2

FOUNDATIONS OF FUZZY APPROACHES

2.1 FROM CRISP TO FUZZY SETS

The mathematical foundations of fuzzy logic rest in *fuzzy set theory*, which can be thought of as a generalization of *classical set theory*. A familiarity with the novel notions, notations, and operations of fuzzy sets is useful in studying fuzzy logic principles and applications; acquiring it will be our main goal in this chapter.

Fuzziness is a property of language. Its main source is the imprecision involved in defining and using symbols. Consider, for example, the set of chairs in a room. In set theory the set of chairs may be established by pointing to every object in a room asking the question, *Is it a chair*? In classical set theory we are allowed to use only two answers: Yes or No. Let us code Yes as 1 and No as 0. Thus, our answers will be in the pair $\{0, 1\}$. If the answer is 1, an element belongs to a set; if the answer is 0, it does not. In the end we collect all the objects whose label is 1 and obtain the *set of chairs in a room*. Suppose, however, that we ask the question, *Which objects in a room may function as a chair*? Again we could point to every object and ask, *Could it function as a chair*? The answer here too could artificially be restricted to $\{0, 1\}$. Yet, the set of objects in a room that may *function* as a chair may include not only chairs but also desks, boxes, parts of the floor, and so on. It is a set not uniquely defined. It all depends on what we mean by the word *function*. Words like *function* have many shades of meaning and can be used in many different ways. Their meaning and use may vary with different persons, circumstances, and purposes; it depends on the specifics of a situation. We say therefore that the *set of objects that may function as a chair* is a *fuzzy set*, in the sense that we may not have crisply defined criteria for deciding membership into the set. Objects such as desks, boxes, and part of the floor may function as *chairs*, to a degree. It should be noted, however, that there is nothing fuzzy about the material objects themselves: Chairs, boxes, and desks are what they are. Fuzziness is a feature of their representation in a milieu of symbols and is generally a property of models, computational procedures, and language.

Let us now review some notions of classical set theory. *Classical sets* are crisply defined collections of distinct elements (numbers, symbols, objects, etc.), and for this reason we also call them *crisp sets*. The elements of all the sets under consideration in a given situation belong to an invariable, constant set, called the *universal set* or *universe* or more often the *universe of discourse*.[1] The fact that elements of a set A either belong or do not belong to a crisp set A can be formally indicated by the *characteristic function of* A, defined as

$$\chi_A(x) \equiv \begin{cases} 1 & \text{iff} \quad x \in A \\ 0 & \text{iff} \quad x \notin A \end{cases} \tag{2.1-1}$$

where the symbols $\in$ and $\notin$ denote that x *is* and *is not* a member of A, respectively, and iff is shorthand for "if and only if." The pair of numbers $\{0, 1\}$ is called the *valuation set*. Another way of writing equation (2.1-1) is

$$\chi_A(x): X \to \{0, 1\} \tag{2.1-2}$$

The notation of equation (2.1-2) is read as follows: *There exists a function* $\chi_A(x)$ *mapping every element of the set* X *(our universe of discourse) to the set* $\{0, 1\}$. It emphasizes that the characteristic function is a mechanism for mapping the set X to the valuation set $\{0, 1\}$. Important operations in crisp sets such as *union*, *intersection*, and *complementation* are familiar to us from elementary mathematics. They are usually represented through Venn diagrams but may also be expressed in terms of the characteristic function.

Fundamentally, sets are *categories*. Defining suitable categories and using operations for manipulating them is a major task of modeling and computation. From image recognition to measurement and control, the notion of *category*, or *set*, is essential in the definition of system variables, parameters, their ranges, and their interactions. The constraint to have a dual degree of membership to a set, an all-or-nothing, is a consequence of a desire to abstract a system description away from the multitude of intricacies and complexities that exist in reality and focus on factors of primary influence. Nevertheless, given our modern-day computational technologies, it may be unduly restrictive. This is particularly the case when it is desired to develop computer models *easily calibrated* to the specifics of a system and endowed with adaptive and self-organizing capabilities (Zadeh, 1973, 1988).

[1]The term *universe of discourse* is used in fuzzy logic; it comes from classical logic and describes the complete set of individual elements able to be referred to or quantified.

2.2 FUZZY SETS

As we saw in the previous section, in classical set theory there is a rather strict sense of membership to a set; that is, an element either *belongs* or *does not belong* to the set. In 1965 Lotfi A. Zadeh introduced *fuzzy sets*, where a more flexible sense of membership is possible (Zadeh, 1965). In fuzzy sets many degrees of membership are allowed. The degree of membership to a set is indicated by a number between 0 and 1—that is, a number in the interval $[0, 1]$. The point of departure for fuzzy sets is simply generalizing the valuation set from the pair of numbers $\{0, 1\}$ to all numbers found in $[0, 1]$. By expanding the valuation set we alter the nature of the characteristic function, now called *membership function* and denoted by $\mu_A(x)$. We no longer have *crisp* sets but instead have *fuzzy* sets. Since the interval $[0, 1]$ contains an infinity of numbers, infinite degrees of membership are possible. Thus, in view of equation (2.1-2) we say that *a membership function maps every element of the universe of discourse* X *to the interval* $[0, 1]$, and we formally write this mapping as

$$\mu_A(x): X \to [0, 1] \tag{2.2-1}$$

Equation (2.2-1) is a generalization of the mapping shown in equation (2.1-2). Membership functions are a simple yet versatile mathematical tool for indicating flexible membership to a set and, as we shall see, for modeling and quantifying the meaning of symbols. A question often asked by people beginning the study of fuzzy sets is, How are membership functions found? Membership functions may represent an individual's (subjective) notion of a vague class—for example, *objects in a room functioning as chairs*, *tall people*, *acceptable performance*, *small contribution to system stability*, *little improvement*, *big benefit*, and so on. In designing and operating controllers or automatic decision-making tools, for example, modeling such notions is a very important task. Membership functions may also be determined on the basis of statistical data or through the aid of neural networks. In Part III of this book we will look at the synergistic relation between neural networks and fuzzy logic toward this end (Kosko, 1992). At this point we can simply say that membership functions are primarily subjective in nature; this does not mean that they are assigned arbitrarily, but rather on the basis of application-specific criteria (Kaufmann, 1975; Dubois and Prade, 1980; Zimmermann, 1985).

There are two commonly used ways of denoting fuzzy sets. If X is a universe of discourse and x is a particular element of X, then a fuzzy set A defined on X may be written as a collection of ordered pairs

$$A = \{(x, \mu_A(x))\}, \qquad x \in X \tag{2.2-2}$$

where each pair $(x, \mu_A(x))$ is called a *singleton* and has x first, followed by its membership in A, $\mu_A(x)$. In crisp sets a singleton is simply the element x by itself. In fuzzy sets a singleton is two things: x and $\mu_A(x)$. For example, the set of *small integers*, A, defined (subjectively) over the universe of discourse of positive integers may be given by the collection of singletons

$$A = \{(1, 1.0), (2, 1.0), (3, 0.75), (4, 0.5), (5, 0.3), (6, 0.3), (7, 0.1), (8, 0.1)\}$$

Thus the fourth singleton from the left tells us that 4 belongs to A to a degree of 0.5. A singleton is also written as $\mu_A(x)/x$—that is, by putting membership first, followed by the marker "/" separating it from x.[2] Singletons whose membership to a fuzzy set is zero may be omitted. The *support set* of a fuzzy set A is the set of its elements that have membership function other than the trivial membership of zero.

An alternative notation, used more often than equation (2.2-2), explicitly indicates a fuzzy as the *union* of all $\mu_A(x)/x$ singletons—that is,

$$A = \sum_{x_i \in X} \mu_A(x_i)/x_i \tag{2.2-3}$$

The *summation* sign in equation (2.2-3) indicates the *union* of all singletons (the union operation in set theory is like "addition"). Equation (2.2-3) assumes that we have a *discrete universe of discourse*. In this alternative notation the set of *small integers* above may be written as

$$\begin{aligned} A &= \mu_A(1)/1 + \mu_A(2)/2 + \mu_A(3)/3 + \mu_A(4)/4 + \mu_A(5)/5 \\ &\quad + \mu_A(6)/6 + \mu_A(7)/7 + \mu_A(8)/8 \\ &= 1.0/1 + 1.0/2 + 0.75/3 + 0.5/4 + 0.3/5 + 0.3/6 + 0.1/7 + 0.1/8 \end{aligned}$$

For a continuous universe of discourse, we write equation (2.2-3) as

$$A = \int_X \mu_A(x)/x \tag{2.2-4}$$

where the *integral* sign in equation (2.2-4) indicates the *union* of all $\mu_A(x)/x$ singletons.[3] Consider, for example, the fuzzy set *small numbers* defined (subjectively) over the set of non-negative real numbers through a continuous

[2] It should be noted that "/" does not indicate "division"; it is merely a marker.

[3] Note that the integral sign is not the same as the integral sign of differential and integral calculus. It is used here in the sense that the integral sign is used in set theory—that is, to indicate the *sum* or *union* of individual *singletons*.

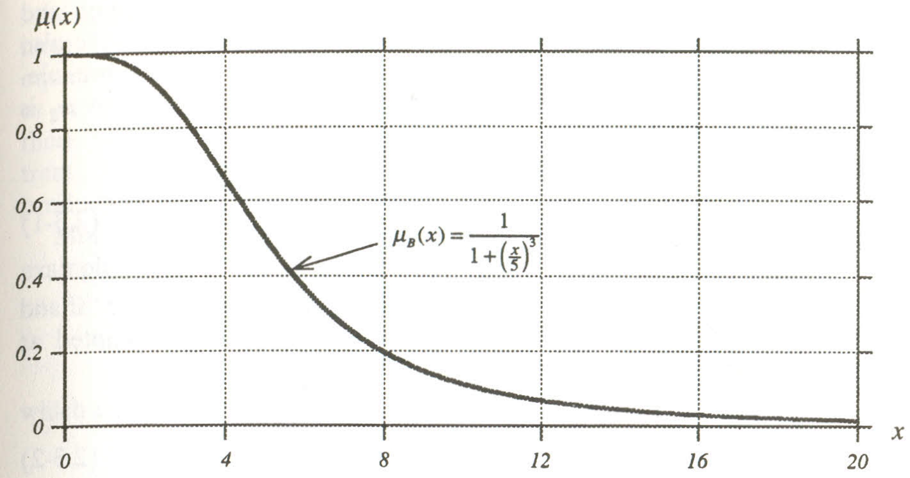


Figure 2.1 Zadeh diagram for the fuzzy set B = {*small numbers*}.

membership function $\mu_B(x)$ given by

$$\mu_B(x) = \frac{1}{1 + \left(\dfrac{x}{5}\right)^3} \tag{2.2-5}$$

Using the form of equation (2.2-4) the fuzzy set B is written as

$$B = \int_{x \geq 0} \mu_B(x)/x = \int_{x \geq 0} \left[\frac{1}{1 + \left(\dfrac{x}{5}\right)^3} \right] \Big/ x \tag{2.2-6}$$

The membership function of fuzzy set B is shown in Figure 2.1.[4] A graph like this is called a *Zadeh diagram.*

2.3 BASIC TERMS AND OPERATIONS

Many fuzzy set operations such as *intersection* and *union* are defined through the min ($\wedge$) and max ($\vee$) operators. *Min* and *max* are analogous to *product* ($\cdot$) and *sum* ($+$) in algebra (Dubois and Prade, 1980; Klir and Folger, 1988; Terano et al., 1992). Let us take a look at how they are used.

[4] Fuzzy sets are sometimes called *fuzzy subsets*, reflecting the fact that they are subsets of a larger set—that is, the *universe of discourse*. Although the term *fuzzy subsets* is factually correct, we will use the standard term *fuzzy set* for convenience.

First, min ($\wedge$) and max ($\vee$) may be used to select the minimum and maximum of two elements—for example, $2 \wedge 3 = 2$, or $2 \vee 3 = 3$. We also write $\min(2,3) = 2$, or $\max(2,3) = 3$. Formally, the minimum of two elements μ_1 and μ_2 denoted either as $\min(\mu_1, \mu_2)$, $\wedge(\mu_1, \mu_2)$, or $\mu_1 \wedge \mu_2$ is defined as

$$\mu_1 \wedge \mu_2 = \min(\mu_1, \mu_2) \equiv \begin{cases} \mu_1 & \text{iff} \quad \mu_1 \leq \mu_2 \\ \mu_2 & \text{iff} \quad \mu_1 > \mu_2 \end{cases} \tag{2.3-1}$$

where, the "$\equiv$" symbol means "by definition" and iff is shorthand for "if and only if." Similarly the maximum of two elements μ_1 and μ_2, denoted as $\max(\mu_1, \mu_2)$ or $\mu_1 \vee \mu_2$, is defined as

$$\mu_1 \vee \mu_2 = \max(\mu_1, \mu_2) \equiv \begin{cases} \mu_1 & \text{iff} \quad \mu_1 \geq \mu_2 \\ \mu_2 & \text{iff} \quad \mu_1 < \mu_2 \end{cases} \tag{2.3-2}$$

Second, min ($\wedge$) and max ($\vee$) may operate on an entire set, selecting the least element (called *infimum* in mathematical analysis) or the greatest element (called *supremum*) of the set. For example, $\wedge$ (0.01, 0.33, 0.44, 0.999) = 0.01 and $\vee$ (0.01, 0.33, 0.44, 0.999) = 0.999. Formally we write this as

$$\mu = \wedge A = \inf A \tag{2.3-3}$$

and

$$\mu = \vee A = \sup A \tag{2.3-4}$$

where μ is an element of A—that is, $\mu \in A$.

In addition, min ($\wedge$) and max ($\vee$) may be used as *functions* operating on single elements or on entire sets, for example, to find the smallest element μ out of a list of elements $(\mu_1, \mu_2, \ldots, \mu_m)$—that is,

$$\mu = \wedge(\mu_1, \mu_2, \ldots, \mu_m) \tag{2.3-5}$$

which is the same as

$$\mu = \mu_1 \wedge \mu_2 \wedge \cdots \wedge \mu_m \tag{2.3-6}$$

We sometimes use a shorthand notation for equations (2.3-5) and (2.3-6) and write them as

$$\mu = \bigwedge_{k=1}^{m} (\mu_k) \tag{2.3-7}$$

This notation is analogous to finite *product* notation in algebra (or finite *summation* when $\vee$ is used). There is in fact a more general analogy between min and max and the operations of *multiplication* and *addition*. They both have the same properties of associativity and distributivity, and thus in equations that involve min and max we may employ them in the same manner as *multiplication* ($\cdot$) and *addition* ($+$). We will see an interesting example of these properties in the composition of fuzzy relations (Chapter 3), where we treat composition as *matrix multiplication* with ($\wedge$) and ($\vee$) in place of product ($\cdot$) and sum ($+$).

Min ($\wedge$) and max ($\vee$) can also operate on a collection of sets as for example in

$$A = \wedge(A_1, A_2, \ldots, A_m) \tag{2.3-8}$$

which can be succinctly written as

$$A = \bigwedge_{k=1}^{m} (A_k) \tag{2.3-9}$$

Using primarily min ($\wedge$) and max ($\vee$), a number of useful notions and operations involving fuzzy sets can be defined.[5]

Empty Fuzzy Set

A fuzzy set A is called *empty* (denoted as $A = \varnothing$) if its membership function is zero everywhere in its universe of discourse X—that is,

$$A \equiv \varnothing \qquad \text{if } \mu_A(x) = 0, \forall x \in X \tag{2.3-10}$$

where "$\forall x \in X$" is shorthand notation indicating "for any element x in X."

Normal Fuzzy Set

A fuzzy set is called *normal* if there is at least one element x_0 in the universe of discourse where its membership function equals one—that is,

$$\mu_A(x_0) = 1 \tag{2.3-11}$$

More than one element in the universe of discourse can satisfy equation (2.3-11).[6]

[5]These operations can also be defined in terms of *T-norms* (see Appendix).

[6]It should be noted that the term *normal* does not refer to the area under the curve of the membership function. It simply means what the definition says: At least one point, maybe more, needs to have full membership value.

Equality of Fuzzy Sets

Two fuzzy sets are said to be *equal* if their membership functions are equal everywhere in the universe of discourse—that is,

$$A \equiv B \qquad \text{if } \mu_A(x) = \mu_B(x) \tag{2.3-12}$$

Union of Two Fuzzy Sets

The *union* of two fuzzy sets A and B defined over the same universe of discourse X is a new fuzzy set $A \cup B$ also on X, with membership function which is the maximum of the grades of membership of every x to A and B—that is,

$$\mu_{A \cup B}(x) \equiv \mu_A(x) \vee \mu_B(x) \tag{2.3-13}$$

The *union* of two fuzzy sets is related to the logical operation of *disjunction* (OR) in fuzzy logic. Equation (2.3-13) can be generalized to any number of fuzzy sets over the same universe of discourse.

Intersection of Fuzzy Sets

The *intersection* of two fuzzy sets A and B is a new fuzzy set $A \cap B$ with membership function which is the minimum of the grades of every x in X to the sets A and B, i.e.,

$$\mu_{A \cap B}(x) \equiv \mu_A(x) \wedge \mu_B(x) \tag{2.3-14}$$

The *intersection* of two fuzzy sets is related to *conjunction* (AND) in fuzzy logic. The definition of *intersection* in (2.3-14) can be generalized to any number of fuzzy sets over the same universe of discourse.

Complement of a Fuzzy Set

The *complement* of a fuzzy set A is a new fuzzy set, $\bar{A}$, with membership function

$$\mu_{\bar{A}}(x) \equiv 1 - \mu_A(x) \tag{2.3-15}$$

Fuzzy set *complementation* is equivalent to *negation* (NOT) in fuzzy logic.

Product of Two Fuzzy Sets

The product of two fuzzy sets A and B defined on the same universe of discourse X is a new fuzzy set, $A \cdot B$, with membership function that equals the algebraic product of the membership functions of A and B,

$$\mu_{A \cdot B}(x) \equiv \mu_A(x) \cdot \mu_B(x) \tag{2.3-16}$$

The product of two fuzzy sets can be generalized to any number of fuzzy sets on the same universe of discourse.

Multiplying a Fuzzy Set by a Crisp Number

We can multiply the membership function of a fuzzy set A by the crisp number a to obtain a new fuzzy set called *product* $a \cdot A$. Its membership function is

$$\mu_{aA}(x) \equiv a \cdot \mu_A(x) \tag{2.3-17}$$

The operations of multiplication and raising a fuzzy set to a power that we see next are useful for modifying the meaning of linguistic terms (Zadeh, 1975).

Power of a Fuzzy Set

We can raise fuzzy set A to a power α (positive real number) by raising its membership function to α. The α *power* of A is a new fuzzy set, A^α, with membership function

$$\mu_{A^\alpha}(x) \equiv [\mu_A(x)]^\alpha \tag{2.3-18}$$

Raising a fuzzy set to the second power is usually taken to be equivalent to linguistically changing it through the modifier *VERY* (Zadeh, 1983) (see Chapter 5). Thus the square of the membership function of B = {*small numbers*} in Figure 2.1 is taken to represent the fuzzy set B^2 = {*VERY small numbers*}.

Raising a fuzzy set to the second power is a particularly useful operation and therefore has its own name. It is called *concentration* or *CON*. Taking the square root of a fuzzy set is called *dilation* or *DIL* (an operation useful for representing analytically the linguistic modifier *MORE OR LESS*).

Example 2.1 Union, Intersection, and Complement of Fuzzy Sets. Consider the Zadeh diagram of fuzzy sets A and B shown in Figure 2.2*a* and defined by membership functions

$$\mu_A(x) = \frac{1}{1 + 0.3(x-8)^2} \quad \text{and} \quad \mu_B(x) = \frac{1}{1 + \left(\frac{x}{5}\right)^3} \tag{E2.1-1}$$

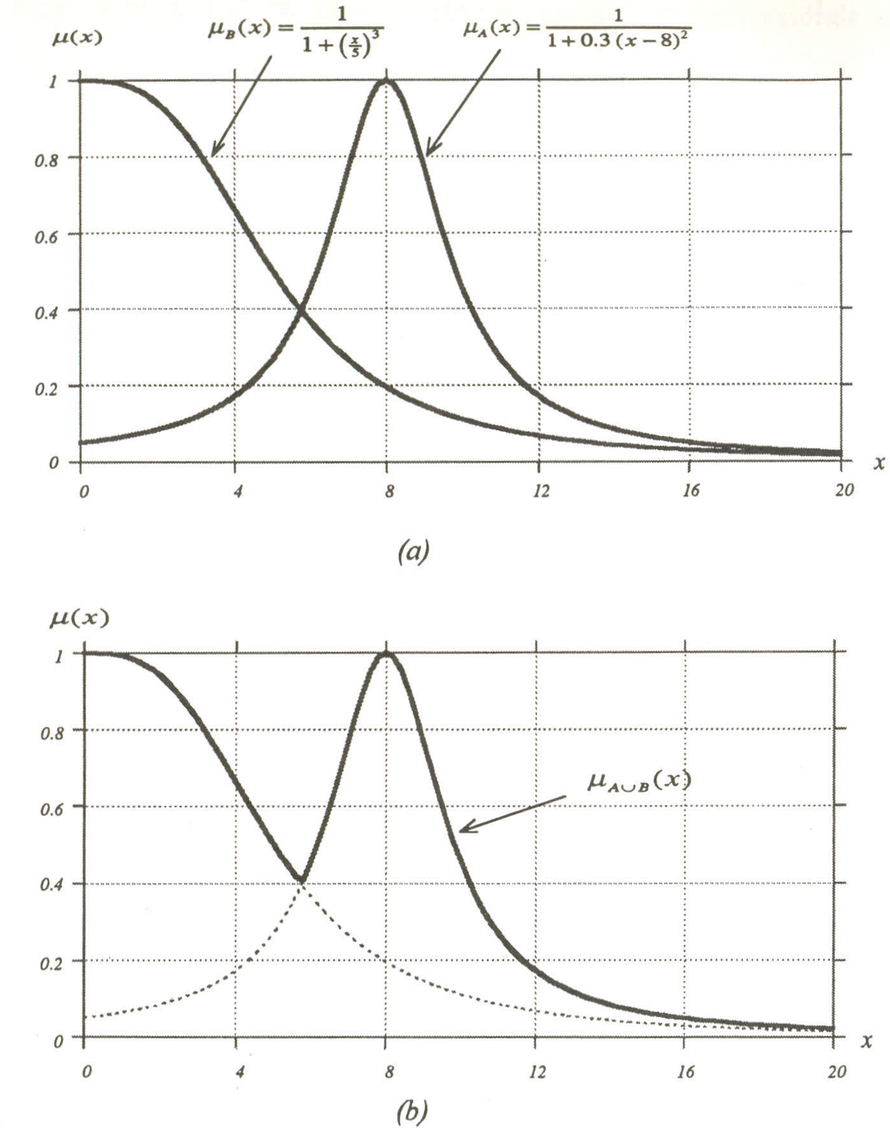


Figure 2.2 Zadeh diagram for (*a*) fuzzy sets *A* and *B* and (*b*) their *union* in Example 2.1.

Fuzzy set A may be thought of as defining the set of numbers "*about* 8," and fuzzy set B may be thought of as defining "*small numbers*." We take numbers between 0 and 20 to be the universe of discourse and, would like to find the *union* and *intersection* of A and B and the *complement* of B.

The membership function of the *union* of fuzzy sets A and B is the maximum grade of membership of each element x of the universe of

discourse to either A or B in accordance with equation (2.3-13). Figure 2.2*b* shows the membership function of the *union* $A \cup B$. The interpretation of $A \cup B$ is "*about* 8 *OR small number*." Similarly the membership function of the *intersection* of fuzzy sets A and B, shown in Figure 2.3*a*, represents the new fuzzy set "*about* 8 *AND small number*." We observe that although the union of A and B is a *normal* fuzzy set, the intersection shown in Figure 2.3*a* is not, because fuzzy set $A \cap B$ has no point in the universe of

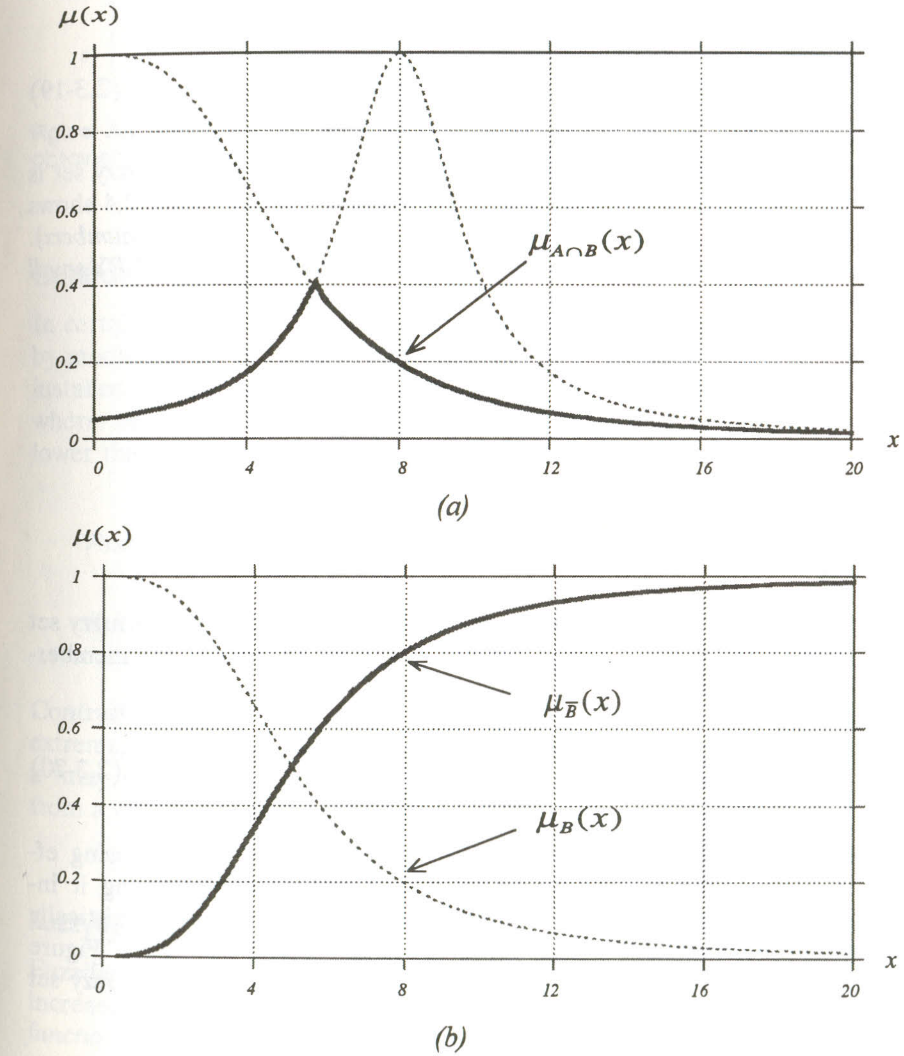


Figure 2.3 Zadeh diagram for (*a*) the *intersection* of fuzzy sets *A* and *B* and (*b*) the *complement* of *B* in Example 2.1.

discourse with grade of membership equal to 1. The complement of fuzzy set B is a new fuzzy set with membership function given by equation (2.3-15). Figure 2.3*b* shows the membership function of the *complement* $\overline{B}$. The complement $\overline{B}$ represents the logical negation (*NOT*) of B—that is, the set "*NOT small numbers.*" □ [7]

Concentration

The *concentration* of a fuzzy set A defined over a universe of discourse, X, is denoted as $CON(A)$ and it is a new fuzzy set with membership function given by

$$\mu_{CON(A)}(x) \equiv (\mu_A(x))^2 \tag{2.3-19}$$

As we said in the previous paragraph, squaring or *concentrating* a fuzzy set is equivalent to linguistically modifying it by the term *VERY*. Figure 2.4 shows the concentration operation applied to the fuzzy set B = {*small numbers*}. The membership function of the new fuzzy set $CON(B) = B^2$ = {*VERY small numbers*} is

$$\mu_{CON(B)}(x) = (\mu_B(x))^2 = \frac{1}{\left[1 + \left(\frac{x}{5}\right)^3\right]^2}$$

Dilation

The *dilation* of a fuzzy set A, denoted as $DIL(A)$, produces a new fuzzy set in X, with membership function defined as the square root of the membership function of A—that is,

$$\mu_{DIL(A)}(x) \equiv \sqrt{\mu_A(x)} \tag{2.3-20}$$

Dilation (*DIL*) and *concentration* (*CON*) are operations with opposing effects. Concentrating a fuzzy set reduces its fuzziness while dilating it increases its fuzziness. The *dilation* operation corresponds to linguistically modifying the meaning of a fuzzy set by the term "*MORE OR LESS.*" Figure 2.4 shows the dilation of B = {*small numbers*}, resulting in a new fuzzy set $DIL(B) = B^{1/2}$ = {*MORE OR LESS small numbers*}.

[7] Here and throughout this book, the end of an example is indicated by the symbol "□."

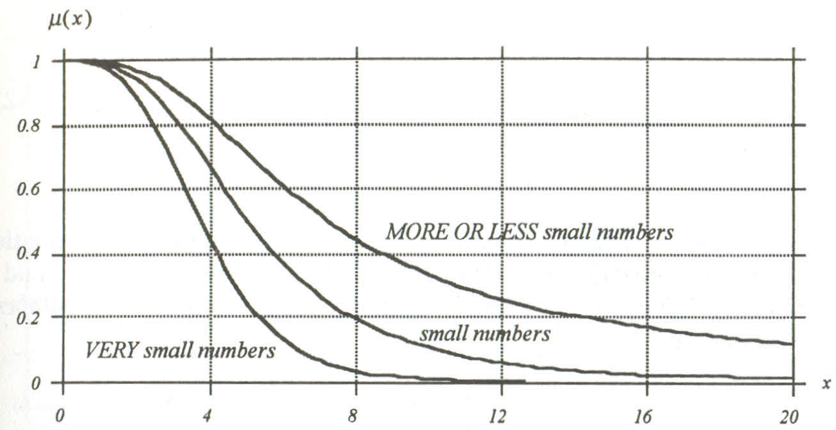


Figure 2.4 The fuzzy sets *VERY small numbers* and *MORE OR LESS small numbers* obtained by *concentrating* and *dilating* the fuzzy set *small numbers*.

Contrast Intensification

In certain applications it is desirable to control the *fuzziness* of a fuzzy set A by modifying the contrast between low and high grades of membership. For instance, we may want to increase the membership function on that part of A where membership values are higher than 0.5, and decrease it for values lower than 0.5. We define the *contrast intensification of* A as

$$\begin{aligned} \mu_{INT(A)}(x) &\equiv 2[\mu_A(x)]^2 && \text{for} \quad 0 \le \mu_A(x) \le 0.5 \\ \mu_{INT(A)}(x) &\equiv 1 - 2[1 - \mu_A(x)]^2, && \text{for } 0.5 \le \mu_A(x) \le 1.0 \end{aligned} \tag{2.3-21}$$

Contrast intensification may be repeatedly applied to a fuzzy set. In the extreme, when the maximum possible contrast is achieved we no longer have a fuzzy set. We are back to a crisp set. The opposite effect—that is, going from a crisp set to fuzzy set—may be achieved through *fuzzification.*

Fuzzyfication

Fuzzification is used to transform a crisp set into a fuzzy set or simply to increase the fuzziness of a fuzzy set. For fuzzification we use a *fuzzyfier function* F that controls the fuzziness of a set. F may be one or more simple parameters. For instance, consider the fuzzy set A that describes *large* *numbers*. We define it (subjectively) through the membership function

$$\mu_{large\ numbers}(x) = \frac{1}{1 + \left(\dfrac{x}{F_2}\right)^{-F_1}} \tag{2.3-22}$$

where x is any positive real number. The membership function in equation (2.3-22) has two fuzzifying parameters: an *exponential fuzzyfier*, F_1, and a *denominational fuzzyfier*, F_2. Through them the fuzzy set $A = \{large\ numbers\}$ can be written as

$$A \equiv \int_X \left[\frac{1}{1 + \left(\dfrac{x}{F_2}\right)^{-F_1}} \right] \Big/ x \tag{2.3-23}$$

The membership function inside the brackets of equation (2.3-23) can be adjusted when needed in order to better represent the meaning of the term *large numbers*. Consider the case when we fix the value of denominational fuzzifier as $F_2 = 50$ and vary the exponential fuzzyfier F_1. The result is a family of fuzzy sets with decreasing fuzziness as F_1 increases. Figure 2.5 shows membership functions that result from such a variation. Note that when F_1 becomes very large, the set A appears almost like a crisp set. The effect of varying the denominational fuzzyfier F_2 while keeping the exponen-

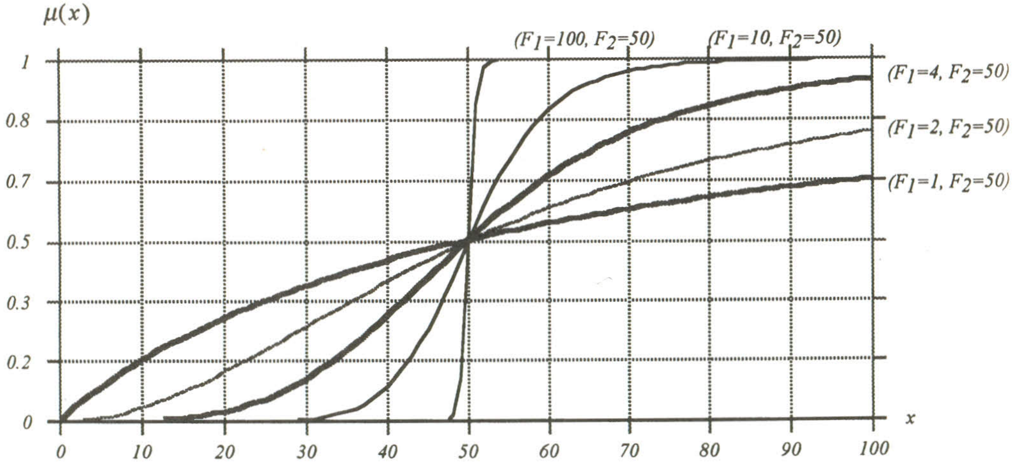


Figure 2.5 The effect of varying the exponential fuzzifier F_1 while keeping the denominational fuzzifier F_2 constant in fuzzifying the set A.

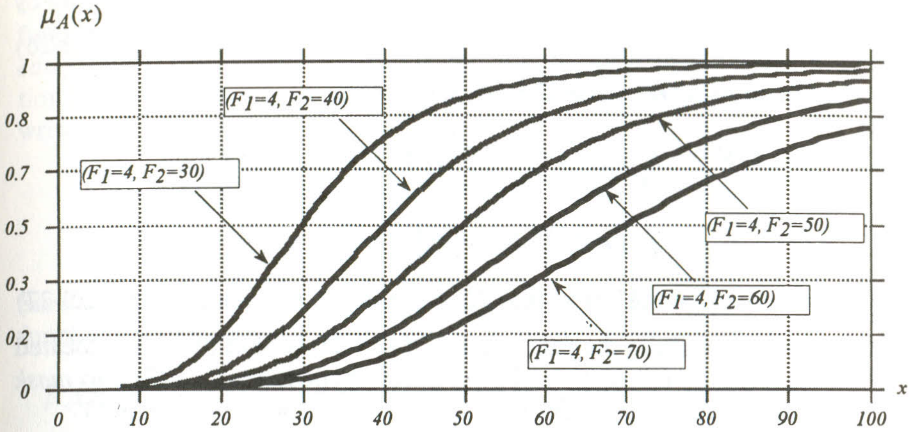


Figure 2.6 The effect of varying the denominational fuzzifier F_2 while keeping the exponential fuzzifier F_1 constant in fuzzifying the set A.

tial fuzzyfier at $F_1 = 4$ is shown in Figure 2.6. Varying F_2 results primarily in translating the membership function left and right, and to a lesser extent it affects the fuzziness of A. Such fuzzifiers are often used in fuzzy pattern recognition and image analysis in defining, for instance, the meaning of the words *vertical*, *horizontal*, and *oblique lines* (Pal and Majumder, 1986).

Fuzzification may be used more systematically by associating a *fuzzyfier F* with another function, namely a *fuzzy kernel*, $K(x)$, which is the fuzzy set that results from the application of F to a singleton x. This is often done in control applications where the input to an on-line control or diagnostic system comes from sensors and is therefore crisp, usually a real number. In order to use it in fuzzy algorithms (see Chapters 5 and 6), it is often necessary to convert a crisp number to a fuzzy set, a step known as *fuzzification*. As a result of the application of K to a fuzzy set A, we have

$$F(A;K) = \int_X \mu_A(x) \cdot \mu_{K(x)}(x)/x \tag{2.3-24}$$

where $F(A;K)$ is a fuzzy set that results from changing the fuzziness of A in accordance with K. The *fuzzy kernel* $K(x)$ is simply a fuzzy set imposed on a singleton. It functions as a "mask" that covers the singleton to produce a fuzzy set. For example, suppose that we have the universe of discourse $X = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ and a fuzzy kernel $K(x)$ that centers a triangular fuzzy set around 5 given by

$$K(5) = 0.33/3 + 0.67/4 + 1.0/5 + 0.67/6 + 0.33/7 \tag{2.3-25}$$

with all other elements of the universe of discourse having trivial (zero) membership. Now suppose that we have the value of 3, which may be a crisp measurement taken at a certain time. We write it as a singleton A given by

$$A = \mu_A(3)/3 = 1.0/3 \tag{2.3-26}$$

We fuzzify A using equation (2.3-24) as follows:

$$\begin{aligned} F(A;K) &= \int_X \mu_A(x) \cdot \mu_{K(x)}(x)/x \\ &= \int_X [\mu_A(3) \cdot \mu_{K(3)}(x)]/x \\ &= 0.33/1 + 0.67/2 + 1.0/3 + 0.67/4 + 0.33/5 \end{aligned} \tag{2.3-27}$$

which results in shifting the fuzzy kernel of (2.3-25) so that its peak is located at the singleton '3'. In other words, the effect of equations (2.3-27) is to *mask* the crisp value '3' by the fuzzy set $K(5)$, shifting its peak from '5' to '3'.

2.4 PROPERTIES OF FUZZY SETS

Fuzzy set properties are useful in performing operations involving membership functions. The properties we list here are valid for crisp and fuzzy sets as well, but some of them are specific to fuzzy sets only; more detailed treatment of properties may be found in Dubois and Prade (1980) and in Klir and Folger (1988). Consider sets A, B, C defined over a common universe of discourse X. We indicate the complement of a set by a bar over it. The following properties are true:

Double Negation Law:	$\overline{(\overline{A})} = A$	(2.4-1)
Idempotency:	$A \cup A = A$ $A \cap A = A$	(2.4-2)
Commutativity:	$A \cap B = B \cap A$ $A \cup B = B \cup A$	(2.4-3)
Associative Property:	$(A \cup B) \cup C = A \cup (B \cup C)$ $(A \cap B) \cap C = A \cap (B \cap C)$	(2.4-4)
Distributive Property:	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	(2.4-5)
Absorption:	$A \cap (A \cup B) = A$ $A \cup (A \cap B) = A$	(2.4-6)
De Morgan's Laws:	$\overline{A \cup B} = \overline{A} \cap \overline{B}$ $\overline{A \cap B} = \overline{A} \cup \overline{B}$	(2.4-7)

In fuzzy sets all these properties can be expressed using the membership function of the sets involved and the definitions of *union*, *intersection*, and *complement*. For example, consider the *associative property* given by equations (2.4-4). In terms of membership functions the associative property is written as

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

Similarly, the *distributive property*, equations (2.4-5), in terms of membership functions is written as

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

De Morgan's law, equation (2.4-7), is written as

$$\overline{\mu_A(x) \vee \mu_B(x)} = \mu_{\overline{A}}(x) \wedge \mu_{\overline{B}}(x)$$

where the bar over the membership functions indicates that we take the complement. *De Morgan's law* says that the intersection of the complement of two fuzzy sets equals the complement of their union; in terms of membership functions, this is the same as saying that the minimum of two membership functions equals the complement of their maximum. There are also some properties generally not valid for fuzzy sets (although valid in crisp sets), such as the *law of contradiction*,

$$A \cap \overline{A} \neq \varnothing \tag{2.4-8}$$

and the *law of the excluded middle*,

$$A \cup \overline{A} \neq X \tag{2.4-9}$$

The law of the excluded middle in crisp sets states that the union of a set with its complement results in the universe of discourse. This is generally not true in fuzzy sets. A property unique to fuzzy sets is

$$A \cap \varnothing = \varnothing \tag{2.4-10}$$

Equation (2.4-10) says that the intersection of a fuzzy set with the empty set —that is, a set with a membership function equal to zero everywhere on the universe of discourse—is also the empty set. In terms of membership functions equation (2.4-10) is written as

$$\mu_A(x) \wedge 0 = 0$$

Also, the union of a fuzzy set A with the empty set, $\varnothing$, is A itself; that is, $A \cup \varnothing = A$ or, equivalently, $\mu_A(x) \vee 0 = \mu_A(x)$. The intersection of a fuzzy set A with the universe of discourse is the fuzzy set A itself; that is, $A \cap X = A$ or, equivalently, $\mu_A(x) \wedge 1 = \mu_A(x)$. The union of a fuzzy set A with the universe of discourse X is the universe of discourse; that is, $A \cup X = X$, which, in terms of the membership function, is written as $\mu_A(x) \vee 1 = 1$. The universe of discourse may be viewed as a fuzzy set whose membership function equals 1 everywhere; that is, $\mu_X(x) = 1$ for all x in X.

2.5 THE EXTENSION PRINCIPLE

While fuzzification operations such as the ones we saw in Section 2.3 are useful for fuzzifying individual sets or singletons, more general mathematical expressions may also be fuzzified when the quantities they involve are fuzzyfied. For example, the output of arithmetic operations when their arguments are fuzzy sets becomes also a fuzzy quantity. The *extension principle* is a mathematical tool for extending crisp mathematical notions and operations to the milieu of fuzziness. It provides the theoretical warranty that fuzzifying the parameters or arguments of a function results in computable fuzzy sets. It is an important principle, and we will use it on several occasions, particularly in conjunction with fuzzy relations (Chapter 3) and fuzzy arithmetic (Chapter 4). We give here an informal heuristic description of the extension principle; detailed formulations may be found in (Zadeh (1975), and in Dubois and Prade (1980).

Suppose that we have a function f that maps elements $x_1, x_2, \ldots, x_n$ of a universe of discourse X to another universe of discourse Y—that is,

$$\begin{aligned} y_1 &= f(x_1) \\ y_2 &= f(x_2) \\ &\cdots \\ y_n &= f(x_n) \end{aligned} \qquad (2.5\text{-}1)$$

Now suppose that we have a fuzzy set A defined on $x_1, x_2, x_3, \ldots, x_n$ (the input to the function f). A is given by

$$A = \mu_A(x_1)/x_1 + \mu_A(x_2)/x_2 + \cdots + \mu_A(x_n)/x_n \qquad (2.5\text{-}2)$$

We then ask the question, If the input to our function f becomes fuzzy—for example, the set A of equation (2.5-2)—what happens to the output? Is the output also fuzzy? In other words, is there an output fuzzy set B that can be computed by inputting A to f. Well, the extension principle tells us that there is indeed such an output fuzzy set B and that it is given by

$$B = f(A) = \mu_A(x_1)/f(x_1) + \mu_A(x_2)/f(x_2) + \cdots + \mu_A(x_n)/f(x_n) \qquad (2.5\text{-}3)$$

where every single image of x_i under f—that is, $y_i = f(x_i)$—becomes fuzzy to a degree $\mu_A(x_i)$. Recalling that functions are generally *many-to-one* mappings, it is conceivable that several x's may map to the same y. Thus for a certain y_0 we may have more than one x: Let us say that both x_2 and x_{13} in (2.5-1) are mapping to y_0. Hence, we have to decide which of the two membership values, $\mu_A(x_2)$ or $\mu_A(x_{13})$, we should take as the membership value of y_0. The extension principle says that the *maximum* of the membership values of these elements in the fuzzy set A ought to be chosen as the grade of membership of y_0 to the set B—that is,

$$\mu_B(y_0) = \mu_A(x_2) \vee \mu_A(x_{13}) \qquad (2.5\text{-}4)$$

If, on the other hand, no element x in X is mapped to y_0—that is, no inverse image of y_0 exists—then the membership value of the set B at y_0 is zero. Having accounted for these two special cases (many x's mapping to the same y and no inverse image for a certain y), we can compute the set B—that is, the grades of membership of elements y in Y produced by the mapping $f(A)$—using equation (2.5-3).

In a more general case where we have several variables, $u, v, \ldots, w$, from different universes of discourse $U, V, \ldots, W$ and m different fuzzy sets $A_1, A_2, \ldots, A_m$ defined on the product space $U \times V \times \cdots \times W$, the multivariable function, $y = f(u, v, \ldots, w)$, may also be used to fuzzify the space Y through the extension principle. In this case, the grade of membership of any y equals the minimum of the membership values of $u, v, \ldots, w$ in $A_1, A_2, \ldots, A_m$, respectively. The membership function of B is given by

$$\mu_B(y) = \int_{U \times V \times \cdots \times Y} \left[\mu_{A_1}(u) \wedge \mu_{A_2}(v) \wedge \cdots \wedge \mu_{A_m}(w) \right] / f(u, v, \ldots, w) \qquad (2.5\text{-}5)$$

where there is also a max ($\vee$) operation implicit in the union operation [the integral sign in equation (2.5-5) indicates a union ($\vee$) operation]. The max operation is performed over all $u, v, \ldots, w$ such that $y = f(u, v, \ldots, w)$. This is indicated by the union over the product space $U \times V \times \cdots \times W$ of all the universes on which the m-tuples $u, v, \ldots, w$ are defined under the integral sign. If the inverse image does not exist, then the membership function is simply zero.

In many engineering applications, the interpretation of numerical data may not be precisely known. We consider this type of data to be fuzzy. Using the extension principle, it is quite possible to adapt ordinary algorithms, which are used with precise data, to the case where the data are fuzzy. Example 2.2 is a mathematical illustration of the extension principle.

Example 2.2 Using the Extension Principle. As an illustration of how the extension principle may be used, consider the function f that maps points from the x axis to y axis in the Cartesian plane according to the equation

$$y = f(x) = \sqrt{1 - \frac{x^2}{4}} \tag{E2.2-1}$$

Figure 2.7a shows the function y of equation (E2.2-1). It is the upper half of an ellipse located on the center of the plane with major axis, $a = 2$, and minor axis (height), $b = 1$. The general equation of the ellipse shown in Figure 2.7a is

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{E2.2-2}$$

In our case with $a = 2$ and $b = 1$, equation (E2.2-2) becomes

$$\frac{x^2}{4} + y^2 = 1 \tag{E2.2-3}$$

Equation (E2.2-1) is one of the two solutions of equation (E2.2-3).

Now suppose that we define a fuzzy set A on X as shown in Figure 2.7b: We fuzzify the x's of equation (E2.2-1) by specifying a grade of membership $\mu_A(x)$ for each x to fuzzy set A—that is, $\mu_A(x) = \frac{1}{2}|x|$ and

$$A = \int_{-2 \le x \le 2} \left[\tfrac{1}{2}|x|\right]/x \tag{E2.2-4}$$

where $|x|$ is the absolute value of x, and we limit the support of A between -2 and $+2$ as indicated by the limits under the *integration sign* (*union*) of equation (E2.2-4).

Having the x values fuzzyfied by the fuzzy set A, we want to know the effect of fuzzification on y. The extension principle tells us that the fuzziness of A will be extended to y as well. In other words, we will have a fuzzy set B on Y derived by equations (2.5-3) or (2.5-5). To avoid the case where more than one x will map to the same y, we consider first the function f in the first quadrant of the plane (where both x and y are positive). Later we will look at the entire function. The fuzzy set, B, defined on Y is

$$B = f(A) = \int_Y \mu_B(y)/y \tag{E2.2-5}$$

We need to find $\mu_B(y)$ in equation (E2.2-5). In terms of the membership function of A and according to the extension principle, equation (2.5-3), the

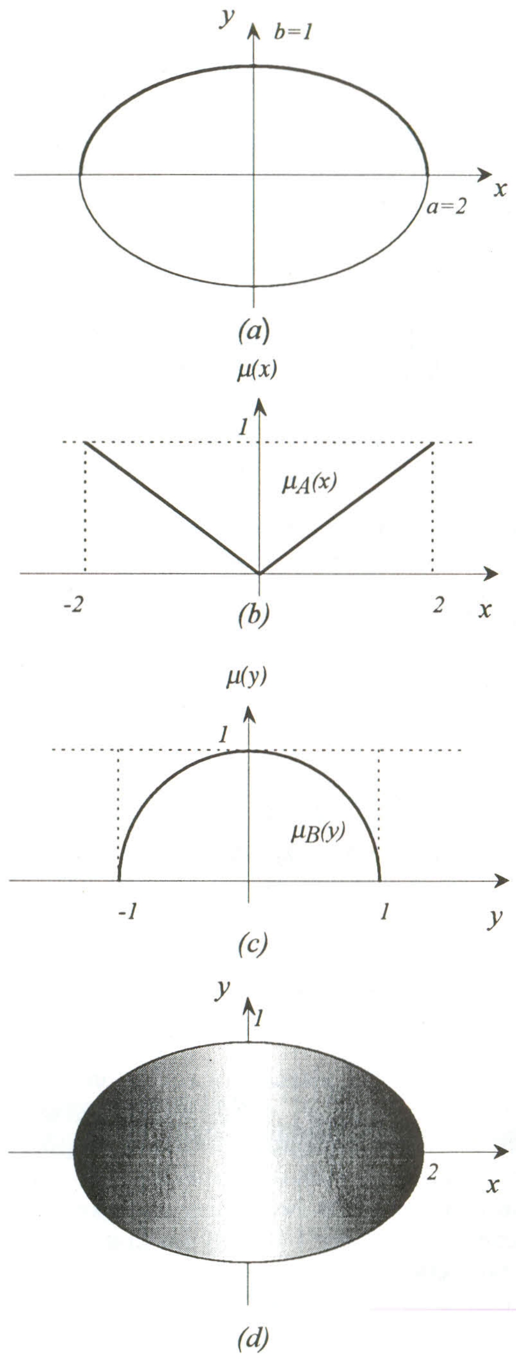


Figure 2.7 Graphs for Example 2.2. (*a*) The function *y*, which is the upper part of the ellipse shown. (*b*) The membership function of the set *A*. (*c*) The membership function of *B*. (*d*) Fuzzifying the interior of the ellipse.

set B will be

$$B = f(A) = \int_Y \mu_A(x)/f(x) \tag{E2.2-6}$$

Of course we want to transform the x variable to y in equation (E2.2-6) since the union (integration) is formed with respect to Y, the universe of discourse for B. We use equation (E2.2-1) to solve for x:

$$x = 2\sqrt{1 - y^2} \tag{E2.2-7}$$

Then we substitute (E2.2-7) in (E2.2-6), noting that $f(x) = y$ and that $\mu_A(x)$ is given by (E2.2-4). Thus we obtain the fuzzy set B:

$$B = \int_{0 \le y \le 1} \sqrt{1 - y^2} \Big/ y \tag{E2.2-8}$$

Now if we consider negative values for x as well, we would have to take the maximum of the membership value of A at (x) and $(-x)$ in accordance with equation (2.5-5). Due to the symmetry of the problem these values are actually the same and therefore B is still as derived in (E2.2-8). The membership function of B is,

$$\mu_B(y) = \sqrt{1 - y^2} \tag{E2.2-9}$$

as shown in Figure 2.7*c*. Figure 2.7*d* shows the geometric interpretation of fuzzyfying the interior of the ellipse in accordance with the fuzzy sets A and B above. The result is a kind of fuzzy elliptic region, strongest near the x axis and particularly at its $x = \pm 2$ sides and weakest near the origin and the $y = \pm 1$ sides. □

2.6 ALPHA-CUTS

With any fuzzy set A we can associate a collection of crisp sets known as *α-cuts* (*alpha-cuts*) or *level sets* of A. An α-cut is a crisp set consisting of elements of A which belong to the fuzzy set at least to a degree α. As we shall see in the next section, α-cuts offer a method for resolving any fuzzy set in terms of constituent crisp sets (something analogous to resolving a vector into its components). In Chapter 4 we will see that α-cuts are indispensable in performing arithmetic operations with fuzzy sets that represent various qualities of numerical data. It should be noted that α-cuts are crisp, *not* fuzzy, sets.[8]

[8]Formally, a distinction is made between two types of α-cuts, the *strong* and the *weak* α-cut (Dubois and Prade, 1980). We use the weak α-cut, simply calling it α-cut.

The α-cut of a fuzzy set A denoted as A_α is the crisp set comprised of all the elements x of a universe of discourse X for which the membership function of A is *greater than or equal to* α; that is,

$$A_\alpha = \{x \in X \mid \mu_A(x) \ge \alpha\} \tag{2.6-1}$$

where α is a parameter in the range $0 < \alpha \le 1$; the vertical bar "|" in equation (2.6-1) is shorthand for "such that."

Consider, for example, a fuzzy set A with trapezoidal membership function as shown in Figure 2.8. The 0.5-cut of A is simply the part of its support where its membership function is greater than 0.5. In Figure 2.8 we can see the 0.5-cut of A. Reflecting the fact that the α-cut is a *crisp set*, its membership function appears like a characteristic function. As another example consider the set A of *small integers* given by

$$A = 1.0/1 + 1.0/2 + 0.75/3 + 0.5/4 + 0.3/5 + 0.3/6 + 0.1/7 + 0.1/8$$

The 0.5-cut of A is simply the crisp set $A_{0.5} = \{1, 2, 3, 4\}$.

In the next section we will see that α-cuts provide a useful way both for resolving a membership function in terms of constituent crisp sets as well as for synthesizing a membership function out of crisp sets.

A fuzzy set can have an extensive support since its membership function can be zero or nearly zero, or very small. In order to deal with situations where small degrees of membership are not worthy of consideration, *level*

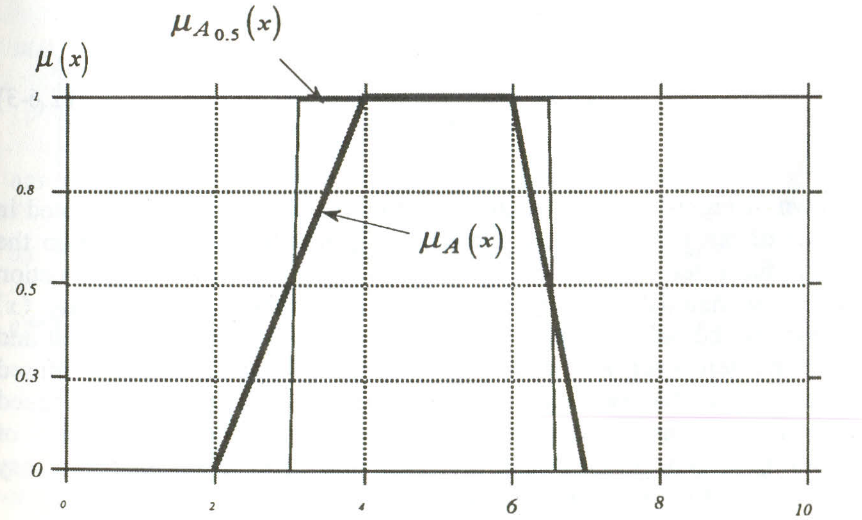


Figure 2.8 A fuzzy set A and its 0.5-cuts.

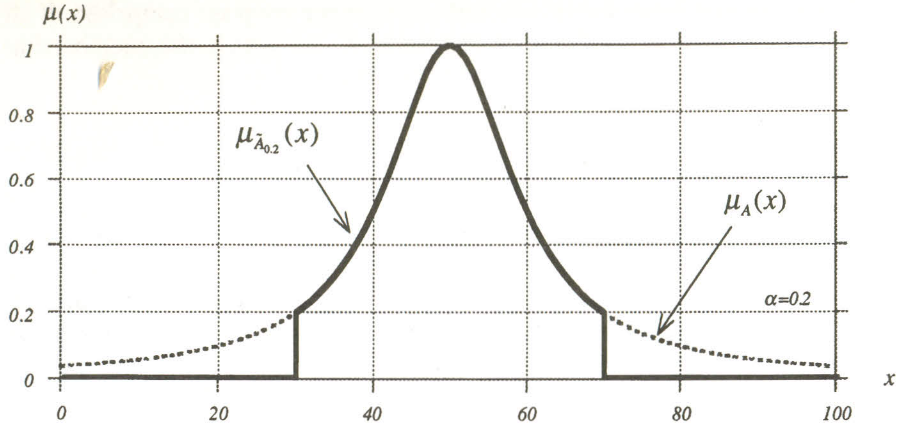


Figure 2.9 The 0.2-level fuzzy set of fuzzy set A.

fuzzy sets were introduced to exclude undesirable grades of membership (Radecki, 1977). We define the level fuzzy sets of a fuzzy set A as fuzzy sets $\tilde{A}_\alpha$ whose membership values are greater than α, where $0 < \alpha < 1$. Formally

$$\tilde{A}_\alpha \equiv \{(x, \mu_A(x)) | x \in A_\alpha\} \tag{2.6-2}$$

where A_α is the α-cut of A. Equation (2.6-2) indicates that for a given α we have a level fuzzy set which is the part of A that has membership greater than α. Let us consider, for example, a fuzzy set A whose membership function is

$$\mu_A(x) = \frac{1}{1 + 0.01(x - 50)^2} \tag{2.6-3}$$

as shown in Figure 2.9 (dotted curve). Suppose that we are not interested in the part of the support that has membership less than 0.2. We obtain the 0.2-level fuzzy set of A by chopping the part of the membership function which is less than 0.2 as shown in the figure. Its membership function $\mu_{\tilde{A}_{0.2}}(x)$ is shown by the solid curve. It is the same as $\mu_A(x)$ between $x = 30$ and $x = 70$ and zero everywhere else. Level fuzzy sets should not be confused with level sets, which is a synonym for α-cuts. Level fuzzy sets are indeed fuzzy sets, whereas α-cuts are crisp sets. They provide a useful way of considering fuzzy sets in the significant part of their support, and hence they save on computing time and storage requirements.

2.7 THE RESOLUTION PRINCIPLE

There are several ways of representing fuzzy sets, and we have already seen a few of them. They all involve two things: identifying a suitable universe of discourse and defining membership functions. One way to represent a fuzzy set would be to list all the elements of the universe of discourse together with the grade of membership of each element (omitting the possibly infinite elements that have zero membership). Alternatively, we can just provide an analytical representation of the membership function. The *resolution principle* offers another way of representing membership to a fuzzy set, namely through its α-cuts. It asserts that the membership function of a fuzzy set A can be expressed in terms of its α-cuts as follows:

$$\mu_A(x) = \bigvee_{0<\alpha\leq 1} \left[\alpha \cdot \mu_{A_\alpha}(x)\right] \tag{2.7-1}$$

where the maximum is taken over all α's. Equation (2.7-1) indicates that the membership function of A is the union (notice the max operator) of all α-cuts, after each one of them has been multiplied by α.

Consider, for example, the fuzzy set A with triangular membership function shown in Figure 2.10. Several α-cuts of A, each multiplied by α, are also shown. Knowing many α's and the α-cuts of A, we can form their products and put them together (in the sense of taking their union) to approximate the function. For example, we multiply the 0.25-cut by 0.25 to get the 0.25-cut pushed down to 0.25, and similarly we multiply the 0.5-cut by 0.5, the 0.75-cut by 0.75, and so on. When put together we have an approximation of the membership function of A as shown in Figure 2.10.

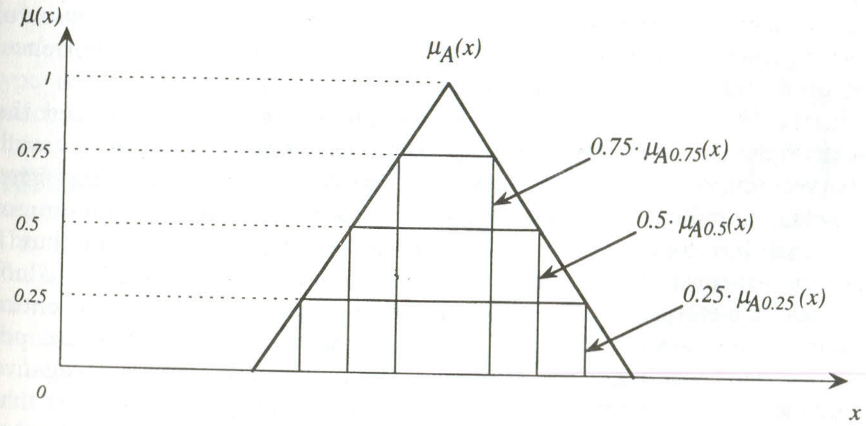


Figure 2.10 Putting many α-cuts of A multiplied by α together approximates the membership function of A.

Thus, a large enough family of α-cuts provides another way of representing a fuzzy set. Although we often know the membership function exactly, in some applications only α-cuts are known and out of them we need to approximate the membership function (see Chapter 4).

2.8 POSSIBILITY THEORY AND FUZZY PROBABILITIES

In the late 1970s Zadeh advanced a theoretical framework for information and knowledge analysis, called *possibility theory*, emphasizing the quantification of the semantic, context-dependent nature of symbols—that is, *meaning* rather than measures of information. The theory of possibility is analogous, and yet conceptually different from the theory of probability. Probability is fundamentally a measure of the frequency of occurrence of an event. Although there are several interpretations of probability (*subjectivistic*, *axiomatic*, and *frequentistic*), probabilities generally have a physical event basis. They are tied to statistical experiments and are primarily useful for quantifying how frequently a sample occurs in a population. Possibility theory, on the other hand, attempts to quantify how accurately a sample resembles an *ideal* element of a population. The ideal element is a prototypical class or a category of the population which we think of as a fuzzy set. In a sense, possibility theory may be viewed as a generalization of the theory of probability with the *consistency principle*, which we will see later on, providing a heuristic connection between the two. Possibility theory focuses more on the *imprecision* intrinsic in language, whereas probability theory focuses more on events that are *uncertain* in the sense of being random in nature. In natural language processing, automatic speech recognition, knowledge-based diagnosis, image analysis, robotics, analysis of rare events, information retrieval, and related areas, major problems are encountered on quantifying the meaning of events—that is, the efficacious and accurate interpretation of their significance and consequence and not the extent of their occurrence. Let us illustrate with a simple example.

In the field of reliability analysis, probabilistic methods have been the basic instrument for quantifying equipment and human reliability as well. Two very important concepts used are the *failure rate* and the *error rate*. Knowing the failure rate of a component amounts to knowing the duration of time that the component may be trusted to operate safely, and thus a schedule for replacement and maintenance activities can be devised. It is not unusual, however, that after a component is fixed or replaced, the entire system breaks down, a problem particularly acute with electronic components. Indeed, such general failures sometimes cause extremely negative consequences, leading to catastrophic accidents. The problem here is that failure rates are not sufficiently meaningful to account for the complex interactions that a human being, such as a maintenance technician or an operator, may have with a machine. In addition, the correct estimation of failure rate and error rate requires a large amount of data, which is often not practically possible to obtain. Is is obviously impractical to melt nuclear reactors to collect failure rate data. Thus, in practice, the failure rate and error rate are estimated by experts based on their engineering judgment (Onisawa, 1990); from this point of view, fuzzy possibilities and probabilities (which we will examine momentarily) can be used to model such judgments in a flexible and efficient way. Engineering judgment enters many areas of systems and reliability analysis including estimating the effect of environmental factors, operator stress, dependence between functions or units, selection of sequence of events, expressing the degree of uncertainty involved in the formulation of safety criteria, assuming parameter ranges, and so on (Shinohara, 1976). Alternatives to failure and error rates have been developed employing the notion of possibility measures, called *failure* and *error possibilities*, and have been applied to the reliability analysis of nuclear power plants, structural damage assessments, and earthquake engineering. Failure possibilities and error possibilities are essentially fuzzy sets on the interval [0, 1] that employ the notions we examine in this section.

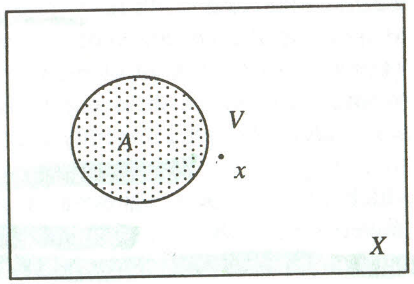
Over the years, two views, or schools of thought, of the definition of fuzziness have emerged. The first view, which we implicitly held in the previous sections, has to do with categorizing or grouping the elements of a universe of discourse into classes or sets whose boundaries of membership are fuzzy. Thus when we defined the set of *small numbers* in Example 2.1 we identified a category of numbers within the universe of all numbers. Implicitly, what we dealt with in the example was the problem of *imprecision*. Our main problem was to find the membership function that most appropriately or accurately described the category of *small numbers*. The other view of fuzziness has to do with the problem of *uncertainty*. Here our main concern is to quantify the certainty of an assertion such as "a number x is a *small number*," where x is an element of the universe of discourse X of numbers (whose location on X is not known in advance) and is therefore called a *nonlocated element*. Possibility theory was advanced in order to address this type of problem. Possibility is more generally known as a *fuzzy measure*, which is a function assigning a value between 0 and 1 to each crisp set of the universe of discourse, signifying the degree of evidence or belief that a particular element belongs to the set. Other types of fuzzy measure are *belief measures*, *plausibility measures*, *necessity measures*, and *probability measures*. The theory of fuzzy measures was advanced in 1974 by Sugeno as part of his Ph.D. dissertation at Tokyo University. Fuzzy measures subsume probability measures as well as belief and plausibility measures used in what is known as the *Dempster–Shafer Theory of Evidence*.

Let us now take a closer look at *possibility*. Possibility is a *fuzzy measure*, which means that possibility is a function with a value between 0 and 1, indicating the degree of evidence or belief that a certain element x belongs to a set (Zadeh, 1978; Dubois and Prade, 1988). A possibility of 0.3 for element x, for example, may indicate a 0.3 degree of evidence or belief that

x belongs to a certain set. How this belief is distributed to elements other that x is quantified through a *possibility distribution*. In possibility theory, the concept of *possibility distribution* is analogous to the notion of *probability distribution* in probability theory. A possibility distribution is viewed as a fuzzy restriction acting as an elastic constraint on the values that may be assigned to a variable. What does this mean? Well, it is best to review the notion of a *variable*, first. Let A be a crisp set defined on a universe X and let V be a variable taking values on some element x of X, a situation illustrated in Figure 2.11. The crisp set A is what in the parlance of probability we call an *event*. Events are comprised out of one or more basic events. Thus, the element x may be thought of as a basic event. If x is within A and x occurs, then we say that the event A has occurred as well. For example, in reliability analysis, equipment failure and human error are considered to be events whose occurrence is based on the occurrence of basic events known as *initiating events*. To say that *V takes its values in A* is to indicate that any element (basic event) of event A could *possibly* be a value of V and that any element outside of A, the complement of A, cannot be a value of V. Thus, the statement *V takes its value in A* can be viewed as inducing a possibility, Π over X, associating with each value x the possibility that x is a value of V. This can be written as

$$\Pi(V := x) = \pi_V(x) = \chi_A(x) = \begin{cases} 1 & \text{if} \quad x \in A \\ 0 & \text{if} \quad x \notin A \end{cases} \tag{2.8-1}$$

where " := " is an assignment symbol indicating that x is assigned to the



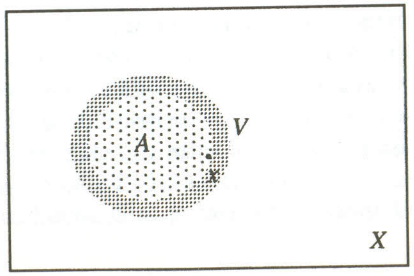


Figure 2.11 The statement about a variable V, "*V takes its values in A*," has a different meaning when the set A is crisp (*top*) than when the set A is fuzzy (*bottom*).

variable V, and $\pi_V(x)$ is the *possibility distribution* associated with V (or the *possibility distribution function* of Π). In equation (2.8-1), $\chi_A(x)$ is the characteristic function of A (see Section 2.1). Mathematically, Π is considered a *measure* which is a special function mapping the universe to the interval [0, 1]. Knowing that the values that V may take are members of A is the same as knowing which values of the universe X are restricted to be values of V and which are restricted not to be values of V. We indicated this in equation (2.8-1) by using the *characteristic function* of the crisp set A. We think of the crisp set A as a *restriction* on the values of the variable V, and in view of the nonfuzzy nature of A this type of restriction is called a *crisp restriction*.

Next, suppose that A is a fuzzy set and that its boundary no longer crisp (i.e., does not sharply divide members from nonmembers) but is instead a fuzzy boundary allowing an element x to be a member of A to some degree. As with any fuzzy set, A is uniquely identified by its membership function $\mu_A(x)$. In terms of events we think of A as a fuzzy event, and we can associate with each basic event x a membership function indicating its membership to A. Let us again consider a variable V whose arguments are elements of X[9]. Now suppose that V is constrained to take values on X. The fuzzy set A also restricts the possible values that the variable V may take, but in a fuzzy manner—that is, to a degree. In such a case we consider the fuzzy set A to act as a *fuzzy restriction* on the possible values of V. Generalizing equation (2.8-1) to the fuzzy case we say that the fuzzy set A induces a possibility Π. The associated possibility distribution $\pi_V(x)$ on the values that V may assume is defined to be equal to the membership function of A, $\mu_A(x)$ and is written as

$$\Pi(V := x) = \pi_V(x) = \mu_A(x) \tag{2.8-2}$$

Thus, the possibility that V is assigned x—that is, $V := x$, which is sometimes indicated as "V is x"—is postulated to be equal to the membership function of A evaluated at x—that is, $\mu_A(x)$. It is important to observe in equation (2.8-2) that *possibility distributions* are fuzzy sets, while *possibilities* are just numbers between 0 and 1. The possibility Π in (2.8.2) is a measure of the compatibility of a given crisp value x that V may take with an *a priori* defined set A. In this way, V becomes a variable associated with the *possibility distribution* $\pi_V(x)$ in much the same way as a random variable is associated with the probability distribution.

What equation (2.8-2) indicates is that in certain situations, such as in the definition of failure and error possibilities, it is of interest to interpret the membership function $\mu_A(x)$ of a fuzzy set as a *possibility distribution* of a variable V. In this sense the fuzzy set A is viewed as the set of more or less possible values for V.

[9] In Chapter 5 the variable V will be generalized to a *fuzzy variable*, which is a variable that takes fuzzy sets as values.

Given a possibility distribution $\pi_V(x)$, the possibility that x may belong to another crisp set B is defined as

$$\Pi(V \subset B) = \bigvee_{x \in B} \pi_V(x) \tag{2.8-3}$$

What equation (2.8-3) indicates is that the possibility of B is the possibility of the most possible elementary event x of B. Generalizing this relationship, it can be shown (Dubois and Prade, 1988; Kandel, 1986) that the possibility measure of the union of two crisp sets B and C is the maximum of the possibilities of B and C and can be written as

$$\Pi(B \cup C) = \Pi(B) \vee \Pi(C) \tag{2.8-4}$$

Given a fuzzy set A and a possibility distribution function, $\pi_V(x)$, the possibility of A, denoted as $\Pi(A)$, is given by

$$\Pi(A) = \bigvee_{x \in X} [\mu_A(x) \wedge \pi_V(x)] \tag{2.8-5}$$

Consider two fuzzy events A and B defined over the universe of discourse X. The possibility of A with respect to B is defined as

$$\Pi(A|B) = \bigvee_{x \in X} [\mu_A(x) \wedge \mu_B(x)] \tag{2.8-6}$$

The possibility measure of A with respect to B reflects the extent to which A and B coincide or overlap. Thus, possibility may be viewed as a measure of comparison of fuzzy sets.

Conditional possibilities have been defined in analogy with conditional probabilities; an entire body of theoretical results has been achieved, known generally as *possibility theory*. It is finding an increasing number of applications in the fields of knowledge representation and applied artificial intelligence (Ragheb and Tsoukalas, 1988). A very comprehensive treatment of possibility may be found in the book entitled *Possibility Theory* by Dubois and Prade (1988). The theory of possibility has assumed particular significance in the field of natural language processing due to the inherent fuzziness of natural language. In the late 1970s Zadeh constructed a universal language called *PRUF*, in which the translation of a proposition expressed in natural language takes the form of a procedure for computing the possibility distribution of a set of fuzzy relations in a database. The procedure, then, may be interpreted as a semantic computation transforming the meaning of a proposition to a computed possibility distribution quantifying the information conveyed by the proposition (Zadeh, 1983).

There are certain differences between probability and possibility measures worth pointing out. Possibility measures are "softer" than probability measures, and the interpretation of probability and possibility is quite different. Probability is used to quantify the frequency of occurrence of an event, while possibility (along with fuzzy tools) is used to quantify the *meaning* of an event. Consider the following example offered by Zadeh (1978). Suppose that we have the proposition "*Hans ate V eggs for breakfast*," where $V = \{1, 2, 3, \ldots\}$. A *possibility distribution* and a *probability distribution* may be associated with V, as shown in the following table:

x	1	2	3	4	5	6	7	8	9
$\pi_V(x)$	1	1	1	1	0.8	0.6	0.4	0.2	0.1
$p_V(x)$	0.1	0.8	0.1	0	0	0	0	0	0

The possibility distribution is interpreted as the *degree of ease* with which Hans can eat x eggs, while the probability distribution might have been determined by observing Hans at breakfast for 100 days. Note that the probability distribution function $p_V(x)$ is given a *frequentistic* interpretation and that it sums to '1', while the possibility distribution function $\pi_V(x)$ is imputed with a *situation* or *context-dependent* interpretation and does not have to sum to '1'.

Possibility is an upper bound for probability: A high degree of possibility does not imply a higher degree of probability. If, however, an event is not possible, it is not also probable. This is referred to as the *probability/possibility consistency principle* (Zadeh, 1978). This heuristic principle is useful for drawing a distinction between the *objectivistic* use of probability measures and the *subjectivist* use of possibility or fuzzy measures. When we attempt to use the two to describe a similar thing, we can use the *possibility/probability consistency principle* as a guide. Possibility measures are more flexible measures useful for epistemic (i.e., cognitive) or context-dependent descriptions. In general, according to Zadeh a variable may be associated with both a possibility distribution and a probability distribution, with the weak connection between the two given by the consistency principle (Zadeh, 1978).

In the language of probability theory the set A in Figure 2.11 may be viewed as a *fuzzy event*. Such a fuzzy event induces a distribution on the values of a variable which we called the possibility distribution function and defined in equation (2.8-2). We can also define the *probability of a fuzzy event A*. Suppose that a fuzzy event A is comprised of elementary events x, and with each x we associate a basic probability $p(x)$.

Zadeh defined the *probability of fuzzy event A* as the mathematical expectation (the first moment) of its membership function, that is,

$$P(A) = \frac{\int_X \mu_A(x) p(x)\, dx}{\int_X p(x)\, dx} \tag{2.8-7}$$

where A is a fuzzy event on the universe X, x is an element of X, also called an elementary event, and $p(x)$ is a probability distribution (Zadeh, 1968). When A is not a fuzzy event, equation (2.8-7) reduces back to the usual crisp probability $P(A)$. In equation (2.8-7) we assume that the probability measure on the entire universe of discourse must equal unity—that is, $\int_X p(x)\,dx = 1$.

In addition, given equation (2.8-7) we can define a *fuzzy mean* as

$$m_A = \frac{1}{P(A)} \int_X x \mu_A(x) p(x)\, dx \tag{2.8-8}$$

and a *fuzzy variance* as

$$\sigma_A^2 = \frac{1}{P(A)} \int_X (x - m_A)^2 \mu_A(x) p(x)\, dx \tag{2.8-9}$$

The probability of a fuzzy event as defined in equation (2.8-7) has been an extremely useful notion with wide application in the field of quantification theory (Terano et al., 1992). Quantification methods are useful in analyzing data involving human judgments which are not normally given numerical expression, as well as in interpreting and understanding such data.

Example 2.3 Possibility Measures and Distributions. Let us illustrate the distinction between *possibility measure* or *possibility* and *possibility distribution*. We consider a possibility distribution induced by the proposition "*V is a small integer*" where the possibility distribution is (subjectively) defined as

$$\pi_V(x) = 1.0/1 + 1.0/2 + 0.75/3 + 0.5/4 + 0.3/5 + 0.3/6 + 0.1/7 + 0.1/8 \tag{E2.3-1}$$

We also consider the crisp set $A = \{3, 4, 5\}$ which we can write as

$$A = \sum_{x \in X} \mu_A(x)/x = 1/3 + 1/4 + 1/5 \tag{E2.3-2}$$

What is the possibility of A? The *possibility measure* $\Pi(A)$ is found using equation (2.8-5); that is,

$$\Pi(A) = \bigvee_{x \in X} [\mu_A(x) \wedge \pi_V(x)] \tag{E2.3-3}$$

Using equations (E2.3-1) and (E2.3-2) in (E2.3-3), we can obtain the possibility of A:

$$\Pi(A) = 0.75 \vee 0.5 \vee 0.3 = 0.75 \tag{E2.3-4}$$

For another fuzzy set B = {*integers that are not small*} given by

$$B = 0.2/3 + 0.3/4 + 0.6/5 + 0.8/6 + 1.0/7$$

using equation (E2.3-3), we could obtain that the possibility of B is

$$\Pi(B) = 0.2 \vee 0.3 \vee 0.3 \vee 0.3 \vee 0.1 = 0.3 \tag{E2.3-5}$$

It should be noted in equations (E2.3-4) and (E2.3-5) that the possibility is simply a number between 0 and 1, whereas the possibility distribution is a fuzzy set—for example, equation (E2.3-1).

Let us now consider a simple instance of how to generate the possibility distribution itself. Let $C = 1/1 + 1/2 + 0.8/3 + 0.6/4 + 0.4/5 + 0.2/6$ be a fuzzy set that represents *small numbers*. Then the proposition "*V is a small number*" associates with V the possibility distribution, $\pi_V(x)$, taken in view of equation (2.8-2) to be equal to the membership function of C—that is,

$$\pi_V(x) = 1/1 + 1/2 + 0.8/3 + 0.6/4 + 0.4/5 + 0.2/6 \tag{E2.3-6}$$

In equation (E2.3-6) a singleton such as $0.6/4$ indicates that the possibility that x is 4, given that x is a *small integer*, is 0.6. □

small number.

REFERENCES

Dubois, D., and Prade, H., *Fuzzy Sets and Systems: Theory and Applications*, Academic Press, Boston, 1980.

Dubois, D., and Prade, H., *Possibility Theory*, Plenum Press, New York, 1988.

Kandel, A., *Fuzzy Mathematical Techniques with Applications*, Addison-Wesley, Reading, MA, 1986.

Kaufmann, A., *Introduction to the Theory of Fuzzy Subsets*, Volume I, Academic Press, New York, 1975.

Klir, G. J., and Folger, T. A., *Fuzzy Sets, Uncertainty, and Information*, Prentice-Hall, Englewood Cliffs, NJ, 1988.

Kosko, B., *Neural Networks and Fuzzy Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1992.

Onisawa, T., An Application of Fuzzy Concepts to Modelling of Reliability Analysis, *Fuzzy Sets and Systems*, No. 37, pp. 267–286, North-Holland, Amsterdam, 1990.

Pal, S. K., and Majumder, D. K. D., *Fuzzy Mathematical Approach to Pattern Recognition*, John Wiley & Sons, New York, 1986.

Radecki, T., Level Fuzzy Sets, *Journal of Cybernetics*, Vol. 7, pp. 189–198, 1977.

Ragheb, M., and Tsoukalas, L. H., Monitoring Performance of Devices Using a Coupled Probability–Possibility Method, *International Journal of Expert Systems*, Vol. 1, pp. 111–130, 1988.

Shinohara, Y., Fuzzy Set Concepts for Risk Assessment, *International Institute for Applied Systems Analysis*, Report WP-76-2, Laxenburg, Austria, January 1976.

Terano, T., Asai, K., and Sugeno, M., *Fuzzy Systems Theory and Its Applications*, Academic Press, Boston, 1992.

Zadeh, L. A., Fuzzy Sets, *Information and Control*, Vol. 8, pp. 338–353, 1965.

Zadeh, L. A., Probability Measure of Fuzzy Events, *Journal of Mathematical Analysis and Applications*, Vol. 23, pp. 421–427, 1968.

Zadeh, L. A., Outline of a New Approach to the Analysis of Complex Systems and Decision Processes, *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3, pp. 28–44, 1973.

Zadeh, L. A., The Concept of a Linguistic Variable and its Application to Approximate Reasoning, *Information Sciences*, Vol. 8, pp. 199–249, 1975.

Zadeh, L. A., Fuzzy Sets as a Basis for Theory of Possibility, *Fuzzy Sets and Systems*, Vol. 1, pp. 3–28, 1978.

Zadeh, L. A., A Computational Approach to Fuzzy Quantifiers in Natural Languages, *Computer and Mathematics*, Vol. 9, pp. 149–184, 1983.

Zadeh, L. A., Fuzzy Logic, *IEEE Computer*, pp. 83–93, April 1988.

Zimmermann, H. J., *Fuzzy Set Theory and its Applications*, Kluwer-Nijhoff, Boston, 1985.

PROBLEMS

1. What happens to the curves in Figure 2.5 if we set $F_2 = 40$ and vary F_1 as in the figure?

2. In Figure 2.6, what is the significance of the intersection between the $\mu_A = 0.5$ line and the curves?

3. In Example 2.2, substitute $y = \sin x$ for equation (E2.2-1) and utilize the extension principle in the same way as in the example. Choose an appropriate range for x and assume any additional information needed as in the example.

4. The fuzzy variable of Figure 2.9 is given by the equation $\mu_A(x) = 1/[1 + 0.3(x - 50)^2]$. Show that the 0.2 level fuzzy set of fuzzy set A can be represented by α-cuts using the resolution principle.

5. The fuzzy sets A and B are given by

$$A = 0.33/6 + 0.67/7 + 1.00/8 + 0.67/9 + 0.33/10$$
$$B = 0.20/3 + 0.60/4 + 1.00/5 + 0.60/6 + 0.20/7$$

   (a) Write an expression for $A \vee B$.
   (b) Write an expression for $A \wedge B$.

6. Different fuzzy symbols are often used to mean similar things.
   (a) Write all symbols or terms that have the same general meaning as *max* ($\vee$).
   (b) Write all symbols or terms that have the same general meaning as *min* ($\wedge$).

7. Given fuzzy set A, describing *pressure p is higher than* 15 *mPa*, through the membership function:

$$\mu_A(x) = \frac{1}{1 + (x - 15)^{-2}} \quad x > 15,$$
$$= 0 \quad x \leq 15,$$

and fuzzy set B, describing *pressure p is approximately equal to* 17 *mPa*, with membership function:

$$\mu_B(x) = \frac{1}{1 + (x - 17)^4}.$$

Find the membership function of the fuzzy set C, describing *pressure p is higher than* 15 *mPa and approximately equal to* 17 *mPa*. Use at least four different norms for interpreting *AND* (see Appendix) and draw all membership functions.

8. Using the data given in Problem 7, find the membership function of the fuzzy set D, describing *pressure p is higher than* 15 *mPa or approximately equal to* 17 *mPa*. Use at least four different norms for interpreting *OR* (see Appendix) and draw all membership functions.

9. Using the data given in Problem 7, find the membership function of the fuzzy set E, describing *pressure p is not higher than* 15 *mPa and approximately equal to* 17 *mPa*. Use four different norms for interpreting *AND* (see Appendix) and draw all membership functions.

10. Determine all α-cuts for the following fuzzy sets, given that $\alpha = 0.0, 0.1, 0.2, \ldots 0.9, 1.0$.
   I. $A = 0.1/3 + 0.2/4 + 0.3/5 + 0.4/6 + 0.5/7 + 0.6/8 + 0.7/9 + 0.8/10 + 1.0/11 + 0.8/12$
   II. $B = \int_{-\infty < x < +\infty} \left[\frac{1}{1 + (x - 15)^{-2}}\right] / x$

   Write a MATLAB program that takes a number of α-cuts (minimum 10) and reconstructs the membership function.

11. Let $X = N \times N$, and the fuzzy sets:

$$\mu_A(x) = \frac{1}{1 + 10(x - 2)^2}$$

$$\mu_B(y) = \frac{1}{1 + 2y^2}$$

Let the mappings $z = f(x, y)$, $f\colon N \times N \to N$ be the following quadric surfaces

(a) $$z = \sqrt{\frac{x^2}{4} + \frac{y^2}{2}}, \quad x \in A,\ y \in B.$$

(b) $$\frac{x^2}{9} + \frac{y^2}{15} - \frac{z^2}{8} = 1$$

(c) $$2y^2 + 12z^2 = x^2$$

Sketch the surfaces and determine the image $f(A \times B)$ by the extension principle, for each of the above.