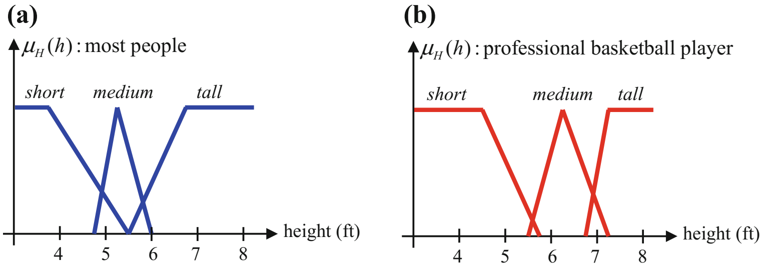


Fig. 2.7 MFs for T(height) = {*short men*, *medium men*, *tall men*}. **a** Most people's MFs, and **b** professional basketball player's MFs (Mendel 1995a © 1995, IEEE)

2.2.5 Returning to Linguistic Labels from Numerical Values of MFs

Sometimes it is necessary to go from MF numerical values for a variable to a linguistic description of that variable. This section examines how to do this for type-1 fuzzy sets.

Consider, for example, the linguistic variable temperature that has been decomposed into five terms {*very negative, medium negative, near zero, medium positive, very positive*}, and the situation depicted in Fig. 2.8 at $x = x'$. This value of x only generates a non-zero membership value in the fuzzy set F_4 = *medium positive*; hence, $x = x'$ can be described linguistically, without any ambiguity, as "*medium positive*."

The situation at $x = x''$ is different, because this value of x generates a non-zero membership value in two fuzzy sets F_4 = *medium positive* and F_5 = *very positive*. It would be very awkward to speak of x'' as "being *medium positive* to degree $\mu_{F_4}(x'')$ and *very positive* to degree $\mu_{F_5}(x'')$." People just don't communicate this way. Instead, $\mu_{F_4}(x'')$ and $\mu_{F_5}(x'')$ are compared to see which is larger, and then x'' is assigned to the set associated with the larger value; hence, in this example, x'' would be described as being "*medium positive.*"

What has just been explained can be descried formally as follows. Given P fuzzy sets F_i with MFs $\mu_{F_i}(x)\ (i = 1, \ldots, p)$. When $x = x'$, evaluate all p MFs at this point, and then compute $\max[\mu_{F_1}(x'), \mu_{F_2}(x'), \ldots, \mu_{F_p}(x')] \equiv \mu_{F_m}(x')$. Let $L(x')$ denote the linguistic label associated with x'. Then, $L(x') \equiv F_m$, i.e.,

$$L(x') = \arg\max_{\forall F_i}\left[\mu_{F_1}(x'), \mu_{F_2}(x'), \ldots, \mu_{F_p}(x')\right] \tag{2.10}$$

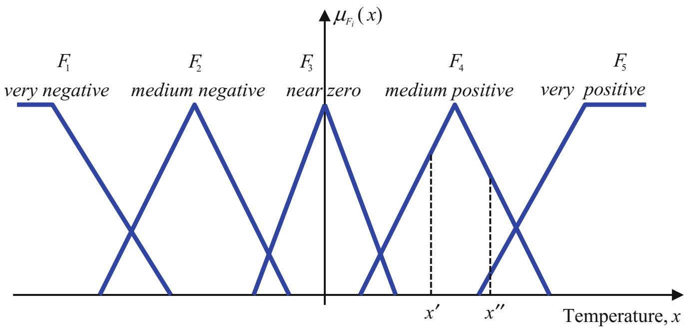


Fig. 2.8 Returning to a linguistic label for type-1 fuzzy sets

2.3 Set Theoretic Operations for Crisp Sets

Now that fuzzy sets have been defined, what can one do with them? The same question could be asked about crisp sets, and one knows that there are lots of things that can be done with them; hence, it is expected that analogous things can be done with fuzzy sets. To begin, the elementary crisp-set operations of union, intersection, and complement are briefly reviewed.

Let A and B be two subsets of X. The *union* of A and B, denoted $A \cup B$, contains all of the elements in either A or B, i.e.,

$$\mu_{A\cup B}(x) = \begin{cases} 1 & \text{if } x \in A \text{ or } x \in B \\ 0 & \text{if } x \notin A \text{ and } x \notin B \end{cases} \tag{2.11}$$

The *intersection* of A and B, denoted $A \cap B$, contains all of the elements that are simultaneously in A and B, i.e.,

$$\mu_{A\cap B}(x) = \begin{cases} 1 & \text{if } x \in A \text{ and } x \in B \\ 0 & \text{if } x \notin A \text{ or } x \notin B \end{cases} \tag{2.12}$$

Let $\bar{A}$ denote the *complement* of A; it contains all the elements that are not in A, i.e.,

$$\mu_{\bar{A}}(x) = \begin{cases} 1 & \text{if } x \notin A \\ 0 & \text{if } x \in A \end{cases} \tag{2.13}$$

From these facts, it is easy to show that:

$$A \cup B \Rightarrow \mu_{A\cup B}(x) = \max[\mu_A(x), \mu_B(x)] \tag{2.14}$$

$$A \cap B \Rightarrow \mu_{A\cap B}(x) = \min[\mu_A(x), \mu_B(x)] \tag{2.15}$$

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x) \tag{2.16}$$

Consider $\mu_{A\cup B}(x)$ for example. In this case, $x \in A$ or $x \in B$ means:

$$(\mu_A(x) = 1, \mu_B(x) = 1) \text{ or } (\mu_A(x) = 1, \mu_B(x) = 0) \text{ or } (\mu_A(x) = 0, \mu_B(x) = 1),$$

and, for each of these situations, $\max[\mu_A(x), \mu_B(x)] = 1$. Additionally, $x \notin A$ and $x \notin B$ means $(\mu_A(x) = 0, \mu_B(x) = 0)$ for which $\max[\mu_A(x), \mu_B(x)] = 0$. Consequently, $\max[\mu_A(x), \mu_B(x)]$ for $\forall x$ does provide the correct MF, given in (2.11), for union.

The formulas in (2.14)–(2.16), for $\mu_{A\cup B}(x)$, $\mu_{A\cap B}(x)$, and $\mu_{\bar{A}}(x)$, are very useful for proving other theoretical properties about crisp sets. Note, also, that the maximum and minimum are not the only ways to describe $\mu_{A\cup B}(x)$ and $\mu_{A\cap B}(x)$. While these formulas are not usually part of conventional set theory, they are essential to fuzzy set theory; however, as has just been demonstrated, they really do occur in conventional set theory. See, e.g., Klir and Folger (1988) and Yager and Filev (1994) for other ways to characterize these operations.

The crisp union and intersection operations satisfy many properties (see Table 2.8 in Appendix 1 to this chapter for an extensive list of these properties), including:

1. Commutative

$$A \cup B = B \cup A$$

2. Associative

$$A \cup B \cup C = (A \cup B) \cup C = A \cup (B \cup C)$$

3. Distributive

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \text{ and } A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

These properties can be proved either by Venn diagrams or by means of the MF definition given in (2.1).

De Morgan's laws for crisp sets are:

$$\bullet\, \overline{A \cup B} = \bar{A} \cap \bar{B}$$
$$\bullet\, \overline{A \cap B} = \bar{A} \cup \bar{B}$$

These laws, which are also very useful in proving things about more complicated operations on sets, can also be proved either by Venn diagrams or by means of the MF definition given in (2.1).

The two fundamental (Aristotelian) laws of crisp set theory are:

1. *Law of Excluded Middle*: $A \cup \overline{A} = X$ (i.e., a set and its complement must comprise the universe of discourse).
2. *Law of Contradiction* $A \cap \overline{A} = \emptyset$ (i.e., an element can either be in its set or its complement; it cannot simultaneously be in both).

Fuzzy sets usually break these Aristotelian laws.

2.4 Set Theoretic Operations for Type-1 Fuzzy Sets

For fuzzy sets, union, intersection, and complement are defined in terms of their MFs. Let fuzzy sets A and B be described by their MFs $\mu_A(x)$ and $\mu_B(x)$. One definition of *fuzzy union* leads to the MF

$$\mu_{A\cup B}(x) = \max[\mu_A(x), \mu_B(x)] \tag{2.17}$$

and one definition of *fuzzy intersection* leads to the MF

$$\mu_{A\cap B}(x) = \min[\mu_A(x), \mu_B(x)] \tag{2.18}$$

Additionally, the MF for *fuzzy complement* is

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x) \tag{2.19}$$

Obviously, these three definitions were motivated by their crisp counterparts in (2.14)–(2.16).

Example 2.9 (Mendel 1995a) In engineering the **damping ratio** is a dimensionless measure describing how oscillations in a system decay after a disturbance. Consider the fuzzy sets A = damping ratio x *considerably larger* than 0.5, and B = damping ratio x *approximately* 0.707. Note that damping ratio is a positive real number, i.e., its universe of discourse, X, is the positive real numbers $0 \le x \le 1$. Consequently, $A = \{(x, \mu_A(x))|x \in X\}$ and $B = \{(x, \mu_B(x))|x \in X\}$, where, for example, $\mu_A(x)$ and $\mu_B(x)$ are specified (by this author), as:

$$\mu_A(x) = \begin{cases} 0 & \text{if } 0 \le x \le 0.5 \\ \dfrac{5(x-0.5)^2}{1+(x-0.5)^2} & \text{if } 0.5 < x \le 1 \end{cases} \tag{2.20}$$

and

$$\mu_B(x) = \frac{1}{[1+(x-0.707)^4]} \quad 0 \le x \le 1 \tag{2.21}$$

Figure 2.9 depicts $\mu_A(x), \mu_B(x), \mu_{A\cup B}(x), \mu_{A\cap B}(x)$ and $\mu_{\bar{B}}(x)$. Observe, from Fig. 2.9a that $\mu_A(0.707) + \mu_B(0.707) > 1$ and from Fig. 2.9d, that the point $x = 0.5$ exists in both B and $\overline{B}$ simultaneously, but to different degrees, because $\mu_B(0.5) \neq 0$ and $\mu_{\bar{B}}(0.5) \neq 0$.

This example demonstrates that for fuzzy sets the classical Laws of Excluded Middle and Contradiction are broken, i.e., *for fuzzy sets*: $A \cup \overline{A} \neq X$ and $A \cap \overline{A} \neq \emptyset$. This has also been observed in Fig. 2.4 for the automobile Example 2.1 (continued). In fact, one of the ways to describe the difference between crisp set theory and fuzzy set theory is to explain that these two laws do not hold in fuzzy set theory.[9]

The maximum and minimum operators are not the only ones that could have been chosen to model fuzzy union and fuzzy intersection. Zadeh, in his pioneering first paper (Zadeh 1965), defined two operators each for fuzzy union and fuzzy intersection, namely:

[9]There is a small subset of type-1 fuzzy set theory that requires both of these laws to be satisfied. This work has had no impact on rule-based fuzzy systems and so it is not discussed in this book.

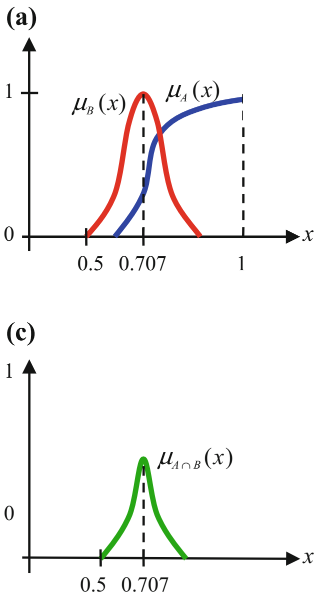


Fig. 2.9 MFs associated with A = damping ratio x *considerably larger* than 0.5, and B = damping ratio x *approximately* 0.707. **a** $\mu_A(x)$ and $\mu_B(x)$, **b** $\mu_{A\cup B}(x)$, **c** $\mu_{A\cap B}(x)$, and **d** $\mu_{\bar{B}}(x)$

1. Fuzzy union: maximum and algebraic sum, where for the latter

$$\mu_{A\cup B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x) \tag{2.22}$$

2. Fuzzy intersection: minimum and algebraic product, where for the latter

$$\mu_{A\cap B}(x) = \mu_A(x)\mu_B(x) \tag{2.23}$$

Later, other operators that have an axiomatic basis (e.g., Klir and Yuan 1995) were introduced (in all cases, $x, y \in [0, 1]$):

1. **t-conorm operators**[10] for fuzzy union (also known as s-norm and denoted $\oplus$). The maximum and algebraic sums are t-conorms; some other examples of t-conorms are:
   - *Bounded sum*: $x \oplus y = \min(1, x + y)$

[10]The axiomatic basis for a *t-conorm* is, for *a, b, d* $\in$ [0,1]: (1) boundary condition, $s(a, 0) = a$; (2) monotonicity, $b \le d \Rightarrow s(a, b) \le s(a, d)$; (3) commutativity, $s(a, b) = s(b, a)$; and, (4) associativity, $s(a, s(b, d)) = s(s(a, b), d)$. Table 3.3 in Klir and Yuan (1995) lists 11 t-conorms.

- *Drastic sum*: $x \oplus y = \begin{cases} x & \text{if } y = 0 \\ y & \text{if } x = 0 \\ 1 & \text{otherwise} \end{cases}$

2. **t-norm operators**[11] for fuzzy intersection (denoted ★). The minimum and algebraic product are t-norms; some other examples of t-norms are:
   - *Bounded product*: $x \bigstar y = \max(0, x + y - 1)$
   - *Drastic product*: $x \bigstar y = \begin{cases} x & \text{if } y = 1 \\ y & \text{if } x = 1 \\ 0 & \text{otherwise} \end{cases}$

There is even an axiomatic basis for the complement (denoted c) of a fuzzy set.[12] In engineering applications, most people use the fuzzy complement whose MF is given in (2.19).

As pointed out in Zimmerman (1991), dual pairs[13] of t-norms and t-conorms with respect to the fuzzy complement in (2.19) satisfy the following generalization of DeMorgan's laws (Bonissone and Decker 1986):

$$s[\mu_A(x), \mu_B(x)] = c\{t[c(\mu_A(x)), c(\mu_B(x))]\} \tag{2.24}$$

$$t[\mu_A(x), \mu_B(x)] = c\{s[c(\mu_A(x)), c(\mu_B(x))]\} \tag{2.25}$$

where $x \in X$. For example,

$$\max[\mu_A(x), \mu_B(x)] = 1 - \min[1 - \mu_A(x), 1 - \mu_B(x)] \tag{2.26}$$

and

$$\min[\mu_A(x), \mu_B(x)] = 1 - \max[1 - \mu_A(x), 1 - \mu_B(x)] \tag{2.27}$$

Note, also, that there are other ways of combining fuzzy sets, e.g., the *fuzzy and*, *fuzzy or*, *compensatory and*, and *compensatory or*; e.g., see Zimmerman (1991) and Yager and Filev (1994).

If at this point you are puzzled by all of the possible choices, a discussion about this is provided in Sect. 2.18.

---

[11]The axiomatic basis for a *t-norm* is, for *a, b, d* ∈ [0,1]: (1) boundary condition, $t(a, 1) = a$; (2) monotonicity, $b \le d \Rightarrow t(a, b) \le t(a, d)$; (3) commutativity, $t(a, b) = t(b, a)$; and, (4) associativity, $t(a, t(b, d)) = t(t(a, b), d)$. Table 3.2 in Klir and Yuan (1995) lists 11 t-norms.

[12]The axiomatic basis for a *fuzzy complement* is: (1) boundary conditions, $c(0) = 1$ and $c(1) = 0$, and (2) monotonicity, for all $a, b \in [0, 1]$, if $a \le b$ then $c(a) \ge c(b)$. There are also many fuzzy complements that additionally satisfy the involutive condition $c(c(a)) = a$.

[13]Some examples of dual pairs with respect to the fuzzy complement (2.19) are: min and max, and product and algebraic sum. See Klir and Yuan (1995, pp. 83–88) for discussions about and properties of dual pairs. Some of their Chap. 3 end-notes provide interesting historical remarks about the origins of t-norms and t-conorms.

Not only are the union, intersection, and complement performed with type-1 fuzzy sets, but sometimes other important set-theoretic operations are performed on them using well-known laws; e.g., commutative, associative, distributive, and De Morgan's laws (see Table 2.8 for a list of all the laws). For this book, an important question that needs to be answered is:

- Is it permissible to use a particular law for type-1 fuzzy sets under maximum t-conorm and either minimum or product t-norms?

Our focus is just on the maximum t-conorm and the minimum or product t-norms, because these are the most widely used ones in the fuzzy system's literature. The question must, of course, be re-examined if one uses other t-conorms and t-norms. Because the studies into the answers to this question, although important, are very technical, their details are presented in Appendix 1. Here, just the results are stated and some conclusions about them are drawn.

The aforementioned question has been very well studied for type-1 fuzzy sets (see Table 2.8) *For maximum t-conorm and minimum t-norm all laws are satisfied*; however, *for maximum t-conorm and product t-norm certain laws are not satisfied*. This means, therefore, that *one must be careful when using maximum t-conorm and product t-norm*. If, for example, the design of a maximum t-conorm and product t-norm type-1 fuzzy system involves the use of any of the violated laws it will be in error. *Fortunately, one usually does not have to use any of the violated laws in the creation and design of a type-1 fuzzy system*. The same cannot be said, in general, for other applications of type-1 fuzzy sets.

2.5 Crisp Relations and Compositions on the Same Product Space

According to Klir and Folger (1988, p. 65): *A crisp relation represents the presence or absence of association, interaction, or interconnectedness between the elements of two or more sets*. Here our attention is limited to relations between two sets U and V, i.e., to binary relations denoted $R(U, V)$. $U \times V$ denotes the Cartesian product of the two crisp sets U and V, i.e.,

$$U \times V = \{(u, v) | u \in U \text{ and } v \in V\} \tag{2.28}$$

$R(U, V)$ is a subset of $U \times V$.

Crisp relation $R(U, V)$ can be defined by the following MF:

$$\mu_R(u, v) = \begin{cases} 1 & \text{if and only if } (u, v) \in R(U, V) \\ 0 & \text{otherwise} \end{cases} \tag{2.29}$$

For binary relations defined over a Cartesian product whose elements come from a discrete universe of discourse, it is convenient to collect the MFs into a relational matrix whose elements are either zero or unity. An equivalent representation for a binary relation is a sagittal diagram, in which the sets U and V are each represented by a set of nodes in the diagram that are clearly distinguished from one another. Elements of $U \times V$ with non-zero membership grade in $R(U, V)$ are represented in the diagram by lines connecting the respective nodes. Although not explicitly shown, the lines have membership values equal to unity.

Example 2.10 (Mendel 1995a) Let R represent the relation of *stability* between the set of all linear, second-order continuous-time systems and the set of the poles of such systems. Of all the possible pairings of linear second-order continuous-time systems and poles, only those pairs whose members are time-invariant with poles lying either in the left-half of the complex s-plane or on the imaginary axis of that plane are known to be stable.

Let $U = \{u_1, u_2\}$ = {linear second-order time-varying continuous-time system, linear second-order time-invariant continuous-time system}, and $V = \{y_1, y_2, y_3\}$ = {poles lie in the left-half s-plane, poles lie on the $j\omega$ axis, poles lie in the right-half s-plane}. The Cartesian product $U \times V$ can be visualized as a 2×3 array of ordered pairs, e.g., the (1, 2) element is (linear second-order time-varying continuous-time system, poles lie on the $j\omega$ axis). The stability relation $R(U, V)$ is the following subset of $U \times V$:

$$\begin{aligned} R(U,V) = \{&(\text{linear second-order time-invariant continuous-time system, poles lie in the left-half } s\text{-plane}),\\ &(\text{linear second-order time-invariant continuous-time system, poles lie on the } j\omega \text{ axis} \end{aligned}$$

The relational matrix for this stability relation is:

$$\begin{array}{cc} & \begin{array}{ccc} v_1 & v_2 & v_3 \end{array} \\ \begin{array}{c} u_1 \\ u_2 \end{array} & \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \end{array}$$

The sagittal diagram for this stability relation is depicted in Fig. 2.10.

Let $R(u, v)$ and $S(u, v)$—R and S for short—be two crisp relations in the same Cartesian product space $U \times V$. The intersection and union of R and S, which are *compositions* of the two relations, are computed using (2.17) and (2.18), because a relation is a set.

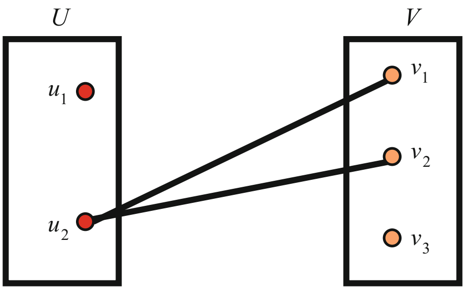


Fig. 2.10 Sagittal diagram for relation of stability between the set of all linear, second-order continuous-time systems and the set of poles of such systems (Mendel 1995a © 1995, IEEE)

2.6 Fuzzy Relations and Compositions on the Same Product Space

A fuzzy relation represents a degree *of presence or absence of association, interaction, or interconnectedness between the elements of two or more fuzzy sets.* Some examples of binary fuzzy relations are:

- x is much larger than y
- y is very close to x
- z is much greener than y
- system 1 is less damped than system 2
- bandwidth of system A is larger than that of system B
- tone C is of higher local signal-to-noise ratio than tone D
- a is more profitable than b.

A binary type-1 fuzzy relation $F(A_1, A_2)$ is (Lin and Lee 1996) a type-1 fuzzy set that is defined on the Cartesian product space of crisp sets A_1 and A_2, where tuples (a_1, a_2) may have varying degrees of membership $\mu_F(a_1, a_2)$ within the relation. More specifically,

$$F(A_1, A_2) = \int_{A_1 \times A_2} \mu_F(a_1, a_2)/(a_1, a_2) \quad a_1 \in A_1 \text{ and } a_2 \in A_2 \tag{2.30}$$

where $\mu_F(a_1, a_2) \in [0, 1]$. It is important to note that the elements of $F(A_1, A_2)$ are numbers and not fuzzy sets. When they are type-1 fuzzy sets, the fuzzy relation becomes a type-2 fuzzy relation (Sect. 7.6).

Example 2.11 (Mendel 1995a) Let U and V be the real numbers, and consider the fuzzy relation "target x is *close* to target y." Here is one MF for this relation:

$$\mu_c(|x - y|) \equiv \max\{(5 - |x - y|)/5, 0\} \tag{2.31}$$

This relational MF is depicted in Fig. 2.11. Note that the distance between the two targets $|x - y|$ is treated as the independent variable.

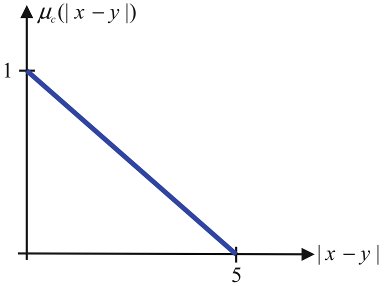


Fig. 2.11 Relational MF $\mu_c(|x - y|)$ (Mendel 1995a © 1995, IEEE)

Because fuzzy relations are fuzzy sets in a Cartesian product space, set theoretic and algebraic operations can be defined for them using the earlier operators for fuzzy union, intersection and complement. Let $R(U, V)$ and $S(U, V)$ (shortened in the sequel to R and S) be two fuzzy relations in the *same* Cartesian product space $U \times V$. The intersection and union of R and S, which are *compositions* of the two relations, are then defined as:

$$\mu_{R \cap S}(x, y) = \mu_R(x, y) \bigstar \mu_S(x, y) \tag{2.32}$$

$$\mu_{R \cup S}(x, y) = \mu_R(x, y) \oplus \mu_S(x, y) \tag{2.33}$$

where $\bigstar$ is any t-norm, and $\oplus$ is any t-conorm.

Example 2.12 Consider the two somewhat contradictory fuzzy relations "u is close to v" *and* "u is smaller than v," and also the less-contradictory relations "u is close to v" *or* "u is smaller than v." All relations are on the same Cartesian product space $U \times V$. For simplicity, it is assumed here that $U = \{u_1, u_2\} = \{2, 12\}$ and $V = \{v_1, v_2, v_3\} = \{1, 7, 13\}$. Let the MFs for *close* and *smaller* be denoted as $\mu_c(u, v)$ and $\mu_s(u, v)$, respectively, where the numbers in $\mu_c(u, v)$ and $\mu_s(u, v)$ have been chosen to agree with a comparison of the numbers in U and V.

$$\mu_c(u, v) \equiv \begin{array}{c c} & \begin{array}{ccc} v_1 & v_2 & v_3 \end{array} \\ \begin{array}{c} u_1 \\ u_2 \end{array} & \begin{pmatrix} 0.9 & 0.4 & 0.1 \\ 0.1 & 0.4 & 0.9 \end{pmatrix} \end{array} \tag{2.34}$$

and

$$\mu_s(u, v) \equiv \begin{array}{c c} & \begin{array}{ccc} v_1 & v_2 & v_3 \end{array} \\ \begin{array}{c} u_1 \\ u_2 \end{array} & \begin{pmatrix} 0 & 0.6 & 1 \\ 0 & 0 & 0.3 \end{pmatrix} \end{array} \tag{2.35}$$

The membership grades for the union and intersection of these relations, assuming minimum t-norm ($\wedge$) and maximum t-conorm ($\vee$), can be found as (i = 1, 2 and j = 1, 2, 3)

$$\mu_{c \cup s}(u_i, v_j) = \mu_c(u_i, v_j) \vee \mu_s(u_i, v_j) \tag{2.36}$$

and

$$\mu_{c \cap s}(u_i, v_j) = \mu_c(u_i, v_j) \wedge \mu_s(u_i, v_j) \tag{2.37}$$

Using (2.36) and (2.37), it is easy to show that

$$\mu_{c\cup s}(u,v) = \begin{array}{c} \\ u_1 \\ u_2 \end{array} \begin{array}{c} \begin{array}{ccc} v_1 & v_2 & v_3 \end{array} \\ \begin{pmatrix} 0.9 & 0.6 & 1 \\ 0.1 & 0.4 & 0.9 \end{pmatrix} \end{array} \tag{2.38}$$

and

$$\mu_{c\cap s}(u,v) = \begin{array}{c} \\ u_1 \\ u_2 \end{array} \begin{array}{c} \begin{array}{ccc} v_1 & v_2 & v_3 \end{array} \\ \begin{pmatrix} 0 & 0.4 & 0.1 \\ 0 & 0 & 0.3 \end{pmatrix} \end{array} \tag{2.39}$$

From (2.38) and (2.39), "u is close to v" *or* "u is smaller than v" is seen to be much more sensible than "u is close to v" *and* "u is smaller than v," because membership values in $\mu_{c\cup s}(u,v)$ are fairly large, whereas those in $\mu_{c\cap s}(u,v)$ are mostly small.

2.7 Crisp Relations and Compositions on Different Product Spaces

Consider[14] two different product spaces, $U \times V$ and $V \times W$, that share a common set and let $R(U,V)$ and $S(V,W)$ be two *crisp* relations on these spaces. The composition of these relations is defined (Klir and Folger 1988, p. 75) as: *a subset* $T(U,W)$ *of* $U \times W$ *such that* $(u,w) \in T$ *if and only if* $(u,v) \in R$ *and* $(v,w) \in S$. This can be expressed as a *max–min composition*, *max–product composition* or, in general, as the following *sup-star composition* for crisp relations:

$$\mu_{R\circ S}(u,w) = \sup_{v\in V}[\mu_R(u,v)\bigstar\mu_S(v,w)]\; u \in U, w \in W \tag{2.40}$$

where $\star$ indicates any suitable t-norm operation. The validity of the *sup-star* composition for crisp set is shown in Wang (1997, p. 54). If R and S are two crisp relations on $U \times W$ and $V \times W$, respectively, then the membership for any pair (u,w), $u \in U$ and $w \in W$, is 1 if and only if there exists at least one $v \in V$ such that $\mu_R(u,v) = 1$ and $\mu_S(v,w) = 1$. In Zadeh (1973) it is shown that this condition is equivalent to having the sup-star composition equal to 1. Because this is a special case of the sup-star composition for fuzzy sets (a crisp set is a special case of a fuzzy set), whose proof is given in Sect. 2.8, the proof of (2.40) for crisp sets is not included here.

[14]Most of this paragraph is taken from Karnik and Mendel (2001, p. 337).

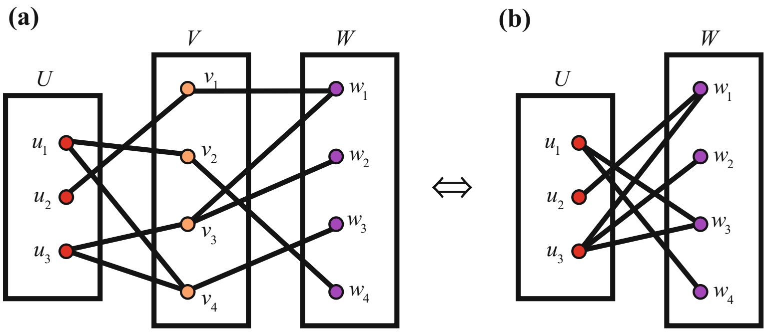


Fig. 2.12 Saggital diagram for Example 2.13. **a** Original diagram for relations $R_1(U, V)$ and $R_2(V, W)$ and **b** compositional diagram for $R_3(U, W)$ (Mendel 1995a © 1995, IEEE)

Example 2.13 Given the saggital diagrams depicted in Fig. 2.12a, b, one concludes that the relational matrices $R_1(U, V)$, $R_2(V, W)$, and $R_3(U, W)$ are:

$$R_1(U,V) = \begin{array}{c} \\ u_1 \\ u_2 \\ u_3 \end{array} \begin{array}{c} \begin{array}{cccc} v_1 & v_2 & v_3 & v_4 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{array} \tag{2.41}$$

$$R_2(V,W) = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \end{array} \begin{array}{c} \begin{array}{cccc} w_1 & w_2 & w_3 & w_4 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array} \tag{2.42}$$

$$R_3(U,W) = \begin{array}{c} \\ u_1 \\ u_2 \\ u_3 \end{array} \begin{array}{c} \begin{array}{cccc} w_1 & w_2 & w_3 & w_4 \end{array} \\ \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \end{array} \tag{2.43}$$

Because it is not efficient to keep describing compositions in terms of sagittal diagrams, a *formula* is needed that conveys the same information.

Definition 2.8 *The* max–min composition *of relations* $R(U, V)$ *and* $S(V, W)$ *is defined by the MF* $\mu_{R \circ S}(u, w)$, *where*

$$\mu_{R \circ S}(u,w) = \left\{ (u,w), \max_{v} [\min(\mu_R(u,v), \mu_S(v,w))] \right\} \tag{2.44}$$

The max-product composition *of relations* $R(U,V)$ *and* $S(V,W)$ *is defined by the MF* $\mu_{R\times S}(u,w)$, *where*

$$\mu_{R\times S}(u,w) = \left\{(u,w), \max_{v}[\mu_R(u,v)\mu_S(v,w)]\right\} \tag{2.45}$$

Clearly, the *max–min* or *max–product* compositions lead to the correct relational matrix $R(U,W)$, because they are special cases of the sup-star composition in (2.40).

Example 2.14 Here (2.44) and (2.45) are verified for the (1, 2) element of $R_3(U,W)$ in (2.43). For this element, (2.44) becomes

$$\begin{aligned}\mu_{R_3}(u_1,w_2) &= \left\{(u_1,w_2), \max_{v}[\min(\mu_{R_1}(u_1,v), \mu_{R_2}(v,w_2))]\right\}\\ &= \{(u_1,w_2), \max[\min(\mu_{R_1}(u_1,v_1), \mu_{R_2}(v_1,w_2)),\\ &\quad \min(\mu_{R_1}(u_1,v_2), \mu_{R_2}(v_2,w_2)), \min(\mu_{R_1}(u_1,v_3), \mu_{R_2}(v_3,w_2)),\\ &\quad \min(\mu_{R_1}(u_1,v_4), \mu_{R_2}(v_4,w_2))]\}\\ &=\{(u_1,w_2), \max[\min(0,0), \min(1,0), \min(0,1), \min(1,0)]\}\\ &=\{(u_1,w_2), \max[0,0,0,0]\} = \{(u_1,w_2), 0\}\end{aligned} \tag{2.46}$$

which agrees with (2.43). Similarly, (2.45) becomes

$$\begin{aligned}\mu_{R_3}(u_1,w_2) &= \left\{(u_1,w_2), \max_{v}[\mu_{R_1}(u_1,v)\mu_{R_2}(v,w_2)]\right\}\\ &= \{(u_1,w_2), \max[\mu_{R_1}(u_1,v_1)\mu_{R_2}(v_1,w_2), \mu_{R_1}(u_1,v_2)\mu_{R_2}(v_2,w_2),\\ &\quad \mu_{R_1}(u_1,v_3)\mu_{R_2}(v_3,w_2), \mu_{R_1}(u_1,v_4)\mu_{R_2}(v_4,w_2)]\}\\ &=\{(u_1,w_2), \max[(0\times 0), (1\times 0), (0\times 1), (1\times 0)]\}\\ &=\{(u_1,w_2), \max[0,0,0,0]\} = \{(u_1,w_2), 0\}\end{aligned} \tag{2.47}$$

which also agrees with (2.43).

The following shortcuts can be used to evaluate the *max-min* or *max-product* compositions that involve relational matrices.

- **Max–min composition**: (1) Write out each element in the matrix product $Q(U,V)P(V,W)$; but, (2) treat each multiplication as a minimum operation; and, then, (3) treat each addition as a maximum operation.
- **Max–product composition**: (1) Write out each element in the *matrix product* $Q(U,V)P(V,W)$; but, (2) treat each multiplication as an algebraic multiplication operation; and, then, (3) treat each addition as a maximum operation.

Example 2.14 (Continued) Here these two shortcuts are used to again verify (2.44) and (2.45), but this time for the (1, 3) element of $R_3(U,W)$ in (2.43). Now applying the shortcut for the max-min composition to the $(1, 3)$ element of $R_1(u, v) \times R_2(v, w)$, one finds

$$\begin{aligned} R_3(u_1, w_3) &= 0 \times 0 + 1 \times 0 + 0 \times 0 + 1 \times 1 \\ &= \min(0, 0) + \min(1, 0) + \min(0, 0) + \min(1, 1) \\ &= \max(0, 0, 0, 1) = 1 \end{aligned} \tag{2.48}$$

Similarly, applying the shortcut for the max-product composition to the (1, 3) element of $R_1(u, v) \times R_2(v, w)$, one finds

$$\begin{aligned} R_3(u_1, w_3) &= 0 \times 0 + 1 \times 0 + 0 \times 0 + 1 \times 1 \\ &= \max(0, 0, 0, 1) = 1 \end{aligned} \tag{2.49}$$

Both of these results agree with the (1, 3) element of $R_3(U, W)$ in (2.43).

The *max–min* and *max–product* compositions are not the only ones that correctly represent $R(U, W)$; however, they seem to be the most widely used ones.

2.8 Fuzzy Relations and Compositions on Different Product Spaces

Next, consider the composition of fuzzy relations from different Cartesian product spaces that share a common set, namely $R(U, V)$ and $S(V, W)$, e.g., u is *smaller* than v, and v is *close* to w. The composition of fuzzy relations from different Cartesian product spaces that share a common set is defined analogously to the crisp composition, except that in the fuzzy case the sets are fuzzy sets. Associated with fuzzy relation R is its MF $\mu_R(u, v)$, where $\mu_R(u, v) \in [0, 1]$; and, associated with fuzzy relation S is its MF $\mu_S(v, w)$, where $\mu_S(v, w) \in [0, 1]$. In this respect, the condition on the composition of crisp relations, that is given below (2.40), can be rephrased as follows:

Theorem 2.1 *If*[15] *R and S are two type-1 fuzzy relations on $U \times V$ and $V \times W$, respectively, then the membership for any pair (u, w), $u \in U$ and $w \in W$, is non-zero if and only if there exists at least one $v \in V$ such that $\mu_R(u, v) \neq 0$ and $\mu_S(v, w) \neq 0$, i.e.:*

$$\mu_{R \circ S}(u, w) = \sup_{v \in V}[\mu_R(u, v) \bigstar \mu_S(v, w)] \; u \in U, v \in V \tag{2.50}$$

(2.50) *is called the* sup-star composition *for fuzzy relations.*

Proof This proof uses the following method: Let **A** be the statement "$\mu_{R \circ S}(u, w) \neq 0$," and **B** be the statement "there exists at least one $v \in V$ such that

[15] This theorem and its proof are taken from Karnik and Mendel, (1998, pp. 61–62).

$\mu_R(u,v) \neq 0$ *and* $\mu_S(v,w) \neq 0$." "**A** if and only if **B**" is proved by first proving that[16] $\bar{\mathbf{B}} \Rightarrow \bar{\mathbf{A}}$ (equivalent to proving that $\mathbf{A} \Rightarrow \mathbf{B}$, i.e., necessity of **B**) and then proving that $\bar{\mathbf{A}} \Rightarrow \bar{\mathbf{B}}$ (equivalent to proving that $\mathbf{B} \Rightarrow \mathbf{A}$, i.e., sufficiency of **B**).

Necessity—If there exists no $v \in V$ such that $\mu_R(u,v) \neq 0$ *and* $\mu_S(v,w) \neq 0$, then this means that for every $v \in V$, either $\mu_R(u,v)$ or $\mu_S(v,w)$ is equal to zero (or both are zero), which in turn implies that $\mu_R(u,v) \bigstar \mu_S(v,w) = 0$ for every $v \in V$; hence, the supremum[17] of $\mu_R(u,v) \bigstar \mu_S(v,w)$ over $v \in V$ is also zero, and therefore $\mu_{R \circ S}(u,w) = 0$, which is $\bar{\mathbf{A}}$.

Sufficiency—If the sup-star composition is zero then it must be true that $\mu_R(u,v) \bigstar \mu_S(v,w) = 0$ for every $v \in V$, which means that for every $v \in V$, either $\mu_R(u,v)$ or $\mu_S(v,w)$ (or both) is zero. This means that there is no $v \in V$ such that $\mu_R(u,v) \neq 0$ and $\mu_S(v,w) \neq 0$, which is $\bar{\mathbf{B}}$.

When R and S are from discrete universes of discourse, then $R \circ S$ can be described either by a sagittal diagram, in which each branch is labeled by its MF value, or a fuzzy relational matrix, in which each element is a positive real number between and including zero and unity. When U, V, and W are discrete universes of discourse, then the supremum operation in (2.50) is the *maximum*. Although it is permissible to use other t-norms, the most commonly used sup-star compositions are the *sup-min* and *sup-product*. The shortcuts for computing the sup-min and sup-product, given in Sect. 2.7, apply also to fuzzy compositions over discrete universes of discourse.

Example 2.15 Consider the type-1 relation "u is close to v" on $U \times V$, where $U = \{u_1, u_2\}$ and $V = \{v_1, v_2, v_3\}$ are given in Example 2.12 as U = {2, 12} and V = {1, 7, 13}, and $\mu_c(u,v)$ is given by (2.34). Now consider another type-1 fuzzy relation "v is much bigger than w" on $V \times W$, where $W = \{w_1, w_2\} = \{4, 8\}$, with the following MF, $\mu_{mb}(v,w)$, for *much bigger*, where the numbers in $\mu_{mb}(v,w)$ have been chosen to agree with a comparison of the numbers in V and W:

$$\mu_{mb}(v,w) \equiv \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \end{array} \begin{array}{c} \begin{array}{cc} w_1 & w_2 \end{array} \\ \begin{pmatrix} 0 & 0 \\ 0.6 & 0 \\ 1 & 0.7 \end{pmatrix} \end{array} \tag{2.51}$$

The statement "u is close to v *and* v is much bigger than w" indicates the composition of these two type-1 relations. This composition can be found by using (2.50) and, e.g., the minimum t-norm, as follows ($i = 1, 2$ and $j = 1, 2, 3$):

[16] Recall that $\overline{\text{at least } (\cdot)} = \text{no } (\cdot)$.

[17] Let S be a set of real numbers. An upper bound for S is a number b such that $x \leq b$ for all $x \in S$. The *supremum* of S, if it exists, is the smallest upper bound for S. An upper bound that actually belongs to the set is called *a maximum*.

$$\begin{aligned}\mu_{comb}(u_i, w_j) = & \left[\mu_c(u_i, v_1) \wedge \mu_{mb}(v_1, w_j)\right] \vee \left[\mu_c(u_i, v_2) \wedge \mu_{mb}(v_2, w_j)\right] \\ & \vee \left[\mu_c(u_i, v_3) \wedge \mu_{mb}(v_3, w_j)\right]\end{aligned} \tag{2.52}$$

where $\wedge$ denotes minimum, and $\vee$ denotes maximum. For example,

$$\begin{aligned}\mu_{comb}(u_1, w_1) &= [\mu_c(u_1, v_1) \wedge \mu_{mb}(v_1, w_1)] \vee [\mu_c(u_1, v_2) \wedge \mu_{mb}(v_2, w_1)] \\ &\quad \vee [\mu_c(u_1, v_3) \wedge \mu_{mb}(v_3, w_1)] \\ &= [0.9 \wedge 0] \vee [0.4 \wedge 0.6] \vee [0.1 \wedge 1] \\ &= 0 \vee 0.4 \vee 0.1 = 0.4\end{aligned} \tag{2.53}$$

Doing all the calculations in a similar manner, one finds (Exercise 2.17):

$$\mu_{comb}(u, w) = \begin{array}{c c} & \begin{array}{cc} w_1 & w_2 \end{array} \\ \begin{array}{c} u_1 \\ u_2 \end{array} & \begin{pmatrix} 0.4 & 0.1 \\ 0.9 & 0.7 \end{pmatrix} \end{array} \tag{2.54}$$

Unlike the case of crisp compositions, for which exactly the same results are obtained using either the max-min or max-product compositions, the same results are not obtained in the case of fuzzy compositions. This is a major difference between fuzzy and crisp compositions.

Suppose fuzzy relation R is just a fuzzy set, in which case $V = U$, so that $\mu_R(u, v)$ just becomes $\mu_R(u)$ [or $\mu_R(v)$], e.g., "v is *medium large* and v is *smaller than* w." What happens to the sup-star composition in this case? Because $V = U$,

$$\sup_{v \in V}[\mu_R(u, v) \bigstar \mu_S(v, w)] = \sup_{u \in U}[\mu_R(u) \bigstar \mu_S(u, w)] \tag{2.55}$$

which is only a function of output variable w; hence, the notation $\mu_{R \circ S}(u, w)$ can be simplified to $\mu_{R \circ S}(w)$, so that *when R is just a fuzzy set*,

$$\mu_{R \circ S}(w) = \sup_{u \in U}[\mu_R(u) \bigstar \mu_S(u, w)] \; w \in W \tag{2.56}$$

Eq. (2.56), which is also known as Zadeh's *compositional rule of inference* (Zadeh 1973), is used a lot in Chap. 3 as the type-1 inference mechanism for a rule.

Example 2.16 Consider again the Example 2.12 relation "u is close to v" on $U \times V$, where $U = \{2, 12\}$ and $V = \{1, 7, 13\}$. The MF for $\mu_c(u, v)$ is given in (2.34). Let the fuzzy set "small" on U be defined as

$$\mu_s(u) \equiv \begin{array}{c} \begin{array}{cc} u_1 & u_2 \end{array} \\ (\, 0.9 \quad 0.1 \,) \end{array} \tag{2.57}$$

The composition of the two statements "u is small *and* u is close to v" can be obtained by using (2.56) as follows ($j = 1, 2, 3$):

$$\mu_{s\circ c}(v_j) = \left[\mu_s(u_1) \wedge \mu_c(u_1, v_j)\right] \vee \left[\mu_s(u_2) \wedge \mu_c(u_2, v_j)\right] \tag{2.58}$$

Using (2.57), it is straightforward to show that

$$\begin{array}{rccc} & v_1 & v_2 & v_3 \\ \mu_{s\circ c}(v) = & (\,0.9 & 0.4 & 0.1\,) \end{array} \tag{2.59}$$

For discrete universes of discourse, the max–min or max-product compositions in (2.56) can be evaluated using the shortcuts described earlier; however, first a row matrix for $\mu_R(u)$ must be created i.e., if $u \in U = \{u_1, u_2, \ldots, u_n\}$ and $R(U) = (\mu_R(u_1), \mu_R(u_2), \ldots, \mu_R(u_n))$, then:

- **Max–min composition**: (1) Write out each element in the matrix product $R(U)S(U, W)$, but (2) treat each multiplication as a minimum operation, and then (3) treat each addition as a maximum operation.
- **Max-product composition**: (1) Write out each element in the *matrix product* $R(U)S(U, W)$, but (2) treat each multiplication as an algebraic multiplication operation, and then (3) treat each addition as a maximum operation.

2.9 Hedges

A *linguistic hedge* or modifier,[18] introduced first in Zadeh (1972), is an operation that modifies the meaning of a term, or more generally, of a fuzzy set. For example, if *weak pressure* is a fuzzy set, then *very weak pressure*, *more-or-less weak pressure*, *extremely weak pressure*, and *not-so weak pressure* are examples of hedges that are applied to this fuzzy set. There are a multitude of hedges, many additional examples of which can be found in Schmucker (1984), Cox (1994).

There are two ways to handle a hedge:

1. They can be viewed as operators that act on a fuzzy set's MF to modify it.
2. They can be treated as new linguistic terms.

By the first approach, one establishes a set of primary terms and their MFs. The hedge operators then operate on some[19] or all of the primary terms, leading to a larger set of terms and their MFs. Three hedge operators introduced in Zadeh (1972) are:

1. *Concentration*: $\mu_{con(F)}(x) \equiv [\mu_F(x)]^2$. If, e.g., *weak pressure* has MF $\mu_{WP}(p)$, then *very weak pressure* is a fuzzy set with MF $[\mu_{WP}(p)]^2$, and *very very weak pressure* is a fuzzy set with MF $[\mu_{WP}(p)]^4$. Because MFs have been assumed to

[18] Some of the material in this section is taken from Mendel and Wu (2010, Sect. 3.6) and Mendel (1995a, p. 356).

[19] Hedges should only operate on primary terms for which the hedged term makes linguistic sense, e.g. the hedge *much* makes no linguistic sense when it is applied to the primary term *low pressure*.

be normalized, it is clear that the operation of concentration leads to a MF that lies within the MF of the original fuzzy set (thus, the term *concentration*); both have the same support, and the same membership values where the value of the original MF equals unity or zero.

2. *Dilation*: $\mu_{dil(F)}(x) \equiv [\mu_F(x)]^{1/2}$. If, e.g., *weak pressure* has MF $\mu_{WP}(p)$, then *more or less weak pressure* is a fuzzy set with MF $[\mu_{WP}(p)]^{1/2}$. The operation of dilation leads to a MF that lies outside of the MF of the original fuzzy set (thus, the term *dilation*); both have the same support, and the same membership values where the value of the original MF equals unity or zero.
3. *Artificial Hedges*: Two hedges that are quite useful are the *plus* and *minus* hedges, whose MFs are $\mu_{plus(F)}(x) \equiv [\mu_F(x)]^{1.25}$ and $\mu_{minus(F)}(x) \equiv [\mu_F(x)]^{0.75}$. These artificial hedges provide milder degrees of concentration and dilation than those associated with the concentration and dilation hedges.

The $\equiv$ sign has been used in these hedge MF formulas to convey the fact that their exponents are quite arbitrary; they can be changed depending upon one's interpretation of the hedges,[20] as already noticed in Zadeh (1972), who stated:

> It should be emphasized, however, that these representations are intended mainly to illustrate the approach rather than to provide accurate definitions of the hedges in question. Furthermore it must be understood that our analysis and its conclusions are tentative in nature and may require modification in later work.

The following example illustrates the use of hedge operators.

Example 2.17 (Adapted from Zadeh 1973, p. 35) In conversations, one frequently uses the phrase *highly unlikely*. Here it is shown how to obtain a MF for it. Let X denote a universe of discourse associated with an appropriate quantity related to the notion of *likely*. X is clarified below. Let $\mu_{LIKELY}(x)$ be the MF for the term *likely*. Then,

$$\mu_{HIGHLY-UNLIKELY}(x) = [1 - \mu_{LIKELY}(x)]^{4\times 0.75} \tag{2.60}$$

To obtain (2.60), the hedge *highly* has been interpreted as *minus very very* (which, of course, is subjective) and the fact that *unlikely* is the complement of *likely* has also been used.

From estimation theory (e.g., Edwards (1972); Mendel (1995b)), it is known that *likelihood is proportional to probability*. This fact helps us to establish the universe of discourse, X, as values of probability (the constant of proportionality between probability and likelihood is irrelevant), i.e., $x \in X = [0, 1]$. As a concrete example, assume the following discrete universe of discourse: $X = \{0, 0.1, 0.2, 0.3, \cdots, 0.9, 1\}$. To evaluate (2.60), $\mu_{LIKELY}(x)$ needs to be specified. Based on my perception of the fuzzy set *likely*, the following ad hoc choice is made for $\mu_{LIKELY}(x)$ (your choice may be different):

[20]Because of the uncertainty about the numerical values of the exponents, hedges might be more appropriately modeled within the framework of type-2 fuzzy sets. This is examined in Sect. 7.10.

$$\mu_{LIKELY}(x) \equiv 1/1 + 1/0.9 + 1/0.8 + 0.8/0.7 + 0.6/0.6 + 0.5/0.5 + 0.3/0.4 + 0.2/0.3 \tag{2.61}$$

Recall that the terms not shown have zero MF values. Evaluating (2.60), one finds that

$$\mu_{HIGHLY-UNLIKELY}(x) \approx 1/0 + 1/0.1 + 1/0.2 + 0.5/0.3 + 0.3/0.4 \tag{2.62}$$

Observe, from (2.61) and (2.62), that the MF $\mu_{HIGHLY-UNLIKELY}(x)$ seems to make sense, i.e., it agrees with the notion that something highly unlikely has a very very small chance (i.e., probability) of occurring. Consequently, large values for $\mu_{HIGHLY-UNLIKELY}(x)$ should and indeed do occur for small values of probability, x.

In Schmucker (1984) one finds the following:

> Representing hedges as operators acting upon the representation of the primary terms has both positive and negative implications. On the positive side, it seems very natural and also allows for an easy implementation of the connection of several hedges…. The negative side of representing hedges as operators is that some hedges don't seem to be easily modeled by such an approach. By this we mean that the way people normally use these hedges entails an implementation considerably different and more complex than that of an operator that acts uniformly upon the fuzzy restrictions that represent the various primary terms.

Finally, in Macvicar-Whelen (1978) there are experimental results that indicate the hedge *very* causes a *shift* in the MF rather than a steepening of the MF as is obtained by the concentration operator; hence, their paper calls into question the use of operators to model hedges.

In the rest of this book, unless otherwise indicated, hedges are treated as new linguistic terms.

2.10 Extension Principle

The Extension Principle was introduced in[21] Zadeh (1975) and is an important tool in fuzzy set theory.[22] Heavy use is made of it in later chapters of this book. It extends mathematical relationships between non-fuzzy variables to fuzzy variables. Suppose, for example, that the MF for the fuzzy set *small* is given and the MF for the fuzzy set $(small)^2$ is desired. The Extension Principle determines the MF for

---

[21]According to Klir and Yuan (1995), the Extension Principle was introduced in Zadeh (1975); however, Zadeh (1975, p. 236, footnote 18) states that the Extension Principle is implicit in a result given in Zadeh (1965).

[22]Actually, there are other Extension Principles (e.g., He, et al. 2000; Arabi, et al. 2001), but the one that is described in this section is the most widely used and is the one used in the rest of this book.

$(small)^2$ by making use of the non-fuzzy mathematical relationship $y = x^2$, in which the fuzzy set *small* plays the role of x, and also the MF for *small*.

Suppose one is given a function of a single variable x, $y = f(x)$, where $x \in U$ and $y \in V$. For illustrative purposes, U is assumed to be a discrete universe of discourse, U_d, and

$$A = \sum_{x \in U_d} \mu_A(x)/x \tag{2.63}$$

The Extension Principle states that (Jang et al. 1997) the image of the fuzzy set A under the mapping $f(\cdot)$ can be expressed as a fuzzy set B, i.e.,

$$\begin{aligned} B = f(A) = f\left(\sum_{x \in U_d} \mu_A(x)/x\right) \\ = \mu_A(x_1)/y_1 + \mu_A(x_2)/y_2 + \cdots + \mu_A(x_N)/y_N \equiv \mu_B(y) \end{aligned} \tag{2.64}$$

where $(i = 1, \ldots, N)$ $y_i = f(x_i)$. Since $x = f^{-1}(y)$, where $f^{-1}(y)$ is the inverse of f (i.e., $f[f^{-1}(y)] = y$), another way to express B is by $\mu_B(y) = \mu_A[f^{-1}(y)]$, $y \in V$.

Example 2.18 As a concrete illustration of (2.64), suppose that U_d = {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}, and $A = small = 1/1 + 0.8/2 + 0.6/3 + 0.3/4$; then, $B = (small)^2 = 1/1 + 0.8/4 + 0.6/9 + 0.3/16$.

The version of the Extension Principle given in (2.64) is valid only if the mapping between y and $f(x)$ is one-to-one. It is quite possible that the same value of y can be obtained for different values of x—a many-to-one mapping—in which case (2.64) needs to be modified, e.g., $f(x_1) = f(x_2) = y$, but $x_1 \neq x_2$ and $\mu_A(x_1) \neq \mu_A(x_2)$. To resolve this ambiguity, the larger one of the two membership values is assigned to $\mu_B(y)$. The general modification to (2.64) is Wang (1997):

$$\mu_B(y) = \max_{x \in f^{-1}(y)} \mu_A(x) \quad y \in V \tag{2.65}$$

where $f^{-1}(y)$ denotes the set of all points $x \in U$ such that $f(x) = y$.

Example 2.19 As an illustration of (2.65), suppose that $U_d = \{-3, -2, -1, 0, 1, 2\}$ and fuzzy set A is characterized by the MF values listed in the second column of Table 2.1. Then $\mu_B(y)$, for $y = f(x) = x^4$, is given in the last column of that table.

Table 2.1 Numerical results for Example 2.19

x	$\mu_A(x)$	$y = f(x) = x^4$	$\mu_B(y)$
−3	0.5	81	$\max\{0.5\} = 0.5$
−2	0.6	16	$\max\{0.6, 0.1\} = 0.6$
−1	1.0	1	$\max\{1, 0.4\} = 1$
0	0.9	0	$\max\{0.9\} = 0.9$
1	0.4	1	$\max\{1, 0.4\} = 1$
2	0.1	16	$\max\{0.6, 0.1\} = 0.6$

Observe that there are two pairs of elements of U that map into the same value of y: -2 and 2 map into 16, and -1 and 1 map into 1. In both cases the membership value of y is obtained by taking the maximum of the membership grades of the respective two elements. From the last two columns of Table 2.1, observe that $B = 0.9/0 + 1/1 + 0.6/16 + 0.5/81$.

So far the Extension Principle has been stated just for a mapping of a single variable. Things get a bit more complicated for a function of more than one variable. Suppose, for example, one has a function of two variables x_1 and x_2, i.e., $y = f(x_1, x_2)$, where $x_1 \in X_{d1}$, $x_2 \in X_{d2}$, $y \in V$, X_{d1} and X_{d2} are assumed to be discrete universes of discourse, and:

$$A_1 = \sum_{x_1 \in X_{d1}} \mu_{A_1}(x_1)/x_1 \tag{2.66}$$

and

$$A_2 = \sum_{x_2 \in X_{d2}} \mu_{A_2}(x_2)/x_2 \tag{2.67}$$

Now it is possible for $y = f(x_1, x_2)$ to be many-to-one, just as it was in the single-variable case; so, the Extension Principle for the two-variable case needs to look something like (2.65). The difference between the two- and one-variable cases is that in the latter there is only one MF that can be evaluated for each value of x, whereas in the former there are two MFs that can be evaluated, namely $\mu_{A_1}(x_1)$ and $\mu_{A_2}(x_2)$. In this case, the Extension Principle becomes:

$$\mu_{f(A_1,A_2)}(y) \equiv \mu_B(y) = \begin{cases} \sup\limits_{(x_1,x_2)\in f^{-1}(y)} & \min\{\mu_{A_1}(x_1), \mu_{A_2}(x_2)\} \\ 0 & \text{if } f^{-1}(y) = \emptyset \end{cases} \tag{2.68}$$

where $f^{-1}(y)$ now denotes the set of all points $x_1 \in X_{d1}$ and $x_2 \in X_{d2}$ such that $f(x_1, x_2) = y$. The condition in (2.68) that $\mu_B(y) = 0$ if $f^{-1}(y) = \emptyset$ means that if there are no values of x_1 and x_2 for which a specific value of y can be reached, then the MF value for that specific value of y is set equal to zero. Only those values of y that satisfy $y = f(x_1, x_2)$ can be reached. Note that (Yager 1986) provides a justification of (2.68) based on the sup-star composition.[23]

Example 2.20 [Adapted from Lin and Lee (1996, p. 30)] As an illustration of (2.68), suppose that $X_{d1} = \{-1, 0, 1\}$ and $X_{d2} = \{-2, 2\}$, and fuzzy sets A_1 and A_2

[23]A *plausibility argument* for the Extension Principle is: (1) $y = f(x_1, x_2)$ can be interpreted literally, as: When $x_1 = x_1'$ *and* $x_2 = x_2'$ then $y = f(x_1', x_2')$, where the *and* in this statement is modeled as a conjunction, which explains the use of the minimum in (2.68); and, (2) when $y = f(x_1, x_2)$ is many-to-one, then this can be interpreted as: For $(x_1, x_2) = (x_1^1, x_2^1)$ *or* (x_1^2, x_2^2) *or* … *or* (x_1^m, x_2^m), the same value is obtained for $y = f(x_1, x_2)$, where the *or*'s in this statement are modeled as disjunctions, which explains the use of the maximum (sup) in (2.68).

Table 2.2 Numerical results for Example 2.20

x_1	$\mu_{A_1}(x_1)$	x_2	$\mu_{A_2}(x_2)$	$y = f(x_1, x_2) = x_1^2 + x_2$	$\mu_{f(A_1,A_2)}(y) \equiv \mu_B(y)$
−1	0.5	−2	0.4	−1	$\max\{0.4, 0.4\} = 0.4$
−1	0.5	2	1.0	3	$\max\{0.5, 0.9\} = 0.9$
0	0.1	−2	0.4	−2	$\max\{0.1\} = 0.1$
0	0.1	2	1.0	2	$\max\{0.1\} = 0.1$
1	0.9	−2	0.4	−1	$\max\{0.4, 0.4\} = 0.4$
1	0.9	2	1.0	3	$\max\{0.5, 0.9\} = 0.9$

are characterized by the MFs listed in the second and fourth columns of Table 2.2. Then the MF for the fuzzy set B that is associated with $\mu_{f(A_1,A_2)}(y)$, where $y = f(x_1, x_2) = x_1^2 + x_2$, is given in the last column of that table. The construction of this table first required determining all x_1 and x_2 pairs for which y is defined. These values constitute the Cartesian product of X_{d1} and X_{d2}, $X_{d1} \times X_{d2}$. By evaluating $y = f(x_1, x_2) = x_1^2 + x_2$ at all these values, it is established that $V = \{-2, -1, 2, 3\}$.

There are two ordered pairs (−1, −2) and (1, −2) that map into the same value of y, namely −1, and, there are also two ordered pairs (−1, 2) and (1, 2) that map into the same value of y = 3. It is for these two sets of ordered pairs that the respective maximum membership grades must be taken in (2.68).

The calculations of $\mu_B(y)$ are illustrated next for $y = -1$:

$$\begin{aligned} \mu_B(-1) &= \max[\min\{\mu_{A_1}(-1), \mu_{A_2}(-2)\}, \min\{\mu_{A_1}(1), \mu_{A_2}(-2)\} \\ &= \max[\min(0.5, 0.4), \min(0.9, 0.4)] = 0.4 \end{aligned} \tag{2.69}$$

From the last two columns of Table 2.2, one concludes that $B = 0.1/-2 + 0.4/-1 + 0.1/2 + 0.9/3$.

Finally, the generalization of the Extension Principle in (2.68) from 2 to r variables is considered. The Cartesian product of r arbitrary non-fuzzy sets $X_1, X_2, \ldots, X_r$, denoted by $X_1 \times X_2 \times \cdots \times X_r$, is the non-fuzzy set of all ordered r-tuples $(x_1, x_2, \ldots, x_r)$ such that $x_i \in X_i$ for $i \in \{1, 2, \ldots, r\}$; i.e., Rudin (1966)

$$X_1 \times \cdots \times X_r = \{(x_1, \ldots, x_r) | x_1 \in X_1, \ldots, x_r \in X_r\}$$

Let f be a mapping from $X_1 \times \cdots \times X_r$ to a universe Y such that $y = f(x_1, \ldots, x_r) \in Y$, and $A_1, A_2, \ldots, A_r$ be type-1 fuzzy sets in $X_1, X_2, \ldots, X_r$, respectively. Then, Zadeh's Extension Principle allows one to induce from the r type-1 fuzzy sets $A_1, A_2, \ldots, A_r$ a type-1 fuzzy set B on Y, through f, i.e., $B = f(A_1, A_2, \ldots, A_r)$, such that (see 2.68)[24]

[24]Equation (2.70) assumes that $x_1, \ldots, x_r$ are *non-interactive* (e.g., if $x_1 = a$ and $x_2 = a^2$, then x_1 and x_2 are interactive) or that there is no joint constraint on $x_1, \ldots, x_r$. For a detailed discussion about this, see Zadeh (1975), Appendix B in Karnik and Mendel (1998) and Rajati and Mendel (2013).

$$\mu_B(y) = \begin{cases} \sup\limits_{(x_1,\ldots,x_r)\in f^{-1}(y)} & \min\{\mu_{A_1}(x_1),\ldots,\mu_{A_r}(x_r)\} \\ 0 & \text{if } f^{-1}(y) = \emptyset \end{cases} \tag{2.70}$$

where $f^{-1}(y)$ denotes the set of all points $x_1 \in X_1,\ldots,x_r \in X_r$ such that $y = f(x_1,\ldots,x_r)$.

To implement (2.70), first the values of $x_1,\ldots,x_r$ must be found for which $y = f(x_1,\ldots,x_r)$, after which $\mu_{A_1}(x_1),\ldots,\mu_{A_r}(x_r)$ and $\min\{\mu_{A_1}(x_1),\ldots,\mu_{A_r}(x_r)\}$ are computed at those values. If more than one set of $x_1,\ldots,x_r$ satisfy $y = f(x_1,\ldots,x_r)$, then this procedure is repeated for all of them and the largest of the minima is chosen as the choice for $\mu_B(y)$.

Zadeh defined the Extension Principle using minimum t-norm and maximum t-conorm (for the supremum operation). Other t-norms and t-conorms can be used, as described, e.g., in Dubois and Prade (1980). In this book, only the maximum t-conorm and either the minimum or product t-norms are used. Note that when the minimum in (2.70) is replaced by another t-norm, the sup-min composition is replaced by the sup-star composition.

When one needs to extend an operation of the form $f(x_1,\ldots,x_r)$ to an operation $f(A_1,\ldots,A_r)$ (e.g., $A_1 + \cdots + A_r$) where A_i are type-1 fuzzy sets, the individual operations like multiplication, addition, etc., involved in f, are not extended. Instead, the following definition is used, which derives directly from (2.70) when the maximum operation is used for the union and a general t-norm ($\bigstar$) is used instead of the minimum operation:

$$f(A_1,\ldots,A_r) = \int_{x_1\in X_1} \cdots \int_{x_r\in X_r} \mu_{A_1}(x_1)\bigstar\cdots\bigstar\mu_{A_r}(x_r)/f(x_1,\ldots,x_r) \tag{2.71}$$

For example, if $f(x_1,x_2) = x_1x_2/(x_1+x_2)$, the extension of f to type-1 fuzzy sets A_1 and A_2 is written as:

$$f(A_1,A_2) = \int_{x_1\in X_1}\int_{x_2\in X_2} \mu_{A_1}(x_1)\bigstar\mu_{A_2}(x_2)\Big/\frac{x_1x_2}{x_1+x_2} \tag{2.72}$$

and **not** as $f(A_1,A_2) = A_1 \times A_2/(A_1+A_2)$.

To compute $f(A_1,\ldots,A_r)$ using (2.71), $f(x_1,\ldots,x_r)$ and $\mu_{A_1}(x_1)\bigstar\cdots\bigstar\mu_{A_r}(x_r)$ must be computed for $\forall x_1 \in X_1,\ldots,\forall x_r \in X_r$. It is easy to write a computer program to do this, although sometimes it can be done analytically, as is demonstrated in the next three examples.

Example 2.21 Let F_1, …, and F_n be type-1 interval fuzzy numbers having domains $[l_1,r_1]$, …, and $[l_n,r_n]$, respectively. Then $\sum_{i=1}^n F_i$ is also a type-1 interval fuzzy number whose domain is $\left[\sum_{i=1}^n l_i, \sum_{i=1}^n r_i\right]$. The proof is by mathematical induction.

(a) For F_1 and F_2, using (2.71), the algebraic sum of F_1 and F_2 can be obtained as

$$F_1 + F_2 = \int_{u \in [l_1, r_1]} \int_{w \in [l_2, r_2]} (1 \bigstar 1)/(u + w) \tag{2.73}$$

Observe from (2.73) that: (1) each term in $F_1 + F_2$ is equal to the sum $u + w$ for some $u \in [l_1, r_1]$ and $w \in [l_2, r_2]$, the smallest term being $(l_1 + l_2)$ and the largest being $(r_1 + r_2)$; and (2) since both F_1 and F_2 have continuous domains, $F_1 + F_2$ has a continuous domain; hence, $F_1 + F_2$ is a type-1 interval fuzzy number with domain $[l_1 + l_2, r_1 + r_2]$.

(b) The proof of the general result is straightforward, and is left to the reader (Exercise 2.26).

Example 2.22 Let $F_1, \ldots,$ and F_n be type-1 interval fuzzy numbers having domains $[l_1, r_1], \ldots,$ and $[l_n, r_n]$, respectively. Then, $\sum_{i=1}^{n} a_i F_i + b$ (where each a_i as well as b is a positive real number) is also a type-1 interval fuzzy number whose domain is $\left[\sum_{i=1}^{n} a_i l_i + b, \sum_{i=1}^{n} a_i r_i + b\right]$. The derivation of these results follows.

Consider $F_i = 1/[l_i, r_i]$. Multiplying F_i by the positive real number a$_i$ (expressed as the type-1 fuzzy set $1/a_i$) yields [use (2.71)][25]

$$a_i F_i = \int_{v \in V} 1/(a_i v) \;\; V = [l_i, r_i] \tag{2.74}$$

Adding the positive real number b (expressed as the type-1 fuzzy set $1/b$) to $a_i F_i$ yields (see 2.73)

$$a_i F_i + b = \int_{v \in V} 1/(a_i v + b) \;\; V = [l_i, r_i] \tag{2.75}$$

Substituting $w = a_i v + b$ into (2.75), it follows that:

$$a_i F_i + b = \int_{w \in W} 1/w \;\; W = [a_i l_i + b, a_i r_i + b] \tag{2.76}$$

Consequently, from Example 2.21 and (2.76), the domain of $\sum_{i=1}^{n} a_i F_i + b$ is $\left[\sum_{i=1}^{n} a_i l_i + b, \sum_{i=1}^{n} a_i r_i + b\right]$, Q. E. D.

Note that, when $[l_i, r_i]$ is expressed in terms of its center and spread, as $[c_i - s_i, c_i + s_i]$, for which $l_i = c_i - s_i$ and $r_i = c_i + s_i$, then $\left[\sum_{i=1}^{n} a_i l_i + b, \sum_{i=1}^{n} a_i r_i + b\right] = \left[\sum_{i=1}^{n} a_i c_i + b - \sum_{i=1}^{n} a_i s_i, \sum_{i=1}^{n} a_i c_i + b + \sum_{i=1}^{n} a_i s_i\right]$, which is sometimes a useful alternate way to express the domain of $\sum_{i=1}^{n} a_i F_i + b$.

[25]Note that $1 \bigstar 1 = 1$ regardless of whether the t-norm is minimum or product.

Exercise 2.27 asks the reader to obtain the comparable results when a_i are positive or negative real numbers.

Example 2.23 Given n type-1 Gaussian fuzzy sets $F_1, \ldots, F_n$, with means $m_1, \ldots, m_n$ and standard deviations $\sigma_1, \ldots, \sigma_n$, i.e.,

$$F_i = \int_{x \in X} \exp\left[-\frac{1}{2}\left(\frac{x - m_i}{\sigma_i}\right)^2\right] \Big/ x \quad i = 1, \ldots, n \tag{2.77}$$

The affine combination $\sum_{i=1}^{n} a_i F_i + b$, where a_i and b are crisp constants, is also a type-1 Gaussian fuzzy set with mean $\sum_{i=1}^{n} a_i m_i + b$ and standard deviation Σ', where

$$\Sigma' = \begin{cases} \sqrt{\sum_{i=1}^{n} a_i^2 \sigma_i^2} & \text{if product t-norm is used} \\ \sum_{i=1}^{n} |a_i \sigma_i| & \text{if minimum t-norm is used} \end{cases} \tag{2.78}$$

The proofs of these results, which can be found in Karnik and Mendel (1998, Appendix C.9), use the results from Exercises 2.24 and 2.25.

2.11 α-Cuts[26]

In the first edition of this book there was no material about α-cuts, because both type-1 and interval type-2 rule-based systems did not need them. Beginning in Sect. 6.7.3, it will be seen that α-cuts play a central role for general type-2 fuzzy sets and systems, something that was not known when the first edition of this book was written.

Definition 2.9 (Zadeh 1975) The *α-cut of type-1 fuzzy set A*, denoted A_α, is an interval of real numbers, defined as:

$$A_\alpha = \{x | \mu_A(x) \geq \alpha\} \tag{2.79}$$

where $\alpha \in [0, 1]$.

Example 2.24 An example of an *α-cut* is depicted in Fig. 2.13, and in this example, $A_\alpha = [1.9, 5.5]$. Observe that the α-cut lies on the x-axis.

Example 2.25 Given a specific type-1 fuzzy set A, it is easy to obtain formulas for the end-points of an α-cut, e.g. see Table 2.3. In order to obtain these formulas, such as the ones for the triangular distribution, solve the two equations $l(x) = \alpha$ for the left end-point and $r(x) = \alpha$ for the right end-point of A_α.

[26]If a reader is interested only in type-1 and interval type-2 fuzzy sets and systems, this section, as well as Sects. 2.12 and 2.13, can be omitted.

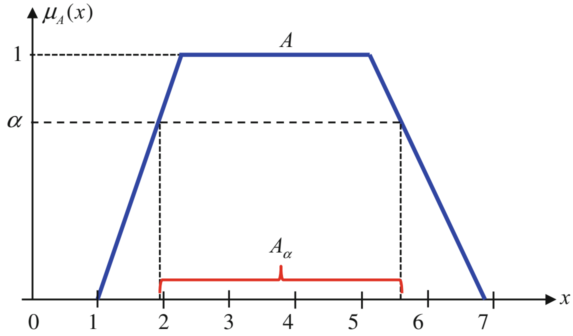
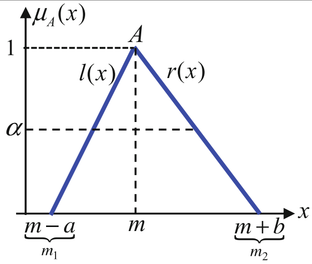
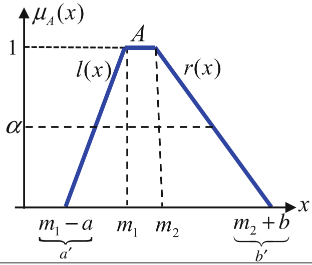


Fig. 2.13 A trapezoidal type-1 fuzzy set and an α-cut (Mendel and Wu 2010 © 2010, IEEE)

Table 2.3 Examples of type-1 fuzzy sets and their α-cut formulas (Mendel and Wu 2010 © 2010, IEEE)

Type-1 fuzzy set	$\alpha-$cut formula
	$A_\alpha = [a_\alpha, b_\alpha] = [m - a(1-\alpha), m + b(1-\alpha)] = [m_1 + (m - m_1)\alpha, m_2 - (m_2 - m)\alpha]$
	$A_\alpha = [a_\alpha, b_\alpha] = [m_1 - a(1-\alpha), m_2 + b(1-\alpha)] = [a' + (m_1 - a')\alpha, b' - (b' - m_2)\alpha]$

Theorem 2.2 *The following set-theoretic properties hold for* α*-cuts*:

$$(A \cap B)_\alpha = A_\alpha \cap B_\alpha \tag{2.80}$$

$$(A \cup B)_\alpha = A_\alpha \cup B_\alpha \tag{2.81}$$

Equations (2.80) and (2.81) state that the α-cut of the intersection (union) of two type-1 fuzzy sets equals the intersection (union) of their α-cuts.

Proof Because the proof of (2.81) is so similar to the proof of (2.80), only the proof (2.80) is provided here; the proof of (2.81) is left as an exercise (Exercise 2.28). This proof is taken from Klir and Yuan (1995, p. 35), and is given for the minimum intersection operator.

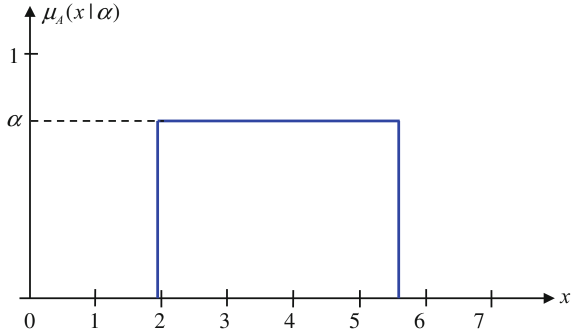


Fig. 2.14 Square-well function $\mu_A(x|\alpha)$ (Mendel and Wu 2010 © 2010, IEEE)

For any $x \in (A \cap B)_\alpha$, it follows from Definition 2.9 that $\mu_{A \cap B}(x) \geq \alpha$; hence, $\min[\mu_A(x), \mu_B(x)] \geq \alpha$. This means $\mu_A(x) \geq \alpha$ and $\mu_B(x) \geq \alpha$ which implies $x \in A_\alpha \cap B_\alpha$, and consequently $(A \cap B)_\alpha \subseteq A_\alpha \cap B_\alpha$.

Conversely, for any $x \in A_\alpha \cap B_\alpha$, $x \in A_\alpha$ and $x \in B_\alpha$. This means, again from Definition 2.9, that $\mu_A(x) \geq \alpha$ and $\mu_B(x) \geq \alpha$; hence, $\min[\mu_A(x), \mu_B(x)] \geq \alpha$ which means $\mu_{A \cap B}(x) \geq \alpha$. This implies $x \in (A \cap B)_\alpha$, and consequently $A_\alpha \cap B_\alpha \subseteq (A \cap B)_\alpha$.

Combining the two parts of this proof, one concludes that $(A \cap B)_\alpha = A_\alpha \cap B_\alpha$.

2.12 Representing Type-1 Fuzzy Sets Using α-Cuts

One of the major roles of α-cuts is their capability to represent a type-1 fuzzy set. In order to do this, first the following *indicator function* is introduced:

$$I_{A_\alpha}(x) = \begin{cases} 1 & x \in A_\alpha \\ 0 & x \notin A_\alpha \end{cases} \tag{2.82}$$

Associated with $I_{A_\alpha}(x)$ is the following *square-well function*:

$$\mu_A(x|\alpha) \equiv \alpha I_{A_\alpha}(x) = \alpha / A_\alpha \tag{2.83}$$

This function, an example of which is depicted in Fig. 2.14, raises the α-cut A_α off of the x-axis to height (level) α.

Theorem 2.3 (Decomposition Theorem) *A type-1 fuzzy set A can be represented as:*

$$\mu_A(x) = \bigcup_{\alpha \in [0,1]} \mu_A(x|\alpha) = \sup_{\alpha \in [0,1]} \{\alpha / A_\alpha\} \; x \in X \tag{2.84}$$

where $\cup$ *is the fuzzy union (i.e., sup over* $[0, 1]$*).*

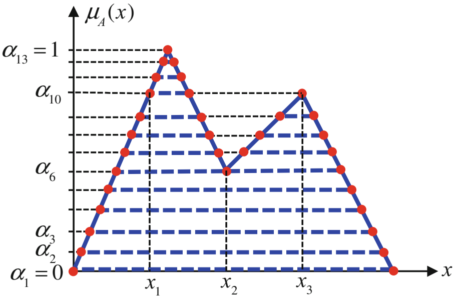


Fig. 2.15 Example to illustrate the Decomposition Theorem when 13 α-cuts are used

This theorem was introduced in Zadeh (1971) and also appears in Zadeh (1975, p. 223), where it is called a *resolution identity*. It is also called a "Decomposition Theorem" because A is decomposed into a collection of square-well functions (i.e., intervals raised to level α) that are then aggregated using the union operation (with respect to α). An example of (2.84) is depicted in Fig. 2.15. In that figure: (1) the blue dashed lines are the α-cuts raised to level α; (2) the red dots show $\mu_A(x)$ computed by using (2.84); and, (3) at x_1, x_2 and x_3, the dashed vertical lines intersect many of the dashed blue lines, but they terminate at their maximum values, the respective red dot, according to (2.84).

Theorem 2.3 holds for continuous and discrete universes of discourse, since (2.84) is valid for both universes, and is valid for convex and non-convex type-1 fuzzy sets. Note that greater resolution is obtained by including more α-cuts, and the calculation of new α-cuts does not affect previously calculated α-cuts. A proof of Theorem 2.3 can be found, e.g., in Klir and Yuan (1995, p. 41) or Wang (1997, p. 369). It is not included herein because, once one understands (2.84), it becomes a rather obvious result.

Example 2.26 (Taken from Mendel, et al. 2014, p. 38) Let $A = 0.2/x_1 + 0.4/x_2 + 0.6/x_3 + 0.8/x_4 + 1/x_5$. Some indicator functions for A are:

$$
\begin{aligned}
I_{A_{0.2}}(x) &= 1/x_1 + 1/x_2 + 1/x_3 + 1/x_4 + 1/x_5 \\
I_{A_{0.4}}(x) &= 0/x_1 + 1/x_2 + 1/x_3 + 1/x_4 + 1/x_5 \\
I_{A_{0.6}}(x) &= 0/x_1 + 0/x_2 + 1/x_3 + 1/x_4 + 1/x_5 \\
I_{A_{0.8}}(x) &= 0/x_1 + 0/x_2 + 0/x_3 + 1/x_4 + 1/x_5 \\
I_{A_{1.0}}(x) &= 0/x_1 + 0/x_2 + 0/x_3 + 0/x_4 + 1/x_5
\end{aligned} \tag{2.85}
$$

Their associated square-well functions are:

$$
\begin{aligned}
\mu_A(x|0.2) &= 0.2/x_1 + 0.2/x_2 + 0.2/x_3 + 0.2/x_4 + 0.2/x_5 \\
\mu_A(x|0.4) &= 0/x_1 + 0.4/x_2 + 0.4/x_3 + 0.4/x_4 + 0.4/x_5 \\
\mu_A(x|0.6) &= 0/x_1 + 0/x_2 + 0.6/x_3 + 0.6/x_4 + 0.6/x_5 \\
\mu_A(x|0.8) &= 0/x_1 + 0/x_2 + 0/x_3 + 0.8/x_4 + 0.8/x_5 \\
\mu_A(x|1.0) &= 0/x_1 + 0/x_2 + 0/x_3 + 0/x_4 + 1/x_5
\end{aligned} \tag{2.86}
$$

Applying (2.84) to these functions, it follows that:

$$A = \mu_A(x|0.2) \cup \mu_A(x|0.4) \cup \mu_A(x|0.6) \cup \mu_A(x|0.8) \cup \mu_A(x|1.0) \tag{2.87}$$

When performing these unions, focus on a specific domain point, e.g. $x = x_4$, for which

$$\mu_A(x_4) = \max(0, 0.2, 0.4, 0.6, 0.8)/x_4 = 0.8/x_4 \tag{2.88}$$

Performing these unions for the five domain points, whose MFs are non-zero, it is straightforward to recover $A = 0.2/x_1 + 0.4/x_2 + 0.6/x_3 + 0.8/x_4 + 1/x_5$.

Example 2.27 For a *convex type-1 fuzzy set*, such as the ones in Table 2.3, $A_\alpha = [a_\alpha, b_\alpha]$ ($\alpha \in [0,1]$), and (2.84) can be expressed as:

$$\mu_A(x) = \sup_{\alpha\in[0,1]} \{\alpha/[a_\alpha, b_\alpha]\} \; x \in X \tag{2.89}$$

The following is a corollary to Theorems 2.2 and 2.3:

Corollary 2.1 *The intersection and union of type-1 fuzzy sets A and B can be computed by using their α-cuts, as follows:*

$$\mu_{A\cap B}(x) = \bigcup_{\alpha\in[0,1]} \alpha/(A_\alpha \cap B_\alpha) \tag{2.90}$$

$$\mu_{A\cup B}(x) = \bigcup_{\alpha\in[0,1]} \alpha/(A_\alpha \cup B_\alpha) \tag{2.91}$$

Proof From Theorem 2.3, it follows that:

$$\mu_{A\cap B}(x) = \bigcup_{\alpha\in[0,1]} \mu_{A\cap B}(x|\alpha) = \bigcup_{\alpha\in[0,1]} \alpha/(A\cap B)_\alpha \tag{2.92}$$

Applying (2.80) to (2.92), it follows that:

$$\mu_{A\cap B}(x) = \bigcup_{\alpha\in[0,1]} \alpha/(A_\alpha \cap B_\alpha) \tag{2.93}$$

which is (2.90). Because the proof of (2.91) is so similar to the proof of (2.90) it is not provided here.

Equation (2.90) is also true when $\cap$ is replaced by t-norm symbol $\bigstar$. An important feature of (2.90) and (2.91) is that, since A_α and B_α are intervals (or multiple intervals) of real numbers, $A_\alpha \cup B_\alpha$ and $A_\alpha \cap B_\alpha$ are easily computed.

Example 2.28 Here (2.91) is applied to the two type-1 fuzzy sets A and B that are depicted in Fig. 2.16a, to verify that the correct answer is obtained for $\mu_{A\cup B}(x)$. The union of A and B, computed as $\max(\mu_A(x), \mu_B(x))$ is depicted as the red curve

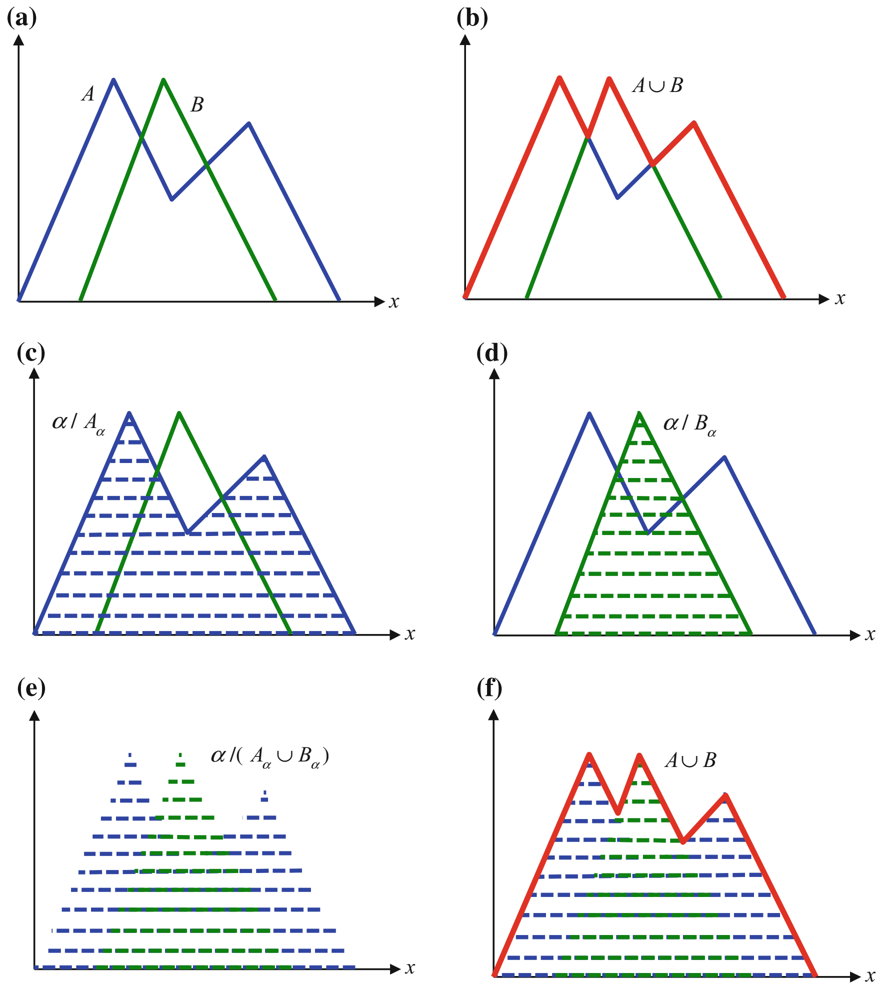


Fig. 2.16 **a** Type-1 fuzzy sets A and B, **b** $A \cup B = \max\{\mu_A(x), \mu_B(x)\}$, **c** α-cuts of A raised to level α **d** α-cuts of B raised to level α, **e** $A_\alpha \cup B_\alpha$ raised to level α, and **f** $A \cup B = \bigcup_{\alpha \in [0,1]} \alpha/(A_\alpha \cup B_\alpha)$

in Fig. 2.16b. Some α-cuts that are raised to level α for A and B, are depicted (as the dashed lines) in Fig. 2.16c, d, respectively. By superimposing all of these dashed lines for α/A_α and α/B_α one obtains[27] $\alpha/(A_\alpha \cup B_\alpha)$ in Fig. 2.16e. The envelope of all of the $\alpha/(A_\alpha \cup B_\alpha)$ in Fig. 2.16e provides the red curve in Fig. 2.16f, which is in agreement with the red curve in Fig. 2.16b. Each point on the red envelope can be obtained by going to a specific value of x, drawing a

[27]Recall that $A_\alpha \cup B_\alpha$ is a set of real numbers that includes all elements in either A_α or B_α.

vertical line up from it, and choosing the height of that line as the value of the highest dashed horizontal line that intersects it.

2.13 Functions of Type-1 Fuzzy Sets Computed by Using α-Cuts

Recall[28] (Sect. 2.10) that the Extension Principle states that when the function $y = f(x_1, x_2, \ldots, x_r)$ is applied to type-1 fuzzy sets X_i $(i = 1, \ldots, r)$ the result is another type-1 fuzzy set, Y, whose MF $\mu_Y(y)$ is given by (2.70). Because $\mu_Y(y)$ is a type-1 fuzzy set, it can, therefore, be expressed in terms of its α-cuts as follows (see (2.79), (2.82)–(2.84), where Y_α plays the role of A_α):

$$Y_\alpha = \{y | \mu_Y(y) \geq \alpha\} \tag{2.94}$$

$$I_{Y_\alpha}(x) = \begin{cases} 1 & y \in Y_\alpha \\ 0 & y \notin Y_\alpha \end{cases} \tag{2.95}$$

$$\mu_Y(y|\alpha) \equiv \alpha I_{Y_\alpha}(y) = \alpha / Y_\alpha \tag{2.96}$$

$$\mu_Y(y) = \bigcup_{\alpha \in [0,1]} \mu_Y(y|\alpha) = \sup_{\alpha \in [0,1]} \{\alpha / Y_\alpha\} \quad y \in D_Y \tag{2.97}$$

In order to implement (2.95)–(2.97), a method is needed to compute Y_α, and this is provided in the following:

Theorem 2.4 (α-Cuts Decomposition Theorem[29]) *Let* $Y = f(X_1, X_2, \ldots, X_r)$ *be an arbitrary (crisp) function, where* X_i $(i = 1, \ldots, r)$ *is a type-1 fuzzy set whose domain is* D_{X_i} *and* α*-cut is* $(X_i)_\alpha$*. Then, under the Extension Principle*:

$$Y_\alpha = f\big((X_1)_\alpha, \ldots, (X_2)_\alpha\big) \tag{2.98}$$

and the height of Y *equals the minimum height of all* X_i*.*

Equation (2.98) shows that the α-cut of a function of type-1 fuzzy sets equals that function applied to the α-cuts of those type-1 fuzzy sets. Theorem 2.4 does not address how to compute $f\big((X_1)_\alpha, \ldots, (X_2)_\alpha\big)$. Example 2.30 below shows how to do this for a specific nonlinear function, and, when this theorem is used in later chapters of this book for other functions, explanations will be given for how to

[28] Much of the material in this section (up to Example 2.29) is taken from Mendel and Wu (2010, Sect. 5A.2, © IEEE 2010).

[29] The statement of this theorem is adapted from Klir and Yuan (1995, Theorem 2.9) and is taken from Mendel and Wu (2010). Zadeh (1975) states this result without a proof for it. Nguyen (1978) seems to be the first to provide necessary and sufficient conditions for (2.98) to hold.

compute $f((X_1)_\alpha, \ldots, (X_2)_\alpha)$ for those functions. It is no exaggeration to say that this theorem is now vitally important for general type-2 fuzzy systems.

Proof For all $y \in D_Y$, from (2.94) it follows that[30]

$$y \in Y_\alpha \Leftrightarrow \mu_Y(y) \geq \alpha \tag{2.99}$$

Under the Extension Principle in (2.70),

$$\mu_Y(y) \geq \alpha \Leftrightarrow \sup_{(x_1,\ldots,x_r)|y=f(x_1,\ldots,x_r)} \min\{\mu_{X_1}(x_1), \ldots, \mu_{X_r}(x_r)\} \geq \alpha \tag{2.100}$$

It follows that:

$$\begin{aligned}
&\sup_{(x_1,\ldots,x_r)|y=f(x_1,\ldots,x_r)} \min\{\mu_{X_1}(x_1), \ldots, \mu_{X_r}(x_r)\} \geq \alpha \\
&\Leftrightarrow (\exists x_{10} \in D_{X_1} \text{ and } \cdots \text{ and } x_{r0} \in D_{X_r}) \text{ such that} \\
&\quad (y = f(x_{10}, \ldots, x_{r0}) \text{ and } \min\{\mu_{X_1}(x_{10}), \ldots, \mu_{X_r}(x_{r0})\} \geq \alpha) \\
&\Leftrightarrow (\exists x_{10} \in D_{X_1} \text{ and } \cdots \text{ and } x_{r0} \in D_{X_r}) \text{ such that} \\
&\quad (y = f(x_{10}, \ldots, x_{r0}) \text{ and } [\mu_{X_1}(x_{10}) \geq \alpha \text{ and } \cdots \text{ and } \mu_{X_r}(x_{r0}) \geq \alpha]) \\
&\Leftrightarrow (\exists x_{10} \in D_{X_1} \text{ and } \cdots \text{ and } x_{r0} \in D_{X_r}) \text{ such that} \\
&\quad \big(y = f(x_{10}, \ldots, x_{r0}) \text{ and } \big[x_{10} \in (X_1)_\alpha \text{ and } \cdots \text{ and } x_{r0} \in (X_r)_\alpha\big]\big) \\
&\Leftrightarrow y \in f((X_1)_\alpha, \ldots, (X_2)_\alpha)
\end{aligned} \tag{2.101}$$

Hence, from the last line of (2.101) and (2.100),

$$\mu_Y(y) \geq \alpha \Leftrightarrow y \in f((X_1)_\alpha, \ldots, (X_2)_\alpha) \tag{2.102}$$

which is (2.98). Because the right-hand side of (2.100) (read from right to the left) indicates that α cannot exceed the minimum height of all $\mu_{X_i}(x_i)$ (otherwise there is no α-cut on one or more X_i), the height of Y must equal the minimum height of all X_i.

Example 2.29 Let[31] $A = [a, b, c]$ and $B = [p, q, r]$ be two triangle type-1 fuzzy numbers whose MFs are:

$$\mu_A(x) = \begin{cases} \frac{x-a}{b-a} & a \leq x \leq b \\ \frac{c-x}{c-b} & b \leq x \leq c \end{cases} \tag{2.103}$$

[30]This proof is similar to the one that is given for Theorem 2.9 in Klir and Yuan (1995), where it is only provided for a function of a single variable. Even so, our proof of Theorem 2.4 follows the proof of their Theorem 2.9 very closely; however, their theorem does not explain how sub-normal type-1 fuzzy sets should be handled. Such sub-normal type-1 fuzzy sets are quite common in type-2 fuzzy sets because many kinds of lower MFs (see Chap. 6) are sub-normal.

[31]This example is adapted from Dutta, et al. (2011).

$$\mu_B(x) = \begin{cases} \frac{x-p}{q-p} & p \le x \le q \\ \frac{r-x}{r-q} & q \le x \le r \end{cases} \tag{2.104}$$

Then the α-cuts of A and B are (use the first row of Table 2.3 in which $m_1 = a$, $m = b$ and $m_2 = c$):

$$A_\alpha = [(b-a)\alpha + a, c - (c-b)\alpha] \tag{2.105}$$

$$B_\alpha = [(q-p)\alpha + p, r - (r-q)\alpha] \tag{2.106}$$

Here the MF of $A + B$, the sum of two type-1 fuzzy numbers, is computed. To begin, the $\alpha -$ cuts of A and B are added using interval arithmetic, namely

$$[r+s] + [t+u] = [r+t, s+u] \tag{2.107}$$

Consequently:

$$\begin{aligned} A_\alpha + B_\alpha &= [(b-a)\alpha + a, c - (c-b)\alpha] + [(q-p)\alpha + p, r - (r-q)\alpha] \\ &= [a + p + (b - a + q - p)\alpha, c + r - (c - b + r - q)\alpha] \end{aligned} \tag{2.108}$$

To find $\mu_{A+B}(x)$ equate to x both the first and second components in (2.108) [note that this is the reverse of what was done to obtain the α-cuts in (2.105) and (2.106)]:

$$x = a + p + (b - a + q - p)\alpha \tag{2.109a}$$

$$x = c + r - (c - b + r - q)\alpha \tag{2.109b}$$

Next, express α in terms of x and then set $\alpha = 0$ and $\alpha = 1$ in (2.109a, 2.109b) to obtain a respective value of α together with the respective domain of x, as:

$$\alpha = \frac{x - (a+p)}{(b+q) - (a+p)}, \quad (a+p) \le x \le (b+q) \tag{2.110a}$$

$$\alpha = \frac{(c+r) - x}{(c+r) - (b+q)}, \quad (b+q) \le x \le (c+r) \tag{2.110b}$$

Because α is the MF grade of $A + B$ (this is a crucial observation) it follows that:

$$\mu_{A+B}(x) = \begin{cases} \frac{x-(a+p)}{(b+q)-(a+p)} & (a+p) \le x \le (b+q) \\ \frac{(c+r)-x}{(c+r)-(b+q)} & (b+q) \le x \le (c+r) \end{cases} \tag{2.111}$$

Observe that $A + B$ is also a type-1 fuzzy number, i.e. $A + B = [(a+p), (b+q), (c+r)]$.

A very interesting exposition about interval computing (e.g., using α - cuts) and fuzzy sets is Kreinovich (2008).

2.14 Multivariable MFs and Cartesian Products

Most discussions in this chapter have been for type-1 fuzzy sets that depend on only one variable. This section describes how to characterize type-1 fuzzy sets that depend on up to p variables, $x_1, x_2, \ldots, x_p$.

For two variables, x_1 and x_2, type-1 fuzzy set A is defined on the Cartesian product $X_1 \times X_2$, i.e.,

$$\begin{aligned} A &= \{((x_1, x_2), \mu_A(x_1, x_2)) | x_1 \in X_1, x_2 \in X_2\} \\ &= \{((x_1, x_2), \mu_A(x_1, x_2)) | (x_1, x_2) \in X_1 \times X_2\} \end{aligned} \tag{2.112}$$

where $\mu_A(x_1, x_2)$ is a general function of x_1 and x_2. When $X_1 \times X_2$ is continuous, then A can also be written as

$$A = \int_{x_1 \in X_1} \int_{x_2 \in X_2} \mu_A(x_1, x_2)/(x_1, x_2) \tag{2.113}$$

or, if $X_1 \times X_2$ is discrete, $X_{1d} \times X_{2d}$, then A can be written as

$$A = \sum\nolimits_{x_1 \in X_{1d}} \sum\nolimits_{x_2 \in X_{2d}} \mu_A(x_1, x_2)/(x_1, x_2) \tag{2.114}$$

When the MF $\mu_A(x_1, x_2)$ is *separable*, which occurs when x_1 and x_2 do not interact with one another, then it is expressed in terms of $\mu_{A_1}(x_1)$ and $\mu_{A_2}(x_2)$, as

$$\mu_A(x_1, x_2) = \mu_{A_1}(x_1) \bigstar \mu_{A_2}(x_2) \tag{2.115}$$

where $\star$ denotes a t-norm such as minimum or product. In this book only separable MFs are used.

The extensions of these two-variable results to more than two variables is straightforward, e.g., for p variables, when the MF $\mu_A(x_1, x_2, \ldots, x_p)$ is separable, then

$$\mu_A(x_1, x_2, \ldots, x_p) = \mu_{A_1}(x_1) \bigstar \mu_{A_2}(x_2) \bigstar \cdots \bigstar \mu_{A_p}(x_p) \tag{2.116}$$

where $x_1 \in X_1, x_2 \in X_2, \ldots, x_p \in X_p$, which can be interpreted as the Cartesian product of the type-1 fuzzy sets $A_1, A_2, \ldots, A_p$ in the product space $X_1 \times X_2 \times \cdots \times X_p$.

Equation (2.116) is frequently written as

$$\mu_A(x_1, x_2, \ldots, x_p) = \mu_{X_1}(x_1) \bigstar \mu_{X_2}(x_2) \bigstar \cdots \bigstar \mu_{X_p}(x_p) \tag{2.117}$$

Using the notation of (2.117), X_i plays a double role as the label of the fuzzy set and as the universe of discourse for x_i. Usually, this does not cause any confusion. (2.117) is widely used in Chap. 3.

2.15 Crisp Logic

According to the *Encyclopedia Britannica*, "Logic is the study of propositions and their use in argumentation." According to *Webster's Dictionary of the English Language*, "logic is the science of formal reasoning, using principles of valid inference," and " logic is the science whose chief end is to ascertain the principles on which all valid reasoning depends, and which may be applied to test the legitimacy of every conclusion that is drawn from premises." Although multi-valued logic exists, most of us are most familiar with two-valued (dual-valued) logic in which a proposition is either *true* or *false*. With the advent of fuzzy logic, this kind of logic is also referred to as *crisp logic*, which was first systematized by Aristotle thousands of years ago, in ancient Athens.

From Fig. 1.2, observe that one of the major components of a fuzzy system is *Rules*. In this book, rules will be expressed as logical implications, i.e., in the forms of IF–THEN statements, e.g.,

$$\text{IF } x \text{ is } A, \text{THEN } y \text{ is } B, \text{where } x \in X \text{ and } y \in Y$$

A rule represents a special kind of *relation* between A and B; its MF is denoted $\mu_{A \to B}(x, y)$. What is a proper and appropriate choice for this MF? Nothing that has been presented so far helps us to answer this question, because an implication resides within a branch of mathematics known as logic, and so far only set theory has been discussed. Fortunately, as stated in Klir and Folger (1988, p. 24):

> It is well established that propositional logic is isomorphic to set theory under the appropriate correspondence between components of these two mathematical systems. Furthermore, both of these systems are isomorphic to a Boolean algebra, which is a mathematical system defined by abstract (interpretation-free) entities and their axiomatic properties. … The isomorphisms between Boolean algebra, set theory, and propositional logic guarantee that every theorem in any one of these theories has a counterpart in each of the other two theories. … These isomorphisms allow us, in effect, to cover all these theories by developing only one of them.

Consequently, not a lot of time will be spent reviewing crisp logic; but, some time must be spent on it, especially on the concept of implication, in order to reach the comparable concept in fuzzy logic.

Rules are a form of propositions.[32] A *proposition* is an ordinary statement involving terms that have been defined, e.g., "The damping ratio is low." Consequently, one could have the following rule: "IF the damping ratio is low, THEN the system's impulse response oscillates a long time before it dies out." In traditional propositional logic, a proposition must be meaningful to call it "true" or "false," whether or not one knows which of these terms properly applies.

Logical reasoning is the process of combining given propositions into other propositions, and then doing this over and over again. Propositions can be combined in many ways, all of which are derived from three fundamental operations: *conjunction* (denoted $p \wedge q$), where one asserts the simultaneous truth of two separate propositions p and q (e.g., damping ratio is low and bandwidth is large); *disjunction* (denoted $p \vee q$), where one asserts the truth of either or both of two separate propositions (e.g., I will design an analog filter or I will design a digital filter); and, *implication* (denoted $p \rightarrow q$), which usually takes the form of an IF–THEN rule, an example of which has been given in the previous paragraph. The IF part of an implication is called its *antecedent* whereas the THEN part is called its *consequent*.

In addition to generating propositions using conjunction, disjunction, or implication, a new proposition can be obtained from a given one by prefixing the clause "it is false that …". This is the operation of *negation* (denoted $\sim p$). Additionally, $p \leftrightarrow q$ is the *equivalence* relation; it means that p and q are both true or false.

In traditional propositional logic an implication is said to be *true* if one of the following holds: (1) antecedent is true, consequent is true, (2) antecedent is false, consequent is false, and (3) antecedent is false, consequent is true. The implication is called *false* when (4) antecedent is true, consequent is false. Situation (1) is the familiar one of common experience. Situation (2) is also reasonable, for if one starts from a false assumption one expects to reach a false conclusion, however, intuition is not always reliable. One may reason correctly from a false antecedent to a true consequent (e.g., IF 1 = 2, THEN 3 = 3; note that 1 = 2 is false, but, adding 2 = 1 to this false statement, lets one correctly conclude that 3 = 3); hence, a false antecedent can lead to a consequent which is either true or false, and thus both situations (2) and (3) are allowed in traditional propositional logic. Finally, situation (4) is in accord with our intuition, for an implication is clearly false if a true antecedent leads to a false consequent.

A logical structure is constructed by applying the aforementioned five operations to propositions. The objective of a logical structure is to determine the truth or falsehood of all propositions that can be stated in the terminology of this structure.

A *truth table* is very convenient for showing relationships between several propositions. The fundamental truth tables for conjunction, disjunction, implication, equivalence, and negation are collected together in Table 2.4, in which symbol T means that the corresponding proposition is true, and symbol F means that it is false.

The fundamental axioms of traditional propositional logic are: (1) every proposition is either true or false, but not both true or false; (2) the expressions

[32]Much of the rest of this section is paraphrased from Allendoerfer and Oakley (1955).

Table 2.4 Truth table for five operations that are frequently applied to propositions (Mendel 1995a © 1995, IEEE)

p	q	$p \wedge q$	$p \vee q$	$p \to q$	$p \leftrightarrow q$	$\sim p$
T	T	T	T	T	T	F
T	F	F	T	F	F	F
F	T	F	T	T	F	T
F	F	F	F	T	T	T

given by defined terms are propositions; and (3) the truth Table 2.4 for conjunction, disjunction, implication, equivalence, and negation. Using truth tables, one can derive many interpretations of the preceding operations and can also prove relationships about them.

A *tautology* is a proposition formed by combining other propositions (*p, q, r, ...*) which is true regardless of the truth or falsehood of *p, q, r, ...* . The most important tautology for our work is:

$$(p \to q) \leftrightarrow \sim [p \wedge (\sim q)] \tag{2.118}$$

A proof of this tautology, using truth tables, is given in Table 2.5. Observe that the entries in the two columns $p \to q$ and $\sim [p \wedge (\sim q)]$ are identical, which proves the tautology. This tautology can also be expressed as

$$(p \to q) \leftrightarrow (\sim p) \vee q \tag{2.119}$$

the truth of which is also demonstrated in Table 2.5. The importance of these tautologies is that they let one express the MF for $p \to q$ in terms of MFs of either propositions p and $\sim q$ or $\sim p$ and q, which is very important for transitioning from crisp to fuzzy logic.

Some of the most important mathematical equivalences between logic and set theory are:

Logic	Set theory
$\wedge$	$\cap$
$\vee$	$\cup$
$\sim$	$\overline{()}$

Table 2.5 Proofs of $(p \to q) \leftrightarrow \sim [p \wedge (\sim q)]$ and $(p \to q) \leftrightarrow (\sim p) \vee q$ (Mendel 1995a © 1995, IEEE)

p	q	$p \to q$	$\sim q$	$p \wedge (\sim q)$	$\sim [p \wedge (\sim q)]$	$\sim p$	$(\sim p) \vee q$
T	T	T	F	F	T	F	T
T	F	F	T	T	F	F	F
F	T	T	F	F	T	T	T
F	F	T	T	F	T	T	T

Table 2.6 Validations of (2.120) and (2.121) (Mendel 1995a © 1995, IEEE)

$\mu_p(x)$	$\mu_q(y)$	$1-\mu_p(x)$	$1-\mu_q(y)$	$1-\min[\mu_p(x), 1-\mu_q(y)]$	$\max[1-\mu_p(x), \mu_q(y)]$
1	1	0	0	1	1
1	0	0	1	0	0
0	1	1	0	1	1
0	0	1	1	1	1

Additionally, as mentioned earlier, there is a correspondence between elementary logic and Boolean Algebra (0, 1). Any statement that is true in one system becomes a true statement in the other, simply by carrying through the following changes in notation:

Logic	Boolean algebra (0, 1)
T	1
F	0
$\wedge$	$\times$
$\vee$	$+$
$\sim$	$'$
$\leftrightarrow$	$=$
$p, q, r, \ldots$	$a, b, c, \ldots$

In this list, $'$ stands for complement, and *a, b, c,* ... are arbitrary elements of the two-element set {0, 1}.

Using the facts that $(p \rightarrow q) \leftrightarrow \sim[p \wedge (\sim q)]$ and $(p \rightarrow q) \leftrightarrow (\sim p) \vee q$, and the equivalence between logic and set theory, two MFs can be obtained for $p \rightarrow q$. The first of these tautologies lets us show that

$$\mu_{p \rightarrow q}(x, y) = 1 - \mu_{p \cap \bar{q}}(x, y) = 1 - \min\left[\mu_p(x), 1 - \mu_q(y)\right] \tag{2.120}$$

and the second of these tautologies lets us show that[33]

$$\text{Kleene-Dienes} : \mu^{KD}_{p \rightarrow q}(x, y) = \mu_{\bar{p} \cup q}(x, y) = \max\left[1 - \mu_p(x), \mu_q(y)\right] \tag{2.121}$$

To validate the truth of these two MFs, construct a Boolean truth table, such as the one in Table 2.6. Observe that the entries in the last two columns agree with the entries in Table 2.4 for $p \rightarrow q$, where the logical T and F are interchanged with Boolean 1 and 0, respectively.

[33]A named implication MF (e.g., Kleene-Dienes) refers to the person or persons attributed to it in Klir and Yuan (1995, Table 11.1, p. 309).

The implication MFs given in (2.120) and (2.121) are by no means the only ones that give agreement with $p \to q$. Two others are shown here [see Klir and Yuan (1995, Table 11.1) for many more]:

$$\text{Reichenbach}: \mu^{R}_{p\to q}(x,y) = 1 - \mu_p(x)[1-\mu_q(y)] \tag{2.122}$$

and

$$\text{Lukasiewicz}: \mu^{L}_{p\to q}(x,y) = \min[1, 1-\mu_p(x)+\mu_q(y)] \tag{2.123}$$

The MF in (2.122) is similar to the one in (2.120), except that a *product* operation is used for conjunction instead of the minimum operation.

In logic, an *inference rule* is a logical form consisting of a function that takes premises, analyzes their syntax, and returns a conclusion. In traditional propositional (crisp) logic there are two very important inference rules, *Modus Ponens* and *Modus Tollens*:

Modus Ponens:

Premise: x is A
Implication: IF x is A THEN y is B
Consequence: y is B.

Modus Ponens is associated with the implication "A implies B" ($A \to B$). In terms of propositions p and q, Modus Ponens is expressed as $(p \wedge (p \to q)) \to q$.

Modus Tollens:

Premise: y is not B
Implication: IF x is A THEN y is B
Consequence: x is not A.

In terms of propositions p and q, Modus Tollens is expressed as $(\overline{q} \wedge (p \to q)) \to \overline{p}$.

Whereas Modus Ponens plays a central role in engineering applications of logic, due in large part to cause and effect, Modus Tollens does not seem to have yet played much of a role.

2.16 From Crisp Logic to Fuzzy Logic

Fuzzy logic is a type of logic that includes more than just true or false values. It is the logic that deals with situations where one cannot give a clear yes/no (true/false) answer. In fuzzy logic, propositions are represented with *degrees of truthfulness or falsehood*, i.e., fuzzy logic uses a continuous range of truth values in the interval [0, 1] rather than just true or false values. In fuzzy logic, Aristotle's laws of the Excluded Middle and Contradiction are usually broken.

Fuzzy logic begins by borrowing notions from crisp logic, just as fuzzy set theory borrows from crisp set theory. As in our extension of crisp set theory to fuzzy set theory, our extension of crisp logic to fuzzy logic is made by replacing the bivalent MFs of crisp logic with their fuzzy MFs. That is all there is to it; hence, the IF–THEN statement "IF x is A, THEN y is B," where $x \in X$ and $y \in Y$, has a MF $\mu_{A\to B}(x, y)$ where $\mu_{A\to B}(x, y) \in [0, 1]$. Note that $\mu_{A\to B}(x, y)$ measures the degree of truth of the implication relation between x and y, and it resides in the Cartesian product space $X \times Y$. Examples of such MFs are:

$$\mu_{A\to B}(x, y) = 1 - \min[\mu_A(x), 1 - \mu_B(y)] \tag{2.124}$$

$$\mu^{KD}_{A\to B}(x, y) = \max[1 - \mu_A(x), \mu_B(y)] \tag{2.125}$$

and

$$\mu^{R}_{A\to B}(x, y) = 1 - \mu_A(x)(1 - \mu_B(y)) \tag{2.126}$$

which, of course, are fuzzy versions of (2.120)–(2.122), respectively.

In fuzzy logic, Modus Ponens is extended to *Generalized Modus Ponens*:

Premise: x is A^*
Implication: IF x is A THEN y is B
Consequence: y is B^*.

Compare Modus Ponens and Generalized Modus Ponens to see their subtle differences, namely, in the latter, fuzzy set A^* is not necessarily the same as rule antecedent fuzzy set A, and fuzzy set B^* is not necessarily the same as rule consequent B.

Example 2.30 (Mendel 1995a) Consider the rule "IF a man is short, THEN he will not make a very good professional basketball player." Here fuzzy set A is *short man*, and fuzzy set B is *not a very good professional basketball player*. Given Premise 1, as "This man is under five feet tall," A^* is the fuzzy set *man under five feet tall*. Clearly $A^* \neq A$; but, A^* is similar to A. The following consequence is now drawn: "He will make a poor professional basketball player." Here B^* is the fuzzy set *poor professional basketball player*, and $B^* \neq B$, although B^* is indeed similar to B.

In crisp logic a rule will be fired only if the premise is exactly the same as the antecedent of the rule, and the result of such rule firing is the rule's actual consequent. In fuzzy logic, on the other hand, a rule is fired so long as there is a non-zero degree of similarity between the premise and the antecedent of the rule, and the result of such rule firing is a consequent that has a non-zero degree of similarity to the rule's consequent.

Generalized Modus Ponens is a fuzzy composition where the first fuzzy relation is merely the fuzzy set A^*. Consequently, using (2.56), $\mu_{B^*}(y)$ is obtained from the following *sup-star composition* (also called the *compositional rule of inference*):

$$\mu_{B^*}(y) = \sup_{x \in X}[\mu_{A^*}(x) \bigstar \mu_{A \to B}(x, y)] \;\; y \in Y \tag{2.127}$$

To help us understand the meaning of (2.127), some examples will be considered. In all these examples the fuzzy set A^* is assumed to be a fuzzy singleton, i.e.,

$$\mu_{A^*}(x) = \begin{cases} 1 & x = x' \\ 0 & x \neq x' \text{ and } \forall x \in X \end{cases} \tag{2.128}$$

In Chap. 3 this will be called a *singleton fuzzifier* and one will learn why it is so popular. For the singleton fuzzifier, (2.127) becomes:

$$\begin{aligned} \mu_{B^*}(y) &= \sup_{x \in X}[\mu_{A^*}(x) \bigstar \mu_{A \to B}(x, y)] \\ &= \sup[\mu_{A \to B}(x', y), 0] = \mu_{A \to B}(x', y) \;\; y \in Y \end{aligned} \tag{2.129}$$

Eq. (2.129) is true regardless of whether one uses minimum or product for $\bigstar$. Observe that for the singleton fuzzifier the supremum operation is very easy to evaluate, because $\mu_{A^*}(x)$ is non-zero at only one point, x'.

Example 2.31 To begin, the result of using (2.129) for $\mu_{A \to B}(x', y)$ in (2.120) is examined, i.e.,

$$\mu_{B^*}(y) = \mu_{A \to B}(x', y) = 1 - \min[\mu_A(x'), 1 - \mu_B(y)] \;\; y \in Y \tag{2.130}$$

A graphical interpretation of this result is given in Fig. 2.17. Starting with $\mu_B(y)$ in (a), $1 - \mu_B(y)$ is computed as shown in (b), and, for the given level of $\mu_A(x')$ shown in (b), $\min[\mu_A(x'), 1 - \mu_B(y)]$, also shown in (b), is then constructed. Note that the level shown for $\mu_A(x')$ in (b) was chosen arbitrarily (by the author), where $\mu_A(x') \in [0, 1]$. Finally, $1 - \min[\mu_A(x'), 1 - \mu_B(y)]$ is constructed, as shown in (c).

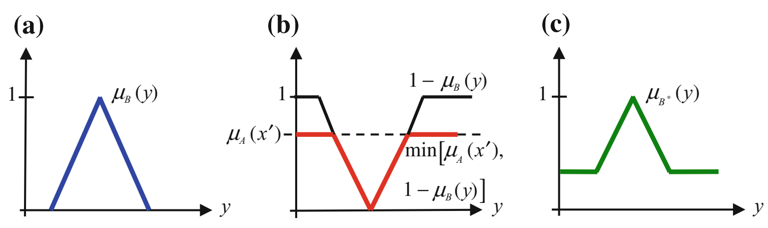


Fig. 2.17 Construction of $\mu_{B^*}(y)$ in (2.130). **a** Consequent MF $\mu_B(y)$, **b** construction of $\min[\mu_A(x'), 1 - \mu_B(y)]$, and **c** $\mu_{B^*}(y)$ (Mendel 1995a © 1995, IEEE)

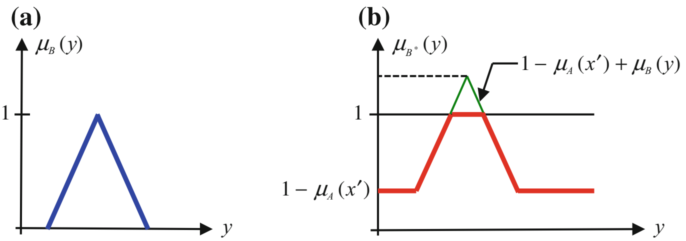


Fig. 2.18 Construction of $\mu_{B^*}(y)$ in (2.132). **a** Consequent MF $\mu_B(y)$, **b** construction of $\mu_{B^*}(y)$ (Mendel 1995a © 1995, IEEE)

The result shown in (c) is disturbing for an engineering application, i.e., given a specific input $x = x'$, the result of firing a specific rule, whose consequent is associated with a specific fuzzy set of finite support [the base of the triangle in (a)], is a fuzzy set whose support is infinite. Somehow a bias (constant) has gotten into the output so that regardless of x' the output is never zero [unless $\mu_A(x') = 1$]. This does not seem desirable for engineering applications.

Example 2.32 Perhaps the problem experienced in Example 2.31 was a result of a poor choice for $\mu_{A\to B}(x', y)$. Therefore, the result of using $\mu^L_{A\to B}(x', y)$ obtained from (2.123), is examined next, i.e.,

$$\mu^L_{A\to B}(x', y) = \min[1, 1 - \mu_A(x') + \mu_B(y)] \;\; y \in Y \tag{2.131}$$

which, by the way, is the implication MF given in the important paper (Zadeh 1973). Substituting this expression for $\mu^L_{A\to B}(x', y)$ into (2.129), it follows that:

$$\mu_{B^*}(y) = \mu^L_{A\to B}(x', y) = \min[1, 1 - \mu_A(x') + \mu_B(y)] \;\; y \in Y \tag{2.132}$$

A graphical interpretation of this result is given in Fig. 2.18. As in Example 2.31, the level shown for $\mu_A(x')$—and subsequently for $1 - \mu_A(x')$—was chosen arbitrarily. Once again, a result has been obtained in Fig. 2.18b that includes a bias. It is easy to demonstrate that all of the other choices provided earlier for $\mu_{A\to B}(x, y)$ have the same problem. Even many choices not listed here have the same problem.

2.17 Mamdani (Engineering) Implications

Mamdani (1974) seems to have been the first one to recognize the problem just demonstrated. Based on simplifying the computations associated with (2.127), he chose to work with the following *minimum implication* (inference)

Table 2.7 Demonstration that minimum and product implications do not agree with $\mu_{p \to q}(x, y)$ (Mendel 1995a © 1995, IEEE)

$\mu_p(x)$	$\mu_q(y)$	$\min[\mu_p(x), \mu_q(y)]$	$\mu_p(x)\mu_q(y)$	$\mu_{p \to q}(x, y)$
1	1	1	1	1
1	0	0	0	0
0	1	0	0	1
0	0	0	0	1

$$\mu_{A \to B}(x, y) \equiv \min[\mu_A(x), \mu_B(y)] \quad x \in X,\ \mathrm{y} \in \mathrm{Y} \tag{2.133}$$

Later, Larsen (1980) proposed the following *product implication* (inference)

$$\mu_{A \to B}(x, y) \equiv \mu_A(x)\mu_B(y) \quad x \in X, y \in Y \tag{2.134}$$

Again, the reason for this choice was simplicity of computation.[34]

Equations (2.133) and (2.134) can be expressed collectively as

$$\mu_{A \to B}(x, y) \equiv \mu_A(x) \bigstar \mu_B(y) \quad x \in X,\ y \in Y \tag{2.135}$$

where $\bigstar$ is a t-norm, product, or minimum, and is frequently referred to as a *Mamdani implication* regardless of whether the t-norm used is the minimum or product.

Today, minimum and product implications are the most widely used implications in the engineering applications of fuzzy logic; but, what do they have to do with traditional propositional logic? Table 2.7 demonstrates that neither minimum implication nor product implication agrees with the accepted propositional logic definition of implication; hence, minimum and product implications have nothing to do with traditional propositional logic. Consequently, minimum and product implications—Mamdani implications—can be thought of as *engineering implications*.

Because of the use of engineering implication functions in rule-based fuzzy systems and their disconnect from material implication, I now believe it would be better to call such systems "fuzzy systems" rather than "fuzzy logic systems". Hence, *in this book "fuzzy system" is used instead of "fuzzy logic system"*, but "fuzzy system" is not abbreviated to FS, because to do so would confuse it with a fuzzy set.

[34]There is a paragraph in the lower right-hand column on p. 359 of Mendel (1995a) that contains an error. Observe that the derivation of (2.129) has accounted for all values of x, including $x \neq x'$, because it uses (2.128). For some reason that I cannot recall, in the erroneous paragraph, I claim that for all $x \neq x'$, $\mu_{B^*}(y) = 1$, which I then interpret as a form of non-causality, i.e., a rule will be fired for all $x \neq x'$. I then argue for the use of a Mamdani or Larsen implication on the basis of their causality. This is incorrect; however, it does not affect anything else in the 1995 tutorial.

Example 2.33 The purpose of this example is to demonstrate that both the minimum and product implications lead to output fuzzy sets that seem quite reasonable from an engineering perspective, in that they only alter the shape of $\mu_B(y)$ and do not introduce a bias. As in Examples 2.31 and 2.32, singleton fuzzification is assumed, i.e., $\mu_{A^*}(x)$ is given by (2.128).

Considering minimum implication first, (2.129) becomes

$$\mu_{B^*}(y) = \min[\mu_A(x'), \mu_B(y)]\, y \in Y \tag{2.136}$$

A graphical interpretation of this result is given in Fig. 2.19. As in those earlier examples, the level shown for $\mu_A(x')$ was chosen arbitrarily. Observe from Fig. 2.19b that given a specific antecedent $x = x'$ the result of firing a specific rule is a fuzzy set whose support is finite and whose shape is a clipped version of $\mu_B(y)$.

Considering the product implication next, (2.129) becomes:

$$\mu_{B^*}(y) = \mu_A(x')\mu_B(y)\, y \in Y \tag{2.137}$$

A graphical interpretation of this result is given in Fig. 2.20. Similar conclusions are drawn from this figure as were drawn for minimum implication. In this case, the shape of the fuzzy output set is a scaled (attenuated) version of $\mu_B(y)$.

Overall conclusions are that minimum and product implications are, indeed, useful engineering implications, and, that $\mu_{B^*}(y)$ can be expressed as

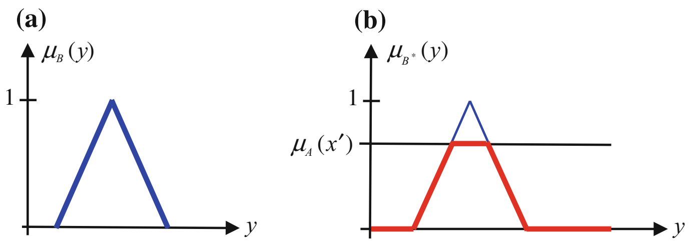


Fig. 2.19 Construction of $\mu_{B^*}(y)$ in (2.136). **a** Consequent MF $\mu_B(y)$, **b** construction of $\mu_{B^*}(y)$ (Mendel 1995a © 1995, IEEE)

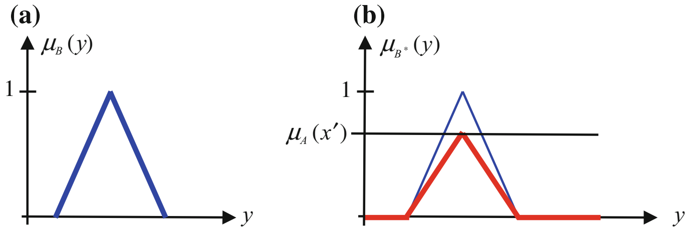


Fig. 2.20 Construction of $\mu_{B^*}(y)$ in (2.137). **a** Consequent MF $\mu_B(y)$, **b** construction of $\mu_{B^*}(y)$ (Mendel 1995a © 1995, IEEE)

$$\mu_{B^*}(y) = \mu_A(x')\bigstar\mu_B(y)\; y \in Y \tag{2.138}$$

where $\bigstar$ is either the minimum or product.

Example 2.34 When there is some uncertainty about the measurement of input variable x, then the measurement can be modeled as a type-1 fuzzy number (in Chap. 3 this will be called a *non-singleton fuzzifier*). Let the measured value of x be denoted x'. In this example a type-1 fuzzy number is created that is centered about x' by using the following Gaussian MF for A^*:

$$\mu_{A^*}(x) = \exp\left(-[(x - x')/\sigma_{A^*}]^2 \Big/ 2\right) \tag{2.139}$$

Here only a single antecedent rule is considered, one whose antecedent MF is also assumed to be a Gaussian, namely:

$$\mu_A(x) = \exp\left(-[(x - m_A)/\sigma_A]^2 \Big/ 2\right) \tag{2.140}$$

Mamdani product implication and product t-norm are assumed, and the goal here is to evaluate the sup-star (product) composition in (2.127).

First, it is shown that the sup-star composition in (2.127) can be expressed as $(y \in Y)$

$$\mu_{B^*}(y) = \left(\sup_{x\in X}[\mu_{A^*}(x)\mu_A(x)]\right) \times \mu_B(y) \tag{2.141}$$

Using product implication, $\mu_{A\to B}(x, y) = \mu_A(x)\mu_B(y)$, and using product t-norm $\bigstar = \times$, (2.127) becomes:

$$\mu_{B*}(y) = \sup_{x\in X}[\mu_{A*}(x)\mu_A(x)\mu_B(y)] = \left(\sup_{x\in X}[\mu_{A*}(x)\mu_A(x)]\right) \times \mu_B(y) \tag{2.142}$$

which is (2.141).

Next, the value of x is established where $\sup_{x\in X}[\mu_{A^*}(x)\mu_A(x)]$ occurs. Let $f(x) \equiv \mu_{A*}(x)\mu_A(x)$, and substitute the Gaussian MFs given in (2.139) and (2.140) into it, to see that

$$f(x) = \exp\left\{-\frac{1}{2}\left[\left(\frac{x - x'}{\sigma_{A*}}\right)^2 + \left(\frac{x - m_A}{\sigma_A}\right)^2\right]\right\} \equiv \exp\left\{-\tfrac{1}{2}\varphi(x)\right\} \tag{2.143}$$

To maximize $f(x)$, $\varphi(x)$, must be minimized; hence, one proceeds as follows:

$$\frac{\partial\varphi(x)}{\partial x} = 2\left(\frac{x - x'}{\sigma_{A*}^2}\right) + 2\left(\frac{x - m_A}{\sigma_A^2}\right) \tag{2.144}$$

Note that $\partial^2\varphi(x)/\partial x^2 = 2/\sigma_{A*}^2 + 2/\sigma_A^2 > 0$; hence, setting $\partial\phi(x)/\partial x = 0$ leads to the value of x that *minimizes* $\varphi(x)$, and subsequently maximizes $f(x)$, i.e.: $\partial\varphi(x)/\partial x = 0 \Rightarrow x = x_{\max}$, which leads to $(x_{\max} - x')\sigma_A^2 + (x_{\max} - m_A)\sigma_{A*}^2 = 0$, from which it is straightforward to show that

$$x_{\max} = \frac{\sigma_{A^*}^2 m_A + \sigma_A^2 x'}{\sigma_A^2 + \sigma_{A^*}^2} \tag{2.145}$$

Finally, $f(x_{\max}) = \sup_{x\in X}[\mu_{A^*}(x)\mu_A(x)]$ is computed. Substitute $x = x_{\max}$ into $\mu_{A*}(x)\mu_A(x)$ and use the middle part of (2.143), to obtain:

$$\begin{aligned} f(x_{\max}) &= \sup_{x\in X}[\mu_{A^*}(x)\mu_A(x)] = \mu_{A^*}(x_{\max})\mu_A(x_{\max}) \\ &= \exp\left\{-\frac{1}{2}\left[\left(\frac{x_{\max} - x'}{\sigma_{A^*}}\right)^2 + \left(\frac{x_{\max} - m_A}{\sigma_A}\right)^2\right]\right\} \end{aligned} \tag{2.146}$$

Using (2.145), it follows that:

$$\frac{x_{\max} - x'}{\sigma_{A*}} = \frac{\sigma_{A*}^2 m_A + \sigma_A^2 x' - (\sigma_A^2 + \sigma_{A*}^2)x'}{(\sigma_A^2 + \sigma_{A*}^2)\sigma_{A*}} = \frac{\sigma_{A*}(m_A - x')}{(\sigma_A^2 + \sigma_{A*}^2)} \tag{2.147}$$

$$\frac{x_{\max} - m_A}{\sigma_A} = \frac{\sigma_{A*}^2 m_A + \sigma_A^2 x' - (\sigma_A^2 + \sigma_{A*}^2)m_A}{(\sigma_A^2 + \sigma_{A*}^2)\sigma_A} = \frac{\sigma_A(x' - m_A)}{(\sigma_A^2 + \sigma_{A*}^2)} \tag{2.148}$$

Consequently,

$$\begin{aligned} f(x_{\max}) &= \exp\left\{-\frac{1}{2}\left[\frac{\sigma_{A*}^2(m_A - x')^2 + \sigma_A^2(x' - m_A)^2}{(\sigma_A^2 + \sigma_{A*}^2)^2}\right]\right\} \\ &= \exp\left\{-\frac{1}{2}\left[\frac{(x' - m_A)^2}{(\sigma_A^2 + \sigma_{A*}^2)}\right]\right\} = f(x') \end{aligned} \tag{2.149}$$

Observe that $f(x_{\max})$ depends on the measured value of x, x', and so it can be treated as a function of x'. Observe, also, that $f(x')$ is also a Gaussian function, one that is centered about m_A and has a variance that is equal to $\sigma_A^2 + \sigma_{A*}^2$; hence, this Gaussian is more spread out than either $\mu_{A^*}(x)$ or $\mu_A(x)$. Once can therefore conclude that the effect of uncertainty on the measured input is to spread out the antecedent's MF.

Exercise 2.40 asks the reader to repeat these computations for Mamdani minimum implication and the minimum t-norm.

2.18 Remarks

So far, all discussions about rules have been for rules with single antecedents, e.g., IF x is A, THEN y is B. Chap. 3 and later chapters describe and characterize rules that have more than one antecedent, e.g.,

$$\text{IF } x_1 \text{ is } F_1 \text{ and } x_2 \text{ is } F_2 \text{ and } \ldots \text{ and } x_p \text{ is } F_p, \text{ THEN } y \text{ is } G$$

In such a multiple-antecedent rule, $x_1 \in X_1, \ldots, x_p \in X_p$, $y \in Y$, and $F_1, \ldots, F_p$ and G are fuzzy sets.

Some other topics, which appear frequently in the fuzzy set literature and are sometimes used in engineering applications of fuzzy set and logic, include: cardinality, similarity and subsethood. Because none of them are used in this book, although they could be used in other applications of rule-based systems, they are left for Exercises 2.43, 2.44 and 2.45, respectively.

The different t-norms, t-conorms, and complements that are available from fuzzy set theory provide some (tough) choices that have to be made in a fuzzy system. Zimmerman (1991, pp. 42–43) describes eight criteria that might be helpful in selecting the connective's operator. Unfortunately, I found most of those criteria to be so subjective that I could not use them in my engineering applications of fuzzy sets.

It is very difficult to make a decision about which t-norm or t-co-norm to use in the fuzzy domain because usually different numerical values are obtained for each choice. It is only back in the crisp domain where the same numerical values are obtained for the different choices that one can make a choice based on complexity (simplicity[35]) of the choice. Interestingly, Zadeh seems to only use the minimum or product for conjunction and the maximum for disjunction, the least complex choices.

Most rule-based engineering applications of fuzzy sets use: (1) the minimum or algebraic product t-norm for fuzzy intersection, (2) the maximum t-conorm for fuzzy union, and (3) $1 - \mu_A(x)$ for the MF of the fuzzy complement. These choices are adhered to in this book.

Finally, I want to comment on fuzzy sets and probability.[36] Some people maintain that there is no difference between fuzzy sets and probability. When I am asked about this, often at the beginning of a lecture or course on fuzzy sets and systems, I ask the following question: "How many of you have had a formal course on probability?" Usually, all hands go up. Then I ask: "How many of you have had a formal course on fuzzy sets and systems." Usually, no hands, or only a very small number of hands go up. I then state that in order to explain the differences between fuzzy sets and probability, one must first spend time formally understanding fuzzy

[35]This is based on Ockham's razor principle; see footnote 13 in Chap. 6 (page 272) for a discussion about this principle.

[36]The rest of the material in this section is taken for the most part from Mendel (1995a).

sets. Only then can intelligent comparisons be made between that which one understands (probability) and that which one will understand (fuzzy sets).

Having just read this chapter, fuzzy sets and probability can now be discussed intelligently.

A lot has been written about fuzzy sets and their relation to probability [e.g., (Cheeseman 1988; Kosko 1990; Laviolette and Seaman 1994; Lindley 1982) and, *IEEE Trans. on Fuzzy Systems*, March 1994, Special Issue]. Many fuzzy set theorists maintain that fuzzy sets are quite different than probability, for a wide variety of reasons, including the facts that: the laws of excluded middle and contradiction are broken for fuzzy sets, but are not broken in probability, and, that conditional probability must be defined in probability theory, but can be derived from first principles using fuzzy sets (Kosko 1990, 1992). Others maintain that fuzzy sets subsume probability. Subjective (as distinguished from frequency-based) probabilists on the other hand, maintain that anything one can do with fuzzy sets can also be done with subjective probability, and that the latter is to be preferred because it has an axiomatic basis, whereas fuzzy sets do not. They bemoan the fact that engineers, who are the largest users of fuzzy systems, are not adequately trained in subjective probability.

The fact of the matter is that there is some truth to both sides of *fuzziness versus probability*. While it is of great intellectual interest to establish the proper connections between fuzzy sets and probability, this author does not believe that doing so will change the ways in which one solves problems, because both probability and fuzzy sets should be in the arsenal of tools used by engineers. Fuzzy sets will not solve all problems, nor will probability.

That fuzzy sets are a tool of enrichment and not replacement is clearly explained in Bezdek and Pal (1992) who ask the question: "Where do fuzzy models fit in with other models?" They then give the following answer (Bezdek and Pal 1992, © IEEE 1992):

> Fuzzy models belong wherever they can provide collateral or competitively better information about a physical process. … we note that each of the following disciplines provides some information about the dynamics of motion: Newtonian mechanics, relativistic mechanics, statistical mechanics, quantum mechanics, and auto mechanics. These models provide us with different, useful, auxiliary, and sometimes contradictory information about various facets of dynamics. Each contributes something about the physical world, so it is with various classes of models. … From a different point of view, because every hard set is fuzzy but not conversely, the mathematical embedding of conventional set theory into fuzzy sets is as natural as the idea of embedding the real numbers into the complex plane. In both cases we can expect the larger 'space' to contain answers to (real) questions that cannot be found in the smaller one. Thus the idea of fuzziness is one of enrichment not of replacement.

Addressing the fuzziness versus probability issue, Bezdek and Pal also ask: "Isn't fuzziness just a clever disguise for probability?" Their answer is ([5], © IEEE 1992):

> … an emphatic no. There is a strong philosophical argument against regarding fuzziness as the surrogate for (frequency-based) probability. The *spirit* of this argument is contained in (the following) example. Let L = set of all liquids, and let fuzzy subset L = {all (*potable*)

liquids}. Suppose you had been in the desert for a week without a drink and you came upon two bottles marked C and A [bottle C is labeled $\mu_L(C) = 0.91$ and bottle A is labeled $\Pr[A \in L] = 0.91$]. Confronted with this pair of bottles, and given that you must drink from the one you choose, which would you choose to drink from? Most readers when presented with this experiment immediately see that while C could contain, say, swamp water, it would not … contain liquids such as hydrochloric acid. That is *membership* of 0.91 means that the contents of C are fairly similar to perfectly potable liquids (e.g., pure water). On the other hand, the probability that A is potable = 0.91 means that over a long run of experiments, the contents of A are expected to be potable in about 91% of the trials; in the other 9% the contents will be deadly—about a 1 chance in 10. Thus, most subjects will opt for a chance to drink swamp water. … There is another facet to this example, and it concerns the idea of *observation*. Continuing then, suppose we examine the contents of C and A and discover them to be Dixie beer and hydrochloric acid, respectively. Note that, *after observation*, the membership value of C is unchanged while the probability value for A drops from 0.91 to 0.0. This example shows that these two models possess philosophically different kinds of information: fuzzy memberships, which represent similarities of objects to imprecisely defined properties; and probabilities, which convey information about relative frequencies.

Appendix 1: Properties of Type-1 Fuzzy Sets

This appendix presents details about properties/laws of type-1 fuzzy sets and examines the following frequently used laws to see if they remain satisfied under maximum t-conorm and either minimum or product t-norms:

Reflexive, anti-symmetric, transitive, idempotent, commutative, associative, absorption, distributive, involution, De Morgan's, and identity

Our reason for doing this is that rules in a rule-based system may make use of the words "and", "or", "unless", "not", etc., but all of the mathematics for such a system is worked out in this book only for canonical rules that use the words "and" and "or". Section 3.2 shows how the former rules can be transformed into the canonical rules by using some of the above laws. So, it is important to know when or if the use of these laws is correct.

The exact nature of all the preceding laws is given in the second column of Table 2.8. These laws are all satisfied for crisp sets (for the minimum and product t-norms), due to the facts that: $\min(0,0) = 0$ and $0 \times 0 = 0$, $\min(1,0) = 0$ and $1 \times 0 = 0$, $\min(0,1) = 0$ and $0 \times 1 = 0$, and, $\min(1,1) = 1$ and $1 \times 1 = 1$. That they are all satisfied for maximum t-conorm and minimum t-norm (a so-called "dual t-conorm and t-norm pair") is well known (e.g. Klir and Yuan 1995) and proofs for this situation are left to the reader (Exercise 2.41).

The rest of this appendix focuses on the maximum t-norm and product t-norm pairing. Reflexive, anti-symmetric, and transitive laws do not make use of any t-norm; hence, they are automatically satisfied for maximum t-conorm and product t-norm. Commutative and associative laws are also satisfied, because both maximum and product operations are commutative and associative; i.e., for $x \in X$ ($\vee \equiv$ maximum):

Table 2.8 Summary of set-theoretic laws and whether or not they are satisfied for type-1 fuzzy sets under maximum t-conorm and either minimum or product t-norms[a]

Set theoretic laws		Minimum t-norm	Product t-norm
Reflexive	$\mu_A \le \mu_A$	Yes	Yes
Anti-symmetric	$\mu_A \le \mu_B, \mu_B \le \mu_A \Rightarrow \mu_A = \mu_B$	Yes	Yes
Transitive	$\mu_A \le \mu_B, \mu_B \le \mu_C \Rightarrow \mu_A \le \mu_C$	Yes	Yes
Idempotent	$\mu_A \vee \mu_A = \mu_A$	Yes	Yes
	$\mu_A \bigstar \mu_A = \mu_A$	Yes	**NO**
Commutative	$\mu_A \vee \mu_B = \mu_B \vee \mu_A$	Yes	Yes
	$\mu_A \bigstar \mu_B = \mu_B \bigstar \mu_A$	Yes	Yes
Associative	$(\mu_A \vee \mu_B) \vee \mu_C = \mu_A \vee (\mu_B \vee \mu_C)$	Yes	Yes
	$(\mu_A \bigstar \mu_B) \bigstar \mu_C = \mu_A \bigstar (\mu_B \bigstar \mu_C)$	Yes	Yes
Absorption	$\mu_A \bigstar (\mu_A \vee \mu_B) = \mu_A$	Yes	**NO**
	$\mu_A \vee (\mu_A \bigstar \mu_B) = \mu_A$	Yes	Yes
Distributive	$\mu_A \bigstar (\mu_B \vee \mu_C) = (\mu_A \bigstar \mu_B) \vee (\mu_A \bigstar \mu_C)$	Yes	Yes
	$\mu_A \vee (\mu_B \bigstar \mu_C) = (\mu_A \vee \mu_B) \bigstar (\mu_A \vee \mu_C)$	Yes	**NO**
Involution	$\mu_{\bar{\bar{A}}} = \mu_A$	Yes	Yes
De Morgan's Laws	$\overline{\mu_A \vee \mu_B} = \mu_{\bar{A}} \bigstar \mu_{\bar{B}}$	Yes	**NO**
	$\overline{\mu_A \bigstar \mu_B} = \mu_{\bar{A}} \vee \mu_{\bar{B}}$	Yes	**NO**
Identity	$\mu_A \vee 0 = \mu_A$	Yes	Yes
	$\mu_A \bigstar 1 = \mu_A$	Yes	Yes
	$\mu_A \vee 1 = 1$	Yes	Yes
	$\mu_A \bigstar 0 = 0$	Yes	Yes

Adapted from Table 1 of Karnik and Mendel (2001)
[a]Arguments of all MFs have been omitted; hence, μ_A, for example, is short for $\mu_A(x)$

$$\mu_A(x) \vee \mu_B(x) = \mu_B(x) \vee \mu_A(x)$$

$$\mu_A(x) \times \mu_B(x) = \mu_B(x) \times \mu_A(x)$$

$$(\mu_A(x) \vee \mu_B(x)) \vee \mu_C(x) = \mu_A(x) \vee (\mu_B(x) \vee \mu_C(x))$$

$$(\mu_A(x) \times \mu_B(x)) \times \mu_C(x) = \mu_A(x) \times (\mu_B(x) \times \mu_C(x))$$

Under product t-norm, the second part of the absorption laws is satisfied, because $\mu_A(x) \times \mu_B(x) \le \mu_A(x)$, so that $\mu_A(x) \vee (\mu_A(x) \times \mu_B(x)) = \mu_A(x)$. The first part of the distributive laws is satisfied; i.e., product is distributive over maximum. The first part of the idempotent laws is also satisfied; i.e., $\mu_A(x) \vee \mu_A(x) = \mu_A(x)$. The involution law is satisfied, since complement is defined as $\mu_{\bar{A}}(x) = 1 - \mu_A(x)$. And, all the identity laws are satisfied (i.e., $\mu_A(x) \vee 0 = \mu_A(x)$, $\mu_A(x) \times 1 = \mu_A(x)$, $\mu_A(x) \vee 1 = 1$, and $\mu_A(x) \times 0 = 0$).

None of the other laws are satisfied under product t-norm, because:

- Idempotent laws—second part

$$\mu_A(x) \times \mu_A(x) \neq \mu_A(x) \tag{2.150}$$

- Absorption laws—first part: assume, e.g. that $\mu_A(x) > \mu_B(x)$; then,

$$\mu_A(x) \times (\mu_A(x) \vee \mu_B(x)) = \mu_A(x) \times \mu_A(x) = \mu_A^2(x) \neq \mu_A(x) \tag{2.151}$$

- Distributive laws–second part: assume, e.g. that $\mu_A(x) > \mu_B(x)$ and $\mu_A(x) > \mu_C(x)$; then,

$$\begin{aligned} \mu_A(x) \vee (\mu_B(x) \times \mu_C(x)) &= \mu_A(x) \\ &\neq (\mu_A(x) \vee \mu_B(x)) \times (\mu_A(x) \vee \mu_C(x)) = \mu_A^2(x) \end{aligned} \tag{2.152}$$

- De Morgan's laws:

$$\begin{aligned} \overline{\mu_A(x) \vee \mu_B(x)} &= 1 - (\mu_A(x) \vee \mu_B(x)) \\ &\neq \mu_{\bar{A}}(x) \times \mu_{\bar{B}}(x) = (1 - \mu_A(x)) \times (1 - \mu_B(x)) \end{aligned} \tag{2.153}$$

$$\begin{aligned} \overline{\mu_A(x) \times \mu_B(x)} &= 1 - \mu_A(x) \times \mu_B(x) \\ &\neq \mu_{\bar{A}}(x) \vee \mu_{\bar{B}}(x) = \max\{(1 - \mu_A(x)), (1 - \mu_B(x))\} \end{aligned} \tag{2.154}$$

Exercises

2.1 Fuzziness as a concept that lets an object reside in more than one set but to different degrees may be traced back to antiquity. Go on the Internet and find a picture of the statue called the *Guardian Sphinx* (530 BC.).

   (a) What are the three sets for this statue?
   (b) What membership grade would you assign to each of the three sets?

2.2 Fuzziness as a concept that lets an object reside in more than one set but to different degrees has occurred in art, even before Zadeh formalized it. For example, it occurs in the works of the Belgian painter Renè Magritte. Go on the Internet and find the following paintings by him and answer the related question:

   (a) *The Explanation* (1952): What is the degree of similarity between the carrot and the wine bottle?
   (b) *Homage to Alphonse Allais* (1964): What is the degree of similarity between the cigar and the fish?

2.3 Suppose that a car is described by its *color*. What scale could be used for color? Create five terms for color and sketch MFs for each term.

2.4 Establish MFs for:

(a) real numbers close to 10
(b) real numbers approximately equal to 6
(c) integers very far from 10
(d) complex numbers near the origin
(e) light (weight)
(f) heavy (weight).

2.5 List six linguistic variables from the field of acoustics (or any field that is of interest to you).

2.6 Using the rules in Example 2.5 as illustrations, list four more rules and their associated MFs.

2.7 Let X be the set of all men and Y be the set of all women. Consider the linguistic variable "weight," and the set of terms {*very skinny*, *skinny*, *just right*, *heavy*, *very heavy*}. Create MFs for these terms for both men and women.

2.8 Consider the judgments listed here, and assume that they can be mapped onto an interval scale ranging from 0 to 10. Define five fuzzy sets for each of them and sketch what you feel are appropriate MFs for them.

(a) touching
(b) eye contact
(c) smiling
(d) acting witty
(e) flirtation.

2.9 Western logic and thinking has been dominated for the most part by the Aristotelian laws of contradiction and the excluded middle. Eastern thinking has not. Eastern religions and concepts such as the Yin and the Yang (female and male/opposite forces) have caused some to speculate that this is why China and Japan were more receptive to fuzzy logic than were people in the West. For example, it's possible for each of you to reside in Yin and Yang simultaneously, but to different degrees. Explain this in terms of fuzzy sets.

2.10 Prove that, for crisp sets A and B, $\min[\mu_A(x), \mu_B(x)]$ provides the correct MF for intersection, given in (2.12).

2.11 For crisp sets A and B, prove the:

(a) commutative law
(b) associative laws
(c) distributive laws
(d) De Morgan's laws.

2.12 Consider three fuzzy sets, A, B, and C, whose MFs are (unnormalized) Gaussians, i.e., $\mu_A(x) = \exp\left[-\frac{1}{2}(x-3)^2\right]$, $\mu_B(x) = \exp\left[-\frac{1}{2}(x-4)^2\right]$ and $\mu_C(x) = \exp\left[-\frac{1}{2}(x-6)^2\right]$. Sketch each of the following:

(a) $$\mu_{A\cap B\cap C}(x)$$

(b) $$\mu_{A\cup B\cup C}(x)$$

(c) $$\mu_{(A\cup B)\cap C}(x) \text{ and } \mu_{A\cup(B\cap C)}(x)$$

(d) $$\mu_{(A\cap B)\cup C}(x) \text{ and } \mu_{A\cap(B\cup C)}(x)$$

(e) $$\mu_{\overline{A\cup B\cup C}}(x).$$

2.13 Consider the fuzzy sets A and B, where $\mu_A(x) = \exp\left[-\frac{1}{2}(x-3)^2\right]$ and $\mu_B(x) = \exp\left[-\frac{1}{2}(x-4)^2\right]$.

(a) Sketch $\mu_{A\cup B}(x)$ for the following t-conorms: maximum, algebraic sum, bounded sum and drastic sum. Which t-conorm gives the largest and smallest values for $\mu_{A\cup B}(x)$?
(b) Sketch $\mu_{A\cap B}(x)$ for the following t-norms: minimum, algebraic product, bounded product and drastic product. Which t-norm gives the largest and smallest values for $\mu_{A\cap B}(x)$?

2.14 Using (2.34) and (2.35), show that $\mu_{c\cup s}(u,v)$ and $\mu_{c\cap s}(u,v)$ are given by (2.38) and (2.39), respectively.
2.15 Verify the max-min and max-product composition of the crisp relations for the (3, 3) element of $R_3(U,W)$ in (2.43).
2.16 Consider the fuzzy relations "u is lighter than v" *or* "u is about the same weight as v." Assume that $u \in U$ and $v \in V$ where U and V are discrete universes of discourse, and U has four elements whereas V has six elements.

(a) Pick U and V to use in the rest of this exercise.
(b) Establish MFs for *lighter* and *about the same*, i.e., $\mu_l(u,v)$ and $\mu_{ats}(u,v)$, where the numbers in $\mu_l(u,v)$ and $\mu_{ats}(u,v)$ agree with a comparison of the numbers in U and V.
(c) Compute $\mu_{l\cup ats}(u,v)$.

2.17 Perform all of the calculations needed to obtain $\mu_{comb}(u,w)$ given in (2.54).
2.18 Repeat Example 2.15 using the product t-norm. Compare these results with the ones given in (2.54) which were obtained using the minimum t-norm. Are they significantly different?
2.19 Consider the fuzzy relation "u is lighter than v" on $U \times V$, and the fuzzy relation "v is heavier than w" on $V \times W$. Assume that U, V, and W are discrete universes of discourse, and U has four elements, V has six elements, and W has three elements.

(a) Pick U, V, and W to use in the rest of this exercise.
(b) Establish MFs for *lighter* and *heavier*, i.e., $\mu_l(u,v)$ and $\mu_h(v,w)$, where the numbers in $\mu_l(u,v)$ and $\mu_h(v,w)$ agree with a comparison of the numbers in U, V, and W.
(c) Compute $\mu_{l \circ h}(u,w)$ using minimum t-norm.
(d) Compute $\mu_{l \circ h}(u,w)$ using product t-norm.
(e) Compare the results from (c) and (d).

2.20 Consider the fuzzy relation "u is lighter than v" on $U \times V$. Assume that U and V are discrete universes of discourse, and U has four elements and V has six elements.

(a) Pick U and V to use in the rest of this exercise.
(b) Establish a MF for *lighter*, i.e., $\mu_l(u,v)$, where the numbers in $\mu_l(u,v)$ agree with a comparison of the numbers in U and V.
(c) Construct a MF for the fuzzy set *skinny*, $\mu_{skinny}(u)$, on U.
(d) Compute the composition of "u is skinny" *and* "u is lighter than v", $\mu_{skinny \circ l}(v)$.

2.21 Using the same universe of discourse as in Example 2.17, develop MFs for:

(a) very likely
(b) not-too-likely.

2.22 Suppose that $U = \{-5,-4,-3,-2,-1,0,1,2,3,4,5\}$ and fuzzy set A is characterized by the MF

$$\begin{aligned}\mu_A(x) = 0.2/-5 + 0.4/-4 + 0.4/-3 + 0.5/-2 + 0.5/\\ -1 + 0.6/0 + 0.9/1 + 1/2 + 0.8/3 + 0.5/4 + 0.1/5\end{aligned}$$

(a) Determine the MF for the fuzzy set B that is associated with $\mu_{f(A)}(y)$ when $y = f(x) = x^3 + 2x^2$.
(b) Determine the MF for the fuzzy set B that is associated with $\mu_{f(A)}(y)$ when $y = |x|$.

2.23 Suppose that $X_1 = \{1,2,3,4\}$ and $X_2 = \{-1,-2,-3,-4\}$, and fuzzy sets A_1 and A_2 are characterized by the following MFs:

$$\begin{aligned}\mu_{A_1}(x_1) &= 0.5/1 + 0.5/2 + 0/3 + 1/4 \text{ and}\\ \mu_{A_2}(x_2) &= 1/-1 + 0/-2 + 0.25/-3 + 0.5/-4\end{aligned}$$

Determine the MF for the fuzzy set B that is associated with $\mu_{f(A_1A_2)}(y)$, when $y = f(x_1,x_2) = x_1^2 - 2x_2^2$.

2.24 Given the type-1 Gaussian fuzzy set F_i, with mean m_i and standard deviation σ_i, prove that $a_iF_i + b$ is a Gaussian fuzzy set with mean $a_im_i + b$ and standard deviation $|a_i\sigma_i|$. Note that this result does not depend on the kind of t-norm used, since a_i and b are crisp numbers.

2.25 Given n type-1 Gaussian fuzzy sets $F_1, \ldots, F_n$, with means $m_1, \ldots, m_n$ and standard deviations $\sigma_1, \ldots \sigma_n$, as in (2.77), prove that $\sum_{i=1}^{n} F_i$ is a Gaussian fuzzy set with mean $\sum_{i=1}^{n} m_i$ and standard deviation Σ'', where

$$\Sigma'' = \begin{cases} \sqrt{\sum_{i=1}^{n} \sigma_i^2} & \text{if product t-norm is used} \\ \sum_{i=1}^{n} \sigma_i & \text{if minimum t-norm is used} \end{cases}$$

[Hints: (1) First prove the results for two sets and then for three sets; (2) show that the supremum of the minimum of two Gaussians is reached at their point of intersection lying between their means.]

2.26 Complete part (b) in the proof of Example 2.21.

2.27 In Example 2.22, obtain the comparable results when a_i are positive or negative real numbers.

2.28 Prove (2.81).

2.29 Repeat Example 2.28 but now for $\mu_{A \cap B}(x)$.

2.30 Let[37] $X_i (i = 1, \ldots, n)$ be fuzzy sets with Gaussian MFs, $\mu_{X_i}(x_i) = \exp\left(-[(x_i - c_i)/\sigma_i]^2 \big/ 2\right)$, and $w_i \geq 0$ be constant weights with $\sum_{i=1}^{n} w_i = 1$. Using the Extension Principle with the minimum t-norm, prove that $Y_n = \sum_{i=1}^{n} w_i X_i$ is a fuzzy set with MF $\mu_{Y_n}(y_n) = \exp\left(-\left[y_n - \sum_{i=1}^{n} w_i c_i\right]^2 \Big/ \left[\sum_{i=1}^{n} w_i \sigma_i\right]^2\right)$. [Hint: Prove this by using mathematical induction.]

2.31 Let[38] $A = [a, b, c]$ and $B = [p, q, r]$ be two triangle type-1 fuzzy numbers with MF given in (2.103) and (2.104), respectively. Compute the MF of:

(a) $A - B$

(b) $\exp(A)$

(c) $\ln(A)$

(d) $\sqrt{A}$

(e) $(A)^{1/n}$.

2.32 Let[39] $A = [a, b, c]$ and $B = [p, q, r]$ be two positive triangle type-1 fuzzy numbers with MF given in (2.103) and (2.104), respectively. Compute the MF of:

(a) $A \cdot B$

(b) $A \div B$

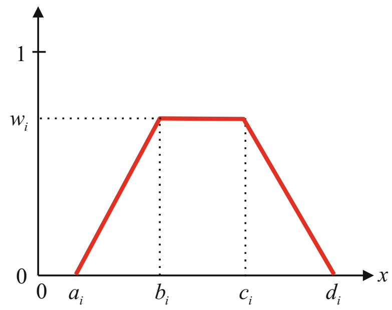
(c) $(A)^{-1}$.

[37] This exercise is adapted from Wang and Mendel (2016).

[38] This exercise is adapted from Dutta et al. (2011).

[39] This exercise is adapted from Dutta et al. (2011).

Fig. 2.21 Type-1 trapezoidal fuzzy number for Exercise 2.33



2.33 For the non-normal type-1 trapezoidal fuzzy number, $A_i = (a_i, b_i, c_i, d_i; w_i)$, whose MF is depicted in Fig. 2.21, prove that (Wei and Chen 2009):

(a) $A_1 + A_2 = (a_1 + a_2, b_1 + b_2, c_1 + c_2, d_1 + d_2; \min(w_1, w_2))$, where a_i, b_i, c_i and d_i are real numbers.
(b) $A_1 - A_2 = (a_1 - d_2, b_1 - c_2, c_1 - b_2, d_1 - a_2; \min(w_1, w_2))$, where a_i, b_i, c_i and d_i are real numbers.
(c) $A_1 \cdot A_2 \approx (a_1 \times a_2, b_1 \times b_2, c_1 \times c_2, d_1 \times d_2; \min(w_1, w_2))$, where a_i, b_i, c_i and d_i are positive real numbers.
(d) $A_1/A_2 \approx (a_1/d_2, b_1/c_2, c_1/b_2, d_1/a_2; \min(w_1, w_2))$, where a_i, b_i, c_i and d_i are non-zero positive real numbers.

In (c) and (d), $\approx$ means that the result is a convex type-1 fuzzy set, as in (2.7), in which $g(x)$ and $h(x)$ are not straight lines.

2.34 Using truth tables show that the following are tautologies [3]:

(a) $$p \wedge (q \vee r) \leftrightarrow (p \wedge q) \vee (p \wedge r)$$

(b) $$p \vee (q \wedge r) \leftrightarrow (p \vee q) \wedge (p \vee r)$$

(c) $$p \wedge (q \wedge r) \leftrightarrow (p \wedge q) \wedge r$$

(d) $$p \vee (q \vee r) \leftrightarrow (p \vee q) \vee r$$

2.35 Use truth tables to determine whether or not the following propositions are tautologies:

(a) $$(p \wedge q) \rightarrow (p \vee q)$$

(b) $$[(p \rightarrow q) \wedge (r \rightarrow s) \wedge (p \vee r)] \rightarrow (q \vee s)$$

(c) $$((p \wedge q) \rightarrow r) \leftrightarrow (p \rightarrow r) \vee (q \rightarrow r)$$

2.36 Prove that $[(A \wedge C) \rightarrow D] \wedge [(B \wedge C) \rightarrow D] \leftrightarrow [(A \vee B) \wedge C \rightarrow D]$[Hint: $(p \rightarrow q) \leftrightarrow (\sim p) \vee q$].
2.37 Validate the truth of the crisp implication MFs given in (2.122) and (2.123).
2.38 Repeat Example 2.32 for the following implication MFs, and indicate which of these has a bias in its output:

(a) Kleene-Dienes in (2.121)
(b) Reichenbach in (2.122)

(c) Gödel : $\mu^{G}_{A\to B}(x', y) = \begin{cases} 1 & \mu_{A}(x') \le \mu_B(y) \\ \mu_{B}(y) & \mu_A(x') > \mu_B(y) \end{cases}$

(d) Gaines Resher: $\mu^{GR}_{A\to B}(x', y) = \begin{cases} 1 & \mu_A(x') \le \mu_B(y) \\ 0 & \mu_A(x') > \mu_B(y) \end{cases}$

2.39

(a) For the upward sloping lines in Fig. 2.22a, show that the sup-min composition between the lines and the triangle always occurs at the intersection of the line and the right-hand leg of the triangle.
(b) For the downward sloping lines in Fig. 2.22b, show that the sup-min composition between the lines and the triangle always occurs at the intersection of the line and the left-hand leg of the triangle.

2.40 Everything is the same as in Example 2.34, except that in this exercise minimum implication and minimum t-norm are used.

(a) Show that, in this case, the sup-star composition in (2.127) can be expressed as

$$\mu_{B^*}(y) = \min\left[\sup_{x\in X}[\min[\mu_{A^*}(x), \mu_A(x)]], \mu_B(y)\right]$$

(b) Show that $\sup_{x\in X}[\min[\mu_{A^*}(x), \mu_A(x)]]$ occurs at the intersection point of the two Gaussian MFs, namely at

$$x = x_{\max} = (\sigma_{A^*} m_A + \sigma_A x')/(\sigma_{A^*} + \sigma_A).$$

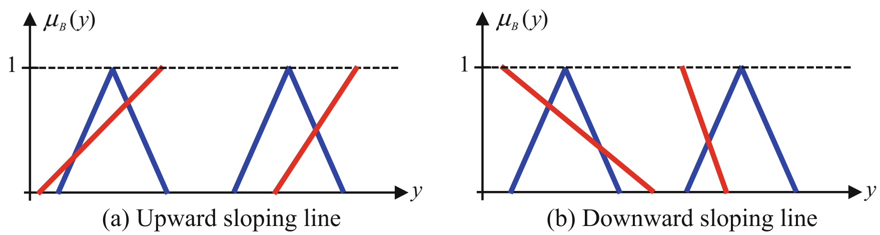


Fig. 2.22 Type-1 fuzzy sets for Exercise 2.39

(c) If possible, obtain a formula for $\sup_{x \in X}[\min[\mu_{A^*}(x), \mu_A(x)]]$.
(d) Assume a Gaussian consequent MF $\mu_B(y)$. Sketch the fired-rule MF $\mu_{B^*}(y)$. How is this obtained directly from sketches of $\mu_{A^*}(x)$, $\mu_A(x)$ and $\mu_B(y)$?
(e) Repeat part (d) for a triangular consequent MF.
(f) Compare the result in part (e) with the result in Fig. 2.19.

2.41 Show that for type-1 fuzzy sets all the set-theoretic laws that are in Table 2.8 are satisfied under maximum t-conorm and minimum t-norm.
2.42 Verify (2.153) and (2.154) numerically.
2.43 As one knows, crisp set A can be defined by using its MF in (2.1). The number of elements that are in A is called its *cardinality*. So, for a crisp set its cardinality can be obtained by summing all of its MF values. Using this idea[40], one can also define the *cardinality of a type-1 fuzzy set A*, $|A|$, analogously (De Luca and Termini 1972), i.e. for a discrete universe, $|A| = \sum_{i=1}^{N} \mu_A(x_i)$, and for a continuous universe, $|A| = \int_X \mu_A(x)dx$. Observe that $|A|$ increases as N increases, and $\lim_{N\to\infty} \sum_{i=1}^{N} \mu_A(x_i)$ does not exist. Wu and Mendel (2007) handle this by defining a *normalized cardinality*, $p(A)$, for a type-1 fuzzy set in which DeLuca and Termini's cardinality definition for continuous universes $|A| = \int_X \mu_A(x)dx$, is discretized, i.e.: $p(A) \equiv \frac{|X|}{N} \sum_{i=1}^{N} \mu_A(x_i)$, where $|X| = x_N - x_1$ is the length of the universe of discourse used in the computation. X can be part of the complete universe of discourse because for some MFs (e.g., Gaussian, Bell) the complete universes of discourse are infinite. Usually x_i $(i = 1, \ldots, N)$ are chosen equally spaced in the domain of x_i; in this case, $p(A)$ converges to its continuous version, $\int_X \mu_A(x)dx$ as N increases.

(a) Compute $|A|$ for the triangle and trapezoidal type-1 fuzzy sets that are in Table 2.3.
(b) Compute $p(A)$ for the same MFs used in (a) for $N = 10, 50, 100$, and compare these results with $|A|$.

2.44 *Similarity* is sometimes used in a rule-based fuzzy system, so this exercise explores similarity for type-1 fuzzy sets. Similarity is about set equality. Two crisp sets A and B are equal if they contain exactly the same elements. In

---

[40]The wording of the rest of this exercise is taken from Wu and Mendel (2007, p. 5383). The following is also taken from Wu and Mendel (2007, pp. 5382–5383): Definitions of the cardinality of type-1 fuzzy sets have been proposed by several authors, including De Luca and Termini (1972), Kaufman (1977), Gottwald (1980), Zadeh (1981), Blanchard (1982), Klement (1982) and Wygralak (1983). Basically there are two kinds of proposals (Dubois and Prade 1985; Wygralak 2003): (1) those that assume that the cardinality of a type-1 fuzzy set should be a precise number, and (2) those that claim it should be a fuzzy integer. De Luca and Termini's definition of cardinality (also called the *power* of a type-1 fuzzy set) is for the first proposal, is the one that is given in the statement of this exercise, and is the most frequently used definition of cardinality.

crisp set theory either two sets are equal or they are different. For fuzzy sets one knows that everything is a matter of degree; thus for two type-1 fuzzy sets A and B, it is reasonable to define a *degree of similarity*. As usual (in this book), crisp sets are our starting point.

As is stated in Nguyen and Kreinovich (2008): It is known that for two crisp sets A and B: (1) $A \cap B \subseteq A \cup B$ (create a Venn diagram to convince yourself of the truth of this), and (2) $A = B$ iff $A \cap B = A \cup B$. So, for crisp sets, to check whether $A = B$ consider the ratio $|A \cap B|/|A \cup B|$ where $|\bullet|$ denotes the cardinality of $\cdot$ (see Exercise 2.43 about cardinality). In general this ratio is between 0 and 1; the smaller the ratio, the more there are elements from $A \cup B$ which are not part of $A \cap B$, and thus elements from one of the sets A and B that do not belong to the other of these two sets. Thus, for crisp sets, this ratio can be viewed as a reasonable measure of degree to which A is equal to B.

Because there are many definitions of cardinality for a type-1 fuzzy set, and because there can be many ways to define the similarity between two type-1 fuzzy sets (Mendel and Wu 2010 mention that there are at least 50 reported expressions for determining how similar two type-1 fuzzy sets are), this exercise focuses on what is arguably the most popular and useful definition of similarity, the so-called *Jaccard similarity measure*, named after P. Jaccard (Jaccard 1908), who is credited with such a formula.[41] The Jaccard similarity measure, $sm_J(A,B)$, for type-1 fuzzy sets A and B, is: $sm_J(A,B) = f(A \cap B)/f(A \cup B)$. Usually, function f is chosen as the cardinality where $\cap$ = min and $\cup$ = max. For a continuous universe of discourse:

$$sm_J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{\int_X \min(\mu_A(x), \mu_B(x))dx}{\int_X \max(\mu_A(x), \mu_B(x))dx}$$

(a) What is the formula for $sm_J(A,B)$ for discrete universes of discourse?
(b) Compute $sm_J(A,B)$ for the two type-1 fuzzy sets that are depicted in Fig. 2.23a.
(c) Compute $sm_J(A,B)$ for the two type-1 fuzzy sets that are depicted in Fig. 2.23b.

[41]Please note that the use of a crisp number for the similarity of type-1 fuzzy sets is not being absolutely advocated for. Arguments can be given for using a type-1 fuzzy set similarity measure just as well as or for using a crisp number for similarity. The application may dictate which kind of measure is preferable. Of greater importance is that a similarity measure should satisfy some desirable properties, otherwise any kind of a measure between two type-1 fuzzy sets could be claimed to be a similarity measure. Four desirable properties for a type-1 fuzzy set similarity measure $sm(A,B)$ are (e.g., Mendel and Wu 2010, Ch. 4): (1) *Reflexivity*: $sm(A,B) = 1 \Leftrightarrow A = B$; (2) *Symmetry*: $sm(A,B) = sm(B,A)$; (3) *Transitivity*: If $C \le A \le B$ (Note: $A \le B$ if $\mu_A(x) \le \mu_B(x)$ for $x \in X$), where C is an arbitrary fuzzy set on domain X, then $sm(C,A) \ge sm(C,B)$; and (4) *Overlapping*: If $A \cap B \ne \emptyset$, then $sm(A,B) > 0$; otherwise, $sm(A,B) = 0$. $sm_J(A,B)$ satisfies these four properties.

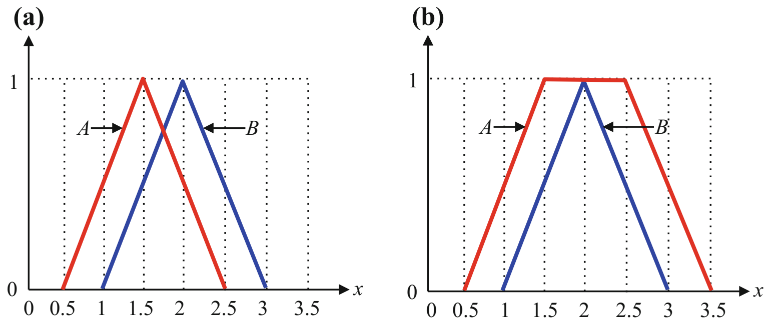


Fig. 2.23 Two type-1 fuzzy sets, A and B, for Exercise 2.44

2.45 *Subsethood* is also sometimes used in a rule-based fuzzy system, so this exercise explores subsethood for type-1 fuzzy sets. Subsethood is about set containment. Containment is dependent on the order of the two sets, A and B, i.e. A can be contained in B but B does not have to be contained in A, e.g. when $A = \{1, 2, 3\}$ and $B = \{1, 2, 3, 4, 5, 6\}$, $A \subset B$ but $B \not\subset A$. For crisp sets, it is only when $A = B$ that A is contained in B and B is contained in A. For fuzzy sets one knows that everything is a matter of degree; thus, for two type-1 fuzzy sets A and B, it is reasonable to define a *degree of subsethood*. As usual (in this book), crisp sets are our starting point.

As is stated in Nguyen and Kreinovich (2008): It is known that for two crisp sets A and B: (1) $A \cap B \subseteq A$ and (2) $A \subseteq B$ iff $A \cap B = A$ (create Venn diagrams to convince yourself of the truth of these). So, for crisp sets, to check whether A is a subset of B consider the ratio $|A \cap B|/|A|$ where $|\cdot|$ denotes the cardinality of $\cdot$ (see Exercise 2.43 about cardinality). In general this ratio is between 0 and 1, and it equals 1 if and only if A is a subset of B. The smaller the ratio the more there are elements from A which are not part of the intersection $A \cap B$ and thus not part of set B. Consequently, for crisp sets, this ratio can be viewed as a reasonable measure of the degree to which A is a subset of B (see, also, Kosko 1990, 1992).

Because there are many definitions of cardinality for a type-1 fuzzy set as well as the intersection of two type-1 fuzzy sets, there can be many ways to define the subsethood[42] between two type-1 fuzzy sets. This exercise focuses on what is

[42]Please note that the use of a crisp number for the subsethood of type-1 fuzzy sets is not being absolutely advocated for. Arguments can be given for using a type-1 fuzzy set subsethood measure just as well as or for using a crisp number for subsethood. The application may dictate which kind of measure is preferable. Of greater importance is that a subsethood measure should satisfy some desirable properties, otherwise any kind of a measure between two type-1 fuzzy sets could be claimed to be a subsethood measure. Three desirable properties for type-1 fuzzy set subsethood measure $ss(A, B)$ are (e.g., Mendel and Wu 2010, Ch. 4): (1) *Reflexivity*: $ss(A, B) = 1 \Leftrightarrow A \leq B$ (Note: $A \leq B$ if $\mu_A(x) \leq \mu_B(x)$ for $x \in X$); (2) *Transitivity*: If $C \leq A \leq B$, then $ss(A, C) \geq ss(B, C)$, where C is an arbitrary fuzzy set on domain X, or if $A \leq B$, then $ss(C, A) \leq ss(C, B)$ for any C;

arguably the most widely used definition of subsethood due to Kosko (1990) and denoted here as $ss_K(A, B)$. For a continuous universe of discourse,

$$ss_K(A, B) = \frac{\int_X \min(\mu_A(x), \mu_B(x))dx}{\int_X \mu_A(x)dx}$$

(a) Explain why $ss_K(A, B) \neq ss_K(B, A)$.
(b) What is the formula for $ss_K(A, B)$ for discrete universes of discourse?
(c) Compute $ss_K(A, B)$ for the two type-1 fuzzy sets that are depicted in Fig. 2.23a.
(d) Compute $ss_K(A, B)$ for the two type-1 fuzzy sets that are depicted in Fig. 2.23b.

References

Aisbett, J., J.T. Rickard, and D.G. Morgenthaler. 2010. Type-2 fuzzy sets as functions on spaces. *IEEE Trans on Fuzzy Systems* 18: 841–844.
Allendoerfer, C.B., and C.O. Oakley. 1955. *Principles of mathematics*. New York: McGraw-Hill.
Arabi, B.N., N. Kehtarnavaz, and C. Lucas. 2001. Restrictions imposed by the fuzzy extension of relations and functions. *Journal of Intelligent and Fuzzy Systems* 11: 9–22.
Bezdek, J.C. 1981. *Pattern recognition with fuzzy objective function algorithms*. New York: Plenum.
Bezdek, J., and S.K. Pal. 1992. *Fuzzy models for pattern recognition*. New York: IEEE Press.
Blanchard, N. 1982. Cardinal and ordinal theories about fuzzy sets. In *Fuzzy Information and Decision Process*, ed. M. M. Gupta and E. Sanchez, pp. 149–157, Amsterdam.
Bonissone, P. P. and K. S. Decker. 1986. Selecting uncertainty calculi and granularity: An experiment in trading off precision and complexity. In *Uncertainty in Artificial Intelligence*, ed. L. N. Kanal and J. F. Lemmer, pp. 217–247, Amsterdam.
Cheeseman, P. 1988. An inquiry into computer understanding. *Computational Intelligence* 4: 57–142 (with 22 commentaries/replies).
Cox, E.A. 1994. *The fuzzy systems handbook*. Cambridge, MA: AP Professional.
Cox, E. A. 1992. Fuzzy fundamentals. *IEEE Spectrum*, pp. 58–61, Oct. 1992.
Dubois, D., and H. Prade. 1980. *Fuzzy sets and systems: Theory and applications*. NY: Academic Press.
Dubois, D., and H. Prade. 1985. Fuzzy cardinality and the modeling of imprecise quantification. *Fuzzy Sets and Systems* 16: 199–230.
Dutta, P. H. Boruah and T. Ali. 2011. Fuzzy arithmetic with and without using the α-cut method: A comparative study. *International Journal of Latest Trends in Computing* 2(1): 99–107 (E-ISSN: 2045-5364).

(Footnote 42 continued)

and (3) *Overlapping*: If $A \cap B \neq \emptyset$, then $ss(A, B) > 0$; otherwise, $ss(A, B) = 0$. $ss_K(A, B)$ satisfies these three properties. The interested reader is referred to, e.g. Young (1996) and Fan et al. (1999).

Edwards, W.F. 1972. *Likelihood*. London: Cambridge Univ. Press.

Fan, J.L., W.X. Xie, and J. Pei. 1999. Subsethood measure: new definitions. *Fuzzy Sets and Systems* 106: 201–209.

Gottwald, S. 1980. A note on fuzzy cardinals. *Kybernetika* 16: 156–158.

He, Q., H.-X. Li, C.L.P. Chen, and E.S. Lee. 2000. Extension principles and fuzzy set categories. *Computers and Mathematics with Applications* 39: 45–53.

Horikawa, S., T. Furahashi and Y. Uchikawa. 1992. On fuzzy modeling using fuzzy neural networks with back-propagation algorithm. *IEEE Transaction on Neural Networks* 3: 801–806.

Jaccard, P. 1908. Nouvelles recherches sur la distribution florale. *Bulletin de la Societe de Vaud des Sciences Naturelles* 44: 223, 1908.

Jang, J.-S.R. 1992. Self-learning fuzzy controllers based on temporal back-propagation. *IEEE Transaction on Neural Networks* 3: 714–723.

Jang, L.-C., and D. Ralescu. 2001. Cardinality concepts for type-two fuzzy sets. *Fuzzy Sets and Systems* 118: 479–487.

Jang, J.-S. R., C-T. Sun and E. Mizutani. 1997. *Neuro-Fuzzy and Soft-Computing*, Prentice-Hall, Upper Saddle River, NJ.

Karnik, N.N., and J.M. Mendel. 2001. Operations on type-2 fuzzy sets. *Fuzzy Sets and Systems* 122: 327–348.

Karnik, N. N. and J. M. Mendel. 1998. *An Introduction to Type-2 Fuzzy Logic Systems*, USC-SIPI Report #418, Univ. of Southern Calif., Los Angeles, CA, June 1998. This can be accessed at: http://sipi.usc.edu/research; then choose "sipi technical reports/418.

Kaufmann, A. 1977. *Introduction a la theorie des sous-ensembles flous, complement et nouvelles applications*, vol. 4. Paris: Masson.

Klement, E. P. 1982. On the cardinality of fuzzy sets. *Proceedings of 6th European Meeting on Cybernetics and Systems Research*, Vienna, pp. 701–704.

Klir, G.J., and T.A. Folger. 1988. *Fuzzy sets, uncertainty, and information*. Englewood Cliffs, NJ: Prentice Hall.

Klir, G.J., and B. Yuan. 1995. *Fuzzy sets and fuzzy logic: Theory and applications*. Upper Saddle River, NJ: Prentice Hall.

Kosko, B. 1990. Fuzziness versus probability. *International Journal of General Systems* 17: 211–240.

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

Kosko. 1986. Fuzzy entropy and conditioning. *Information Sciences* 40: 165–174.

Kreinovich, V. 2008. Relations between interval computing and soft computing. In *Processing with Interval and Soft Computing*, ed. A. de Korvin and V. Kreinovich, 75–97. Springer, London.

Larsen, P. M. 1980. Industrial applications of fuzzy logic control. *International Journal Man-Machine Studies* 12: 3–10.

Laviolette, M., and J.W. Seaman Jr. 1994. The efficacy of fuzzy representations and uncertainty. *IEEE Transaction on Fuzzy Systems* 2: 4–15.

Lin, C.-T., and C.S.G. Lee. 1996. *Neural fuzzy systems: A neuro-fuzzy synergism to intelligent systems*. Upper Saddle River, NJ: Prentice-Hall PTR.

Lindley, B. Y. 1982. Scoring rules and the inevitability of probability. *International Statistical Review* 50: 1–26 (with 7 commentaries/replies).

De Luca, A., and S. Termini. 1972. A definition of non-probabilistic entropy in the setting of fuzzy sets theory. *Information and Computation* 20: 301–312.

Macvicar-Whelen, P. J. 1978. Fuzzy sets, the concept of height, and the hedge 'very'. *IEEE Transaction on Systems, Man, and Cybernetics* SMC-8: 507–511.

Mamdani, E.H. 1974. Applications of fuzzy algorithms for simple dynamic plant. *Proceedings of the IEEE* 121: 1585–1588.

Mendel, J.M. 1995a. Fuzzy logic systems for engineering: A tutorial. *IEEE Proceedings* 83: 345–377.

Mendel, J.M. 1995b. *Lessons in estimation theory for signal processing*. Prentice-Hall PTR, Englewood Cliffs, NJ: Communications and Control.

Mendel, J.M. 2007. Computing with words: Zadeh, Turing, Popper and Occam. *IEEE Computational Intelligence Magazine* 2: 10–17.

Mendel, J.M. 2015. Type-2 fuzzy sets and systems: A retrospective. *Informatik Spektrum* 38 (6): 523–532.

Mendel, J.M., H. Hagras, W.-W. Tan, W.W. Melek, and H. Ying. 2014. *Introduction to type-2 fuzzy logic control*. Hoboken, NJ: John Wiley and IEEE Press.

Mendel, J.M., and D. Wu. 2010. *Perceptual computing: Aiding people in making subjective judgments"*. Hoboken, NJ: Wiley and IEEE Press.

Nguyen, H.T. 1978. A note on the extension principle for fuzzy sets. *Journal of Mathematical Analysis and Applications* 64: 369–380.

Nguyen, H. T. and V. Kreinovich. 2008. Computing degrees of subsethood and similarity for interval-valued fuzzy sets: Fast algorithms. In *Proceedings of 9*th *International Conference on Intelligent Technologies in Tech'08*, pp. 47–55, Samui, Thailand, Oct. 2008.

Rajati, M.R., and J.M. Mendel. 2013. Novel weighted averages versus normalized sums in computing with words. *Information Sciences* 235: 130–149.

Rudin, W. 1966. *Real and complex analysis*. New York: Mc-Graw Hill.

Ruspini, E. 1969. A new approach to clustering. *Information Control* 15: 22–32.

Schmucker, K.S. 1984. *Fuzzy set, natural language computations, and risk analysis*. Rockville, MD: Computer Science Press.

Wang, L.-X. 1997. *A course in fuzzy systems and control*. Upper Saddle River, NJ: Prentice-Hall.

Wang, L.-X., and J.M. Mendel. 2016. Fuzzy opinion networks: A mathematical framework for the evolution of opinions and their uncertainties across social networks. *IEEE Transaction on Fuzzy Systems* 24: 880–905.

Wang, L.-X. and J. M. Mendel. 1992a. Fuzzy basis functions, universal approximation, and orthogonal least squares learning. *IEEE Trans. on Neural Networks* 3: 807–813.

Wang, L.-X. and J. M. Mendel. 1992b. Back-propagation of fuzzy systems as non-linear dynamic system identifiers. *Proceedingas of IEEE Intternational Conference on Fuzzy Systems,* 1409–1418, San Diego, CA.

Wei, S-H. and S-M. Chen. 2009. Fuzzy risk analysis based on interval-valued fuzzy numbers. *Expert Systems with Applications* 36: 2285–2299.

Wu, D., and J.M. Mendel. 2007. Uncertainty measures for interval type-2 fuzzy sets. *Information Sciences* 177: 5378–5393.

Wygralak, M. 1983. A new approach to the fuzzy cardinality of finite fuzzy sets. *Busefal* 15: 72–75.

Wygralak, M. 2003. *Cardinalities of fuzzy sets*. Heidelberg: Springer.

Yager, R.R. 1986. A characterization of the fuzzy extension principle. *Journal Fuzzy Sets and Systems* 18: 205–217.

Yager, R.R., and D.P. Filev. 1994. *Essentials of fuzzy modeling and control*. New York: Wiley.

Young, V.R. 1996. Fuzzy subsethood. *Fuzzy Sets and Systems* 77: 371–384.

Zadeh, L.A. 1965. Fuzzy sets. *Information and Control* 8: 338–353.

Zadeh, L.A. 1971. Similarity relations and fuzzy orderings. *Information Sciences* 3 (2): 177–200.

Zadeh, L.A. 1972. A fuzzy-set-theoretic interpretation of iinguistic hedges. *Journal of Cybernetics* 2: 4–34.

Zadeh, L.A. 1981. Possibility theory and soft data analysis. In *Mathematical frontiers of the social and policy sciences*, ed. L. Cobb, and R.M. Thrall, 69–129. CO: Westview Press, Boulder.

Zadeh, L. A. 1973. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transaction on Systems, Man, and Cybernetics* SMC-3: 28–44.

Zadeh, L. A. 1975. The concept of a linguistic variable and Its application to approximate reasoning–1. *Information Sciences* 8: 199–249.

Zadeh, L. A. 1999. From computing with numbers to computing with words—from manipulation of measurements to manipulation of perceptions. *IEEE Transaction on Circuits and Systems–I: Fundamental Theory and Applications*, 4: 105–119.

Zimmermann, H.J. 1991. *Fuzzy set theory and its applications*, 2nd ed. Boston, MA: Kluwer Academic Publ.

Uncertain Rule-Based Fuzzy Systems
Introduction and New Directions, 2nd Edition
Mendel, J.M.
2017, XXII, 684 p. 215 illus., 192 illus. in color.,
Hardcover
ISBN: 978-3-319-51369-0