THE CONCEPT OF A LINGUISTIC VARIABLE AND ITS APPLICATION TO APPROXIMATE REASONING

by

L. A. Zadeh

Memorandum No. ERL-M411

15 October 1973

THE CONCEPT OF A LINGUISTIC VARIABLE AND ITS APPLICATION TO APPROXIMATE REASONING

by

L. A. Zadeh

Memorandum No. ERL-M411

15 October 1973

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

THE CONCEPT OF A LINGUISTIC VARIABLE AND ITS APPLICATION TO APPROXIMATE REASONING

L. A. Zadeh

Abstract

By a *linguistic variable* we mean a variable whose values are words or sentences in a natural or artificial language. For example, *Age* is a linguistic variable if its values are linguistic rather than numerical, i.e., *young, not young, very young, quite young, old, not very old and not very young*, etc., rather than 20, 21, 22, 23,... .

In more specific terms, a linguistic variable is characterized by a quintuple $(\mathcal{X}, T(\mathcal{X}), U, G, M)$ in which $\mathcal{X}$ is the name of the variable; $T(\mathcal{X})$ is the *term-set* of $\mathcal{X}$, that is, the collection of its linguistic values; U is a universe of discourse; G is a *syntactic rule* which generates the terms in $T(\mathcal{X})$; and M is a *semantic rule* which associates with each linguistic value X its *meaning*, M(X), where M(X) denotes a fuzzy subset of U.

The meaning of a linguistic value X is characterized by a *compatibility function*, $c : U \rightarrow [0,1]$, which associates with each u in U its compatibility with X. Thus, the compatibility of age 27 with *young* might be 0.7 while that of 35 might be 0.2. The function of the semantic rule is to relate the compatibilities of the so-called *primary* terms in a composite linguistic value - e.g., *young* and *old* in *not very young and not very old* - to the compatibility of the composite value. To this end, the hedges such as *very*, *quite*, *extremely*, etc., as well as the connectives *and* and *or* are treated as nonlinear operators which modify the meaning

of their operands in a specified fashion.

The concept of a linguistic variable provides a means of approximate characterization of phenomena which are too complex or too ill-defined to be amenable to description in conventional quantitative terms. In particular, treating Truth as a linguistic variable with values such as true, very true, completely true, not very true, untrue, etc., leads to what is called fuzzy logic. By providing a basis for approximate reasoning, that is, a mode of reasoning which is not exact nor very inexact, such logic may offer a more realistic framework for human reasoning than the traditional two-valued logic.

It is shown that probabilities, too, can be treated as linguistic variables with values such as likely, very likely, unlikely, etc. Computation with linguistic probabilities requires the solution of nonlinear programs and leads to results which are imprecise to the same degree as the underlying probabilities.

The main applications of the linguistic approach lie in the realm of humanistic systems - especially in the fields of artificial intelligence, linguistics, human decision processes, pattern recognition, psychology, law, medical diagnosis, information retrieval, economics and related areas.

THE CONCEPT OF A LINGUISTIC VARIABLE AND ITS APPLICATION TO APPROXIMATE REASONING

L. A. Zadeh

Computer Sciences Division,
Department of Electrical Engineering and Computer Sciences
and the Electronics Research Laboratory
University of California, Berkeley, California 94720

1. Introduction

One of the fundamental tenets of modern science is that a phenomenon cannot be claimed to be well understood until it can be characterized in quantitative terms.[1] Viewed in this perspective, much of what constitutes the core of scientific knowledge may be regarded as a reservoir of concepts and techniques which can be drawn upon to construct mathematical models of various types of systems and thereby yield quantitative information concerning their behavior.

Given our veneration for what is precise, rigorous and quantitative, and our disdain for what is fuzzy, unrigorous and qualitative, it is not surprising that the advent of digital computers has resulted in a rapid expansion in the use of quantitative methods throughout most fields of human knowledge. Unquestionably, computers have proved to be highly

---

This work was supported in part by the Navy Electronic Systems Command under Contract N00039-71-C-0255, the Army Research Office, Durham, N.C., under Grant DA-ARO-D-31-124-71-G174, and the National Science Foundation under Grant GK-10656X3. The writing of the paper was completed while the author was participating in a Joint Study Program with the Systems Research Department, IBM Research Laboratory, San Jose, California.

[1] As expressed by Lord Kelvin in 1883 [1], "In physical science a first essential step in the direction of learning any subject is to find principles of numerical reckoning and practicable methods for measuring some quality connected with it. I often say that when you can measure what you are speaking about and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind: it may be the beginning of knowledge but you have scarcely, in your thoughts, advanced to the state of science, whatever the matter may be."

effective in dealing with <u>mechanistic</u> systems, that is, with inanimate systems whose behavior is governed by the laws of mechanics, physics, chemistry and electromagnetism. Unfortunately, the same cannot be said about <u>humanistic</u> systems,[2] which - so far at least - have proved to be rather impervious to mathematical analysis and computer simulation. Indeed, it is widely agreed that the use of computers has not shed much light on the basic issues arising in philosophy, psychology, literature, law, politics, sociology and other human-oriented fields. Nor have computers added significantly to our understanding of human thought processes - excepting, perhaps, some examples to the contrary that can be drawn from artificial intelligence and related fields [2], [3], [4], [5], [51].

It may be argued, as we have done in [6] and [7], that the ineffectiveness of computers in dealing with humanistic systems is a manifestation of what might be called the <u>principle of incompatibility</u> - a principle which asserts that high precision is incompatible with high complexity.[3] Thus, it may well be the case that the conventional techniques of system analysis and computer simulation - based as they are on precise manipulation of numerical data - are intrinsically incapable of coming to grips with the great complexity of human thought processes and decision-making. The acceptance of this premise suggests that, in order to be able to make significant assertions about the behavior of

[2] By a <u>humanistic</u> system we mean a system whose behavior is strongly influenced by human judgement, perception or emotions. Examples of humanistic systems are: economic systems, political systems, legal systems, religious systems, etc. A single individual and his thought processes may also be viewed as a humanistic system.

[3] Stated somewhat more concretely, the complexity of a system and the precision with which it can be analyzed bear a roughly inverse relation to one another.

humanistic systems, it may be necessary to abandon the high standards of rigor and precision that we have become conditioned to expect of our mathematical analyses of well-structured mechanistic systems, and become more tolerant of approaches which are approximate in nature. Indeed, it is entirely possible that only through the use of such approaches could computer simulation become truly effective as a tool for the analysis of systems which are too complex or too ill-defined for the application of conventional quantitative techniques.

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. For example, in speaking of age, when we say "John is young," we are less precise than when we say, "John is 25." In this sense, the label _young_ may be regarded as a _linguistic value_ of the variable _Age_, with the understanding that it plays the same role as the numerical value 25 but is less precise and hence less informative. The same is true of the linguistic values _very young_, _not young_, _extremely young_, _not very young_, etc. as contrasted with the numerical values 20, 21, 22, 23, ...

If the values of a numerical variable are visualized as points in a plane, then the values of a linguistic variable may be likened to ball-parks with fuzzy boundaries. In fact, it is by virtue of the employment of ball-parks rather than points that linguistic variables acquire the ability to serve as a means of approximate characterization of phenomena which are too complex or too ill-defined to be susceptible of description in precise terms. What is also important, however, is that by the use of a so-called extension principle, much of the existing mathematical apparatus of systems analysis can be adapted to the manipulation of linguistic variables. In this way, we may be able to develop an approximate calculus of linguistic variables which could be of use in a wide variety of practical applications.

The totality of values of a linguistic variable constitute its term-set, which in principle could have an infinite number of elements. For example, the term-set of the linguistic variable Age might read

Age = young + not young + very young + not very young + very very young + ... + old + not old + very old + not very old + ... + not very young and not very old + ... + middle-aged + not middle-aged + ... + not old and not middle-aged + ... + extremely old + ... (1)

in which + is used to denote the union rather than the arithmetic sum. Similarly, the term-set of the linguistic variable Appearance might be

Appearance = beautiful + pretty + cute + handsome + attractive + not beautiful + very pretty + very very handsome + more or less pretty + quite pretty + quite handsome + fairly handsome + not very attractive and not very unattractive + ...

In the case of the linguistic variable Age, the numerical variable age whose values are the numbers 0, 1, 2, 3, ..., 100 constitutes what may be called the base variable for Age. In terms of this variable, a linguistic value such as young may be interpreted as a label for a fuzzy restriction on the values of the base variable. This fuzzy restriction is what we take to be the meaning of young.

A fuzzy restriction on the values of the base variable is characterized by a compatibility function which associates with each value of the base variable a number in the interval [0,1] which represents its compatibility with the fuzzy restriction. For example, the compatabilities of the numerical ages 22, 28 and 35 with the fuzzy restriction labeled young might be 1, 0.7 and 0.2, respectively. The meaning of young, then, would be represented by a graph of the form shown in Fig. 1.1, which is a plot of the compatibility function of young with respect to the base variable age.

The conventional interpretation of the statement "John is young," is that John is a member of the class of young men. However, considering that the class of young men is a fuzzy set, that is, there is no sharp transition from being young to not being young, the assertion that John is a member of the class of young men is inconsistent with the precise mathematical definition of "is a member of." The concept of a linguistic variable allows us to get around this difficulty in the following manner.

The name "John" is viewed as a name of a composite linguistic variable whose components are linguistic variables named Age, Height, Weight, Appearance, etc. Then, the statement "John is young," is interpreted as an assignment equation (Fig. 1.2)

Age = young

which assigns the value young to the linguistic variable Age. In turn, the value young is interpreted as a label for a fuzzy restriction on the base variable age, with the meaning of this fuzzy restriction defined by its compatibility function. As an aid in the understanding of the concept of a linguistic variable, Fig. 1.3 shows the hierarchical structure of the relation between the linguistic variable Age, the fuzzy restrictions which represent the meaning of its values, and the values of the base variable age.

There are several basic aspects of the concept of a linguistic variable that are in need of elaboration.

First, it is important to understand that the notion of compatibility is distinct from that of probability. Thus, the statement that the compatibility of, say, 28 with young is 0.7, has no relation to the probability of the age-value 28. The correct interpretation of the compatibility-value 0.7 is that it is merely a subjective indication of the extent to which the age-value 28 fits one's conception of the label young. As we shall see in later sections, the rules of manipulation applying to compatibilities are different from those applying to probabilities, although there are certain parallels between the two.

Second, we shall usually assume that a linguistic variable is structured in the sense that it is associated with two rules: (i) a syntactic rule, which specifies the manner in which the linguistic values which are in the term-set of the variable may be generated. In regard to this rule, our usual assumption will be that the terms in the term-set of the variable are generated by a context-free grammar.

The second rule, (ii), is a semantic rule which specifies a procedure for computing the meaning of any given linguistic value. In this connection, we observe that a typical value of a linguistic variable, e.g., not very young and not very old, involves what might be called the primary terms, e.g., young and old, whose meaning is both subjective and context-dependent. We assume that the meaning of such terms is specified a priori.

In addition to the primary terms, a linguistic value may involve connectives such as and, or, either, neither, etc.; the negation not; and the hedges such as very, more or less, completely, quite, fairly, extremely, somewhat, etc. As we shall see in later sections, the connectives, the hedges and the negation may be treated as operators which modify the meaning of their operands in a specified, context-independent, fashion. Thus, if the meaning of young is defined by the compatibility function whose form is shown in Fig. 1.1, then the meaning of very young could be obtained by squaring the compatibility function of young, while that of not young would be given by subtracting the compatibility function of young from unity (Fig. 1.4). These two rules are special instances of a more general semantic rule which is described in Sec. 5.

Third, when we speak of a linguistic variable such as Age, the underlying base variable, age, is numerical in nature. Thus, in this case we can define the meaning of a linguistic value such as young by a compatibility function which associates with each age in the interval [0,100] a number in the interval [0,1] which represents the compatibility of that age with the label young.

On the other hand, in the case of the linguistic variable Appearance, we do not have a well-defined base variable; that is, we do not know how to express the degree of beauty, say, as a function of some physical measurements. We could still associate with each member of a group of ladies, for example, a grade of membership in the class of beautiful women, say 0.9 with Fay, 0.7 with Adele, 0.8 with Kathy and 0.9 with Vera, but these values of the compatibility function would be based on impressions which we may not be able to articulate or formalize in explicit terms. In other words, we are defining the compatibility function not on a set of mathematically well-defined objects, but on a set of labeled impressions. Such definitions are meaningful to a human but not - at least directly - to a computer.[4]

As we shall see in later sections, in the first case, where the base variable is numerical in nature, linguistic variables can be treated in a reasonably precise fashion by the use of the extension principle for fuzzy sets. In the second case, their treatment becomes much more qualitative. In both cases, however, some computation is involved - to a lesser or greater degree. Thus, it should be understood that the linguistic approach is not entirely qualitative in nature. Rather, the computations are performed behind the scene and, at the end, linguistic approximation is employed to convert numbers into words (Fig. 1.5).

A particularly important area of application for the concept of a linguistic variable is that of approximate reasoning, by which we mean a type of reasoning which is neither very precise nor very imprecise. As an illustration, the following inference would be an instance of approximate reasoning:

---

[4]The basic problem which is involved here is that of abstraction from a set of samples of elements of a fuzzy set. A discussion of this problem may be found in [8].

x is small

x and y are approximately equal

therefore

y is more or less small.

The concept of a linguistic variable enters into approximate reasoning as a result of treating Truth as a linguistic variable whose truth-values form a term-set such as shown below

Truth = true + not true + very true + completely true + more or less true + fairly true + essentially true + ... + false + very false + neither true nor false + ...

The corresponding base variable, then, is assumed to be a number in the interval [0,1], and the meaning of a primary term such as true is identified with a fuzzy restriction on the values of the base variable. As usual, such a restriction is characterized by a compatibility function which associates a number in the interval [0,1] with each numerical truth-value. For example, the compatibility of the numerical truth-value 0.7 with the linguistic truth-value very true might be 0.6. Thus, in the case of truth-values, the compatibility function is a mapping from the unit interval to itself. (Fig. 6.1.)

Treating truth as a linguistic variable leads to a fuzzy logic which may well be a better approximation to the logic involved in human decision processes than the classical two-valued logic.[5] Thus, in fuzzy

[5] Expositions of alternative approaches to vagueness may be found in [9] - [18].

logic it is meaningful to assert what would be inadmissibly vague in classical logic, e.g.,

The truth-value of "Berkeley is close to San Francisco," is <u>quite true</u>.

The truth-value of "Palo Alto is close to San Francisco," is <u>fairly true</u>.

Therefore, the truth-value of "Palo Alto is more or less close to Berkeley," is <u>more or less true</u>.

Another important area of application for the concept of a linguistic variable lies in the realm of probability theory. If probability is treated as a linguistic variable, its term-set would typically be:

<u>Probability</u> = <u>likely</u> + <u>very likely</u> + <u>unlikely</u> + <u>extremely likely</u> + <u>fairly likely</u> + ... + <u>probable</u> + <u>improbable</u> + <u>more or less probable</u> + ...

By legitimizing the use of linguistic probability-values, we make it possible to respond to a question such as, "What is the probability that it will be a warm day a week from today," with an answer such as <u>fairly high</u>, instead of, say, 0.8. The linguistic answer would, in general, be much more realistic, considering, first, that <u>warm day</u> is a fuzzy event, and, second, that our understanding of weather dynamics is not sufficient to allow us to make unequivocal assertions about the underlying probabilities.

In the following sections, the concept of a linguistic variable and its applications will be discussed in greater detail. To place the concept of a linguistic variable in a proper perspective, we shall begin our discussion with a formalization of the notion of a conventional (nonfuzzy) variable. For our purposes, it will be helpful to visualize such a variable as a tagged valise with rigid (hard) sides. (Fig. 2.1.) Putting an object into the valise corresponds to assigning a value to the variable, and the restriction on what can be put in corresponds to a subset of the universe of discourse which comprises those points which can be assigned as values to the variable. In terms of this analogy, a fuzzy variable, which is defined in Sec. 4, may be likened to a tagged valise with soft rather than rigid sides. (Fig. 4.1.) In this case, the restriction on what can be put in is fuzzy in nature, and is defined by a compatibility function which associates with each object a number in the interval [0,1] representing the degree of ease with which that object can be fitted in the valise. For example, given a valise named X, the compatibility of a coat with X would be 1, while that of a record-player might be 0.7.

As will be seen in Sec. 4, an important concept in the case of fuzzy variables is that of noninteraction, which is analogous to the concept of independence in the case of random variables. This concept arises when we deal with two or more fuzzy variables, each of which may be likened to a compartment in a soft valise. Such fuzzy variables are interactive if the assignment of a value to one affects the fuzzy restrictions placed on the others. This effect may be likened to the interference between objects which are put into different compartments of a soft valise. (Fig. 4.3.)

A linguistic variable is defined in Sec. 5 as a variable whose values are fuzzy variables. In terms of our valise analogy, a linguistic variable corresponds to a hard valise into which we can put soft valises, with each soft valise carrying a name tag which describes a fuzzy restriction on what can be put into that valise. (Fig. 5.2.)

The application of the concept of a linguistic variable to the notion of Truth is discussed in Sec. 6. Here we describe a technique for computing the conjunction, disjunction and negation for linguistic truth-values and lay the groundwork for fuzzy logic.

In Sec. 7, the concept of a linguistic variable is applied to probabilities, and it is shown that linguistic probabilities can be used for computational purposes. However, because of the constraint that the numerical probabilities must add up to unity, the computations in question involve the solution of nonlinear programs and hence are not as simple to perform as computations involving numerical probabilities.

The last section is devoted to a discussion of the so-called <u>compositional rule of inference</u> and its application to approximate reasoning. This rule of inference is interpreted as the process of solving a simultaneous system of so-called <u>relational assignment equations</u> in which linguistic values are assigned to fuzzy restrictions. Thus, if a statement such as "x is small" is interpreted as an assignment of the linguistic value <u>small</u> to the fuzzy restriction on x, and the statement "x and y are approximately equal," is interpreted as the assignment of a fuzzy relation labeled <u>approximately equal</u> to the fuzzy restriction on the ordered pair (x, y), then the conclusion "y is more or less small," may be viewed as a linguistic approximation to the solution of the simultaneous equations

$R(x)$ = small

$R(x,y)$ = approximately equal

in which $R(x)$ and $R(x,y)$ denote the restrictions on x and (x,y), respectively. (Fig. 8.3.)

The compositional rule of inference leads to a generalized modus ponens, which may be viewed as an extension of the familiar rule of inference: If A is true and A implies B, then B is true. The section closes with an example of a fuzzy theorem in elementary geometry and a brief discussion of the use of fuzzy flowcharts for the representation of definitional fuzzy algorithms.

The material in Secs. 2, 3 and 4 is intended to provide a mathematical basis for the concept of a linguistic variable, which is introduced in Sec. 5. For those readers who may not be interested in the mathematical aspects of the theory, it may be expedient to proceed directly to Sec. 5 and refer where necessary to the definitions and results described in the preceding sections.

2. The Concept of a Variable

In the preceding section, our discussion of the concept of a linguistic variable was informal in nature. To set the stage for a more formal definition, we shall focus our attention in this section on the concept of a conventional (nonfuzzy) variable. Then, in Sec. 3 we shall extend the concept of a variable to fuzzy variables and subsequently will define a linguistic variable as a variable whose values are fuzzy variables.

Although the concept of a (nonfuzzy) variable is very elementary in nature, it is by no means a trivial one. For our purposes, the following formalization of the concept of a variable provides a convenient basis for later extensions.

<u>Definition 2.1</u> A variable is characterized by a triple (X,U,R(X;u)), in which X is the name of the variable; U is a universe of discourse (finite or infinite set); u is a generic[1] name for the elements of U; and R(X;u) is a subset of U which represents a <u>restriction</u>[2] on the values of u imposed by X. For convenience, we shall usually abbreviate R(X;u) to R(X) or R(u) or R(x), where x denotes a generic name for the values of X, and will refer to R(X) simply as the restriction <u>on</u> u or the restriction <u>imposed</u> <u>by</u> X.

---

[1]A generic name is a single name for all elements of a set. For simplicity, we shall frequently use the same symbol for both a set and the generic name for its elements, relying on the context for disambiguation.

[2]In conventional terminology, R(X) is the range of X. Our use of the term <u>restriction</u> is motivated by the role played by R(X) in the case of fuzzy variables.

In addition, a variable is associated with an assignment equation

$$x = u: \quad R(X) \tag{2.1}$$

or equivalently

$$x = u, \quad u \in R(X) \tag{2.2}$$

which represents the assignment of a value u to x subject to the restriction R(X). Thus, the assignment equation is satisfied iff (if and only if) $u \in R(X)$.

Example 2.2 As a simple illustration consider a variable named age. In this case, U might be taken to be the set of integers 0, 1, 2, 3, ... and R(X) might be the subset 0, 1, 2, ..., 100.

More generally, let $X_1,\ldots,X_n$ be n variables with respective universes of discourse $U_1,\ldots,U_n$. The ordered n-tuple $X = (X_1,\ldots,X_n)$ will be referred to as an n-ary composite (or joint) variable. The universe of discourse for X is the cartesian product

$$U = U_1 \times U_2 \times \ldots \times U_n \tag{2.3}$$

and the restriction $R(X_1,\ldots,X_n)$ is an n-ary relation in $U_1 \times \ldots \times U_n$. This relation may be defined by its characteristic (membership) function $\mu_R\colon U_1 \times \ldots \times U_n \to \{0,1\}$, where

$$\begin{aligned} \mu_R(u_1,\ldots,u_n) &= 1 \quad \text{if} \quad (u_1,\ldots,u_n) \in R(X_1,\ldots,X_n) \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{2.4}$$

and u_i is a generic name for the elements of U_i, $i = 1,\ldots,n$. Correspondingly, the n-ary assignment equation assumes the form

$$(x_1,\ldots,x_n) = (u_1,\ldots,u_n): \quad R(X_1,\ldots,X_n) \tag{2.5}$$

which is understood to mean that

$$x_i = u_i, \quad i = 1,\ldots,n \tag{2.6}$$

subject to the restriction $(u_1,\ldots,u_n) \in R(X_1,\ldots,X_n)$, with x_i, $i = 1, \ldots, n$, denoting a generic name for values of X_i.

Example 2.3 Suppose that $X_1 \triangleq$ age of father[3], $X_2 \triangleq$ age of son, and $U_1 \triangleq U_2 = \{1,2,\ldots,100\}$. Furthermore, suppose that $x_1 \geq x_2 + 20$ (x_1 and x_2 are generic names for values of X_1 and X_2). Then, $R(X_1,X_2)$ may be defined by

$$\begin{aligned} \mu_R(u_1,u_2) &= 1 \quad \text{for} \quad 21 \leq u_1 \leq 100, \quad u_1 \geq u_2 + 20 \\ &= 0 \quad \text{elsewhere} \end{aligned} \tag{2.7}$$

Marginal and Conditioned Restrictions

As in the case of probability distributions, the restriction $R(X_1,\ldots,X_n)$ imposed by $(X_1,\ldots,X_n)$ induces marginal restrictions $R(X_{i_1},\ldots,X_{i_k})$ imposed by composite variables of the form $(X_{i_1},\ldots, X_{i_k})$, where the index sequence $q = (i_1,\ldots,i_k)$ is a subsequence of the index sequence $(1,2,\ldots,n)$.[4] In effect, $R(X_{i_1},\ldots,X_{i_k})$ is the smallest (i.e., most restrictive) restriction imposed by $(X_{i_1},\ldots,X_{i_k})$ which satisfies the implication

$$(u_1,\ldots,u_n) \in R(X_1,\ldots,X_n) \Rightarrow (u_{i_1},\ldots,u_{i_k}) \in R(X_{i_1},\ldots,X_{i_k}) \tag{2.8}$$

[3]The symbol $\triangleq$ stands for "denotes" or is "equal by definition."

[4]In the case of a binary relation $R(X_1,X_2)$, $R(X_1)$ and $R(X_2)$ are usually referred to as the domain and range of $R(X_1,X_2)$.

Thus, a given k-tuple $u_{(q)} \triangleq (u_{i_1},\ldots,u_{i_k})$ is an element of $R(X_{i_1},\ldots,X_{i_k})$ iff there exists an n-tuple $u \triangleq (u_1,\ldots,u_n) \in R(X_1,\ldots,X_n)$ whose i_1th, $\ldots,i_k$th components are equal to $u_{i_1},\ldots,u_{i_k}$, respectively. Expressed in terms of the characteristic functions of $R(X_1,\ldots,X_n)$ and $R(X_{i_1},\ldots,X_{i_k})$, this statement translates into the equation

$$\mu_{R(X_{i_1},\ldots,X_{i_k})}(u_{i_1},\ldots,u_{i_k}) = \vee_{u(q')}\mu_{R(X_1,\ldots,X_n)}(u_1,\ldots,u_n) \quad (2.9)$$

or more compactly

$$\mu_{R(X_{(q)})}(u_{(q)}) = \vee_{u_{(q')}}\mu_{R(X)}(u) \quad (2.10)$$

where q' is the complement of the index sequence $q = (i_1,\ldots,i_k)$ relative to $(1,\ldots,n)$, $u_{(q')}$ is the complement of the k-tuple $u_{(q)} \triangleq (u_{i_1},\ldots,u_{i_k})$ relative to the n-tuple $u \triangleq (u_1,\ldots,u_n)$, $X_{(q)} \triangleq (X_{i_1},\ldots,X_{i_k})$ and $\vee_{u_{(q')}}$ denotes the supremum of its operand over the u's which are in $u_{(q')}$. (Throughout this paper, the symbols $\vee$ and $\wedge$ stand for Max and Min, respectively; thus, for any real a, b

$$\begin{aligned} a \vee b = \mathrm{Max}(a,b) &= a \quad \text{if} \quad a \geq b \\ &= b \quad \text{if} \quad a < b \end{aligned} \quad (2.11)$$

and

$$\begin{aligned} a \wedge b = \mathrm{Min}(a,b) &= a \quad \text{if} \quad a \leq b \\ &= b \quad \text{if} \quad a > b \end{aligned}$$

Consistent with this notation, the symbol $\vee_z$ should be read as "supremum over the values of z.") Since μ_R can take only two values - 0 or 1 - (2.10) means that $\mu_{R(X_{(q)})}(u_{(q)})$ is 1 iff there exists a $u_{(q')}$ such that $\mu_{R(X)}(u) = 1$

Comment 2.4 There is a simple analogy which is very helpful in clarifying the notion of a variable and related concepts. Specifically, a nonfuzzy variable in the sense formalized in Definition 2.1 may be likened to a tagged valise having rigid (hard) sides, with X representing the name on the tag, U corresponding to a list of objects which could be put in a valise, and R(X) representing a sublist of U which comprises those objects which can be put into valise X. (For example, an object like a boat would not be in U, while an object like a typewriter might be in U but not in R(X), and an object like a cigarette box or a pair of shoes would be in R(X).) In this interpretation, the assignment equation

$$x = u: \quad R(X)$$

signifies that an object u which satisfies the restriction R(X) (i.e., is on the list of objects which can be put into X) is put into X. (Fig. 2.1.)

An n-ary composite variable $X \triangleq (X_1,\ldots,X_n)$ corresponds to a valise carrying the name-tag X which has n compartments named $X_1,\ldots,X_n$ with adjustable partitions between them. The restriction $R(X_1,\ldots,X_n)$ corresponds to a list of n-tuples of objects $(u_1,\ldots,u_n)$ such that u_1 can be put in compartment X_1, u_2 in compartment $X_2,\ldots,$ and u_n in compartment X_n simultaneously. (See Fig. 2.2.) In this connection, it should be noted that n-tuples on this list could be associated with different arrangements of partitions. If n = 2, for example, then for a particular placement of the partition we could put a coat in compartment X_1 and a suit in compartment X_2, while for some other placement we could put the coat in compartment X_2 and a box of shoes in compartment X_1. In this event, both (coat,suit) and (shoes, coat) would be included in the list of pairs of objects which can be put in X simultaneously.

In terms of the valise analogy, the n-ary assignment equation

$$(x_1,\ldots,x_n) = (u_1,\ldots,u_n): \; R(X_1,\ldots,X_n)$$

represents the action of putting u_1 in $X_1,\ldots,$ and u_n in X_n simultaneously, under the restriction that the n-tuple of objects $(u_1,\ldots,u_n)$ must be on the $R(X_1,\ldots,X_n)$ list. Furthermore, a marginal restriction such as $R(X_{i_1},\ldots,X_{i_k})$ may be interpreted as a list of k-tuples of objects which can be put in compartments $X_{i_1},\ldots,X_{i_k}$ simultaneously, in conjunction with every allowable placement of objects in the remaining compartments.

<u>Comment 2.5</u> It should be noted that (2.9) is analogous to the expression for a marginal distribution of a probability distribution, with $\vee$ corresponding to summation (or integration). However, this analogy should not be construed to imply that $R(X_{i_1},\ldots,X_{i_k})$ is in fact a marginal probability distribution.

It is convenient to view the right-hand member of (2.9) as the characteristic function of the projection[5] of $R(X_1,\ldots,X_n)$ on $U_{i_1} \times \cdots \times U_{i_k}$. Thus, in symbols

$$R(X_{i_1},\ldots,X_{i_k}) = \text{Proj } R(X_1,\ldots,X_n) \text{ on } U_{i_1} \times \ldots \times U_{i_k} \tag{2.12}$$

or more simply

[5]The term projection as used in the literature is somewhat ambivalent in that it could denote either the operation of projecting or the result of such operation. To avoid this ambivalence in the case of fuzzy relations, we will occasionally employ the term <u>shadow</u> [19] to denote the relation resulting from applying an operation of projection to another relation.

$$R(X_{i_1},\ldots,X_{i_k}) = P_q R(X_1,\ldots,X_n)$$

where P_q denotes the operation of projection on $U_{i_1} \times \ldots \times U_{i_k}$ with $q = (i_1,\ldots,i_k)$.

Example 2.6 In the case of Example 2.3, we have

$$R(X_1) = P_1\ R(X_1,X_2) = \{21,\ldots,100\}$$

$$R(X_2) = P_2\ R(X_1,X_2) = \{1,\ldots,80\}$$

Example 2.7 Fig. 2.3 shows the restrictions on u_1 and u_2 induced by $R(X_1,X_2)$.

An alternative way of describing projections is the following. Viewing $R(X_1,\ldots,X_n)$ as a relation in $U_1 \times \ldots \times U_n$, let $q' = (j_1,\ldots,j_m)$ denote the index sequence complementary to $q = (i_1,\ldots,i_k)$, and let $R(X_{i_1},\ldots,X_{i_k}|u_{j_1},\ldots,u_{j_m})$ or, more compactly, $R(X_{(q)}|u_{(q')})$ denote a restriction in $U_{i_1} \times \ldots \times U_{i_k}$ which is <u>conditioned on</u> $u_{j_1},\ldots,u_{j_m}$. The characteristic function of this conditioned restriction is defined by

$$\mu_{R(X_{i_1},\ldots,X_{i_k}|u_{j_1},\ldots,u_{j_m})}(u_{i_1},\ldots,u_{i_k}) = \mu_{R(X_1,\ldots,X_n)}(u_1,\ldots,u_n) \tag{2.13}$$

or more simply (see (2.10)),

$$\mu_{R(X_{(q)}|u_{(q')})}(u_{(q)}) = \mu_{R(X)}(u)$$

with the understanding that the arguments $u_{j_1},\ldots,u_{j_m}$ in the right-hand member of (2.14) are treated as parameters. In consequence of this understanding, although the characteristic function of the conditioned restriction is numerically equal to that of $R(X_1,\ldots,X_n)$, it defines a fuzzy relation in $U_{i_1} \times \dots \times U_{i_k}$ rather than in $U_1 \times \dots \times U_n$.

In view of (2.9), (2.12) and (2.13), the projection of $R(X_1,\dots,X_n)$ on $U_{i_1} \times \dots \times U_{i_k}$ may be expressed as

$$P_q R(X_1,\dots,X_n) = \cup_{u(q')} R(X_{i_1},\dots,X_{i_k} | u_{j_1},\dots,u_{j_m}) \qquad (2.14)$$

where $\cup_{u(q')}$ denotes the union of the family of restrictions $R(X_{i_1},\dots,X_{i_k} | u_{j_1},\dots,u_{j_m})$ parametrized by $u_{(q')} \triangleq (u_{j_1},\dots,u_{j_m})$. Consequently, (2.14) implies that the marginal restriction $R(X_{i_1},\dots,X_{i_k})$ in $U_{i_1} \times \dots \times U_{i_k}$ may be expressed as the union of conditioned restrictions $R(X_{i_1},\dots,X_{i_k} | u_{j_1},\dots,u_{j_m})$, i.e.,

$$R(X_{k_1},\dots,X_{i_k}) = \cup_{u(q')} R(X_{i_1},\dots,X_{i_k} | u_{j_1},\dots,u_{j_m}) \qquad (2.15)$$

or more compactly

$$R(X_{(q)}) = \cup_{u(q')} R(X_{(q)} | u_{(q')})$$

<u>Example 2.8</u> As a simple illustration of (2.15), assume that $U_1 = U_2 \triangleq \{3,5,7,9\}$ and that $R(X_1,X_2)$ is characterized by the following relation matrix. (In this matrix, the (i,j)th entry is 1 iff the ordered pair (ith element of U_1, jth element of U_2) belongs to $R(X_1,X_2)$. In effect, the relation matrix of a relation R constitutes a tabulation of the characteristic function of R.)

R	3	5	7	9
3	0	0	1	0
5	1	0	1	0
7	1	0	1	1
9	1	0	0	1

In this case,

$$R(X_1,X_2|u_1 = 3) = \{7\}$$

$$R(X_1,X_2|u_1 = 5) = \{3,7\}$$

$$R(X_1,X_2|u_1 = 7) = \{3,7,9\}$$

$$R(X_1,X_2|u_1 = 9) = \{3,9\}$$

and hence

$$\begin{aligned} R(X_2) &= \{7\} \cup \{3,7\} \cup \{3,7,9\} \cup \{3,9\} \\ &= \{3,7,9\} \end{aligned}$$

Interaction and noninteraction

A basic concept that we shall need in later sections is that of the interaction between two or more variables - a concept which is analogous to the dependence of random variables. More specifically, let the variable $X = (X_1,\ldots,X_n)$ be associated with the restriction $R(X_1,\ldots,X_n)$, which induces the restrictions $R(X_1),\ldots,R(X_n)$ on $u_1,\ldots,u_n$, respectively. Then we have

Definition 2.9 $X_1,\ldots,X_n$ are noninteractive variables under $R(X_1,\ldots,X_n)$ iff $R(X_1,\ldots,X_n)$ is separable, i.e.,

$$R(X_1,\ldots,X_n) = R(X_1) \times \ldots \times R(X_n) \tag{2.16}$$

where, for $i = 1, \ldots, n$,

$$\begin{aligned} R(X_i) &= \text{Proj } R(X_1,\ldots,X_n) \text{ on } U_i \\ &= \cup_{u_{(q')}} R(X_i|u_{(q')}) \end{aligned} \tag{2.17}$$

with $u_{(q)} \triangleq u_i$ and $u_{(q')} \triangleq$ complement of u_i in $(u_1,\ldots,u_n)$.

Example 2.10 Fig. 2.4a shows two noninteractive variables X_1 and X_2 whose restrictions $R(X_1)$ and $R(X_2)$ are intervals; in this case, $R(X_1,X_2)$ is the cartesian product of the intervals in question. In Fig. 2.4b, $R(X_1,X_2)$ is a proper subset of $R(X_1) \times R(X_2)$ and hence X_1 and X_2 are interactive. Note that in Example 2.3 X_1 and X_2 are interactive.

As will be shown in a more general context in Sec. 4, if $X_1,\ldots,X_n$ are noninteractive then an n-ary assignment equation

$$(x_1,\ldots,x_n) = (u_1,\ldots,u_n): \; R(X_1,\ldots,X_n) \tag{2.18}$$

can be decomposed into a sequence of n unary assignment equations

$$\begin{aligned} x_1 &= u_1: \; R(X_1) \\ x_2 &= u_2: \; R(X_2) \\ &\ldots\ldots\ldots\ldots\ldots \\ x_n &= u_n: \; R(X_n) \end{aligned} \tag{2.19}$$

where $R(X_i)$, $i = 1,\ldots,n$, is the projection of $R(X_1,\ldots,X_n)$ on U_i, and by Definition 2.9

$$R(X_1,\ldots,X_n) = R(X_1) \times \ldots \times R(X_n). \tag{2.20}$$

In the case where $X_1,\ldots,X_n$ are interactive, the sequence of n unary assignment equations assumes the following form (see also (4.34)).

$$\begin{aligned} x_1 &= u_1: \; R(X_1) \\ x_2 &= u_2: \; R(X_2 | u_1) \\ &\ldots\ldots\ldots\ldots\ldots \\ x_n &= u_n: \; R(X_n | u_1,\ldots,u_{n-1}) \end{aligned} \tag{2.21}$$

where $R(X_i | u_1,\ldots,u_{i-1})$ denotes the induced restriction for X_i conditioned on $u_1,\ldots,u_{i-1}$. The characteristic function of this conditioned restriction is expressed by (see (2.13))

$$\mu_{R(X_i | u_1,\ldots,u_{i-1})}(u_i) = \mu_{R(X_1,\ldots,X_i)}(u_1,\ldots,u_i) \tag{2.22}$$

with the understanding that the arguments $u_1,\ldots,u_{i-1}$ in the right-hand member of (2.22) play the role of parameters.

<u>Comment 2.10</u>. In words, (2.21) means that, in the case of interactive variables, once we have assigned a value u_1 to x_1, the restriction on u_2 becomes dependent on u_1. Then, the restriction on u_3 becomes dependent on the values assigned to x_1 and x_2, and, finally, the restriction on u_n becomes dependent on $u_1,\ldots,u_{n-1}$. Furthermore, (2.22) implies that the restriction on u_i given $u_1,\ldots,u_{i-1}$ is essentially the same as the marginal restriction on $(u_1,\ldots,u_i)$, with $u_1,\ldots,u_{i-1}$ treated as parameters. This is illustrated in Fig. 2.5.

In terms of the valise analogy (see Comment 2.4), $X_1,\ldots,X_n$ are noninteractive if the partitions between the compartments named $X_1,\ldots,X_n$ are not adjustable. In this case, what is placed in a compartment X_i has no influence on the objects that can be placed in the other compartments.

In case the partitions are adjustable, this is no longer true, and $X_1,\ldots,X_n$ become interactive in the sense that the placement of an object, say u_i, in X_i affects what can be placed in the complementary compartments. From this point of view, the sequence of unary assignment equations (2.21) describes the way in which the restriction on compartment X_i is influenced by the placement of objects $u_1,\ldots,u_{i-1}$ in $X_1,\ldots,X_{i-1}$.

Our main purpose in defining the notions of noninteraction, marginal restriction, conditioned restriction, etc. for nonfuzzy variables is (a) to indicate that concepts analogous to statistical independence, marginal distribution, conditional distribution, etc., apply also to nonrandom, nonfuzzy variables; and (b) to set the stage for similar concepts in the case of fuzzy variables. As a preliminary, we shall turn our attention to some of the relevant properties of fuzzy sets and formulate an extension principle which will play an important role in later sections.

3. Fuzzy Sets and the Extension Principle

As will be seen in Sec. 4, a fuzzy variable X differs from a nonfuzzy variable in that it is associated with a restriction R(X) which is a fuzzy subset of the universe of discourse. Consequently, as a preliminary to our consideration of the concept of a fuzzy variable, we shall review some of the pertinent properties of fuzzy sets and state an extension principle which allows the domain of a transformation or a relation in U to be extended from points in U to fuzzy subsets of U.

Fuzzy Sets - Notation and Terminology

A fuzzy subset A of a universe of discourse U is characterized by a membership function μ_A: $U \to [0,1]$ which associates with each element u of U a number $\mu_A(u)$ in the interval [0,1], with $\mu_A(u)$ representing the grade of membership of u in A.[2] The support of A is the set of points in U at which $\mu_A(u)$ is positive. The height of A is the supremum of $\mu_A(u)$ over U. A crossover point of A is a point in U whose grade of membership in A is 0.5.

Example 3.1 Let the universe of discourse be the interval [0,1], with u interpreted as age. A fuzzy subset of U labeled old may be defined by a membership function such as

$$\mu_A(u) = 0 \quad \text{for} \quad 0 \leq u \leq 50 \qquad (3.1)$$

[1]More detailed discussions of fuzzy sets and their properties may be found in the listed references. (A detailed exposition of the fundamentals together with many illustrative examples may be found in the recent text by A. Kaufmann [20]).

[2]More generally, the range of μ_A may be a partially ordered set (see [21], [22]) or a collection of fuzzy sets. The latter case will be discussed in greater detail in Sec. 6.

$$\mu_A(u) = \left(1 + \left(\frac{u-50}{5}\right)^{-2}\right)^{-1} \quad \text{for } 50 \leq u \leq 100$$

In this case, the support of old is the interval [50,100]; the height of old is effectively unity; and the crossover point of old is 55.

To simplify the representation of fuzzy sets we shall employ the following notation.

A nonfuzzy finite set such as

$$U = \{u_1,\ldots,u_n\} \tag{3.2}$$

will be expressed as

$$U = u_1 + u_2 + \ldots + u_n \tag{3.3}$$

or

$$U = \sum_{i=1}^{n} u_i \tag{3.4}$$

with the understanding that + denotes the union rather than the arithmetic sum. Thus, (3.3) may be viewed as a representation of U as the union of its constituent singletons.

As an extension of (3.3), a fuzzy subset, A, of U will be expressed as

$$A = \mu_1 u_1 + \ldots + \mu_n u_n \tag{3.5}$$

or

$$A = \sum_{i=1}^{n} \mu_i u_i \tag{3.6}$$

where μ_i, $i = 1,\ldots,n$, is the grade of membership of u_i in A. In cases where the u_i are numbers, there might be some ambiguity regarding the identity of the μ_i and u_i components of the string $\mu_i u_i$. In such cases, we shall employ a separator symbol such as / for disambiguation, writing

$$A = \mu_1/u_1 + \ldots + \mu_n/u_n \tag{3.7}$$

or

$$A = \sum_{i=1}^{n} \mu_i/u_i \tag{3.8}$$

<u>Example 3.2</u> Let U = {a,b,c,d} or, equivalently,

$$U = a + b + c + d \tag{3.9}$$

In this case, a fuzzy subset A of U may be represented unambiguously as

$$A = 0.3a + b + 0.9c + 0.5d \tag{3.10}$$

On the other hand, if

$$U = 1 + 2 + \ldots + 100 \tag{3.11}$$

then we shall write

$$A = 0.3/25 + 0.9/3 \tag{3.12}$$

in order to avoid ambiguity.

<u>Example 3.3</u> In the universe of discourse comprising the integers 1,2,...,10, i.e.,

$$U = 1 + 2 + \ldots + 10 \tag{3.13}$$

the fuzzy subset labeled <u>several</u> may be defined as

$$\underline{\text{several}} = 0.5/3 + 0.8/4 + 1/5 + 1/6 + 0.8/7 + 0.5/8 \tag{3.14}$$

Example 3.4 In the case of the countable universe of discourse

$$U = 0 + 1 + 2 + \dots \tag{3.15}$$

the fuzzy set labeled small may be expressed as

$$\underline{\text{small}} = \sum_{0}^{\infty} \left(1 + \left(\frac{u}{10}\right)^2\right)^{-1} / u \tag{3.16}$$

Like (3.3), (3.5) may be interpreted as a representation of a fuzzy set as the union of its constituent fuzzy singletons $\mu_i u_i$ (or μ_i / u_i). From the definition of the union (see 3.34)), it follows that if in the representation of A we have $u_i = u_j$, then we can make the substitution expressed by

$$\mu_i u_i + \mu_j u_i = (\mu_i \vee \mu_j) u_i \tag{3.17}$$

For example,

$$A = 0.3a + 0.8a + 0.5b \tag{3.18}$$

may be rewritten as

$$\begin{aligned} A &= (0.3 \vee 0.8)a + 0.5b \\ &= 0.8a + 0.5b \end{aligned} \tag{3.19}$$

When the support of a fuzzy set is a continuum rather than a countable or a finite set, we shall write

$$A = \int_U \mu_A(u)/u \tag{3.20}$$

with the understanding that $\mu_A(u)$ is the grade of membership of u in A, and the integral denotes the union of the fuzzy singletons $\mu_A(u)/u$, $u \in U$.

Example 3.5 In the universe of discourse consisting of the interval [0,100], with u = age, the fuzzy subset labeled old (whose membership function is given by (3.1)), may be expressed as

$$\underline{\text{old}} = \int_{50}^{100} \left(1 + \left(\frac{u-50}{5}\right)^{-2}\right)^{-1} / u \qquad (3.21)$$

Note that the crossover point for this set, that is, the point u at which

$$\mu_{\underline{\text{old}}}(u) = 0.5 \qquad (3.22)$$

is u = 55.

A fuzzy set A is said to be normal if its height is unity, that is, if

$$\operatorname{Sup}_u \mu_A(u) = 1 \qquad (3.23)$$

Otherwise A is subnormal. In this sense, the set old defined by (3.21) is normal, as is the set several defined by (3.17). On the other hand, the subset of U = 1 + 2 + ... + 10 labeled not small and not large and defined by

$$\underline{\text{not}}\ \underline{\text{small}}\ \underline{\text{and}}\ \underline{\text{not}}\ \underline{\text{large}} = 0.2/2 + 0.3/3 + 0.4/4 + 0.5/5 + 0.4/6 + 0.3/7 + 0.2/8 \qquad (3.24)$$

is subnormal. It should be noted that a subnormal fuzzy set may be normalized by dividing μ_A by $\operatorname{Sup}_u \mu_A(u)$.

A fuzzy subset of U may be a subset of another fuzzy or nonfuzzy subset of U. More specifically, A is a subset of B or is contained in B iff $\mu_A(u) \leq \mu_B(u)$ for all u in U. In symbols

$$A \subset B \Leftrightarrow \mu_A(u) \leq \mu_B(u), \quad u \in U. \tag{3.25}$$

Example 3.6 If $U = a + b + c + d$ and

$$\begin{aligned} A &= 0.5\,a + 0.8b + 0.3d \\ B &= 0.7a + b + 0.3c + d \end{aligned} \tag{3.26}$$

then $A \subset B$.

Level-Sets of a Fuzzy Set

If A is a fuzzy subset of U, then an α-level set of A is a nonfuzzy set denoted by A_α which comprises all elements of U whose grade of membership in A is greater than or equal to α. In symbols

$$A_\alpha = \{u | \mu_A(u) \geq \alpha\} \tag{3.27}$$

A fuzzy set A may be decomposed into its level-sets through the resolution identity[3]

$$A = \int_0^1 \alpha A_\alpha \tag{3.28}$$

or

$$A = \sum_\alpha \alpha A_\alpha \tag{3.29}$$

where αA_α is the product of a scalar α with the set A_α (in the sense of (3.39) and $\int_0^1$ (or $\sum_\alpha$) is the union of the A_α, with α ranging from 0 to 1.

The resolution identity may be viewed as the result of combining

[3] The resolution identity and some of its applications are discussed in greater detail in [23] and [24].

together those terms in (3.5) which fall into the same level-set. More specifically, suppose that A is represented in the form

$$A = 0.1/2 + 0.3/1 + 0.5/7 + 0.9/6 + 1/9 \tag{3.30}$$

Then by using (3.17), A can be rewritten as

$$\begin{aligned} A = 0.1/2 + 0.1/1 + 0.1/7 + 0.1/6 + 0.1/9 \\ + 0.3/1 + 0.3/7 + 0.3/6 + 0.3/9 \\ + 0.5/7 + 0.5/6 + 0.5/9 \\ + 0.9/6 + 0.9/9 \\ + 1/9 \end{aligned}$$

or

$$\begin{aligned} A = {} & 0.1\ (1/2 + 1/1 + 1/7 + 1/6 + 1/9) \\ & + 0.3\ (1/1 + 1/7 + 1/6 + 1/9) \\ & + 0.5\ (1/7 + 1/6 + 1/9) \\ & + 0.9\ (1/6 + 1/9) \\ & + 1/9 \end{aligned} \tag{3.31}$$

which is in the form (3.29), with the level-sets given by (see (3.27))

$$\begin{aligned} A_{0.1} &= 2 + 1 + 7 + 6 + 9 \\ A_{0.3} &= 1 + 7 + 6 + 9 \\ A_{0.5} &= 7 + 6 + 9 \\ A_{0.9} &= 6 + 9 \\ A_{1} &= 9 \end{aligned} \tag{3.32}$$

As will be seen in later sections, the resolution identity - in combination with the extension principle - provides a convenient way of generalizing various concepts associated with nonfuzzy sets to fuzzy sets. This, in fact, is the underlying basis for many of the definitions stated in the sequel.

Operations on Fuzzy Sets

Among the basic operations which can be performed on fuzzy sets are the following.

1. The complement of A is denoted by $\neg$ A (or sometimes by A') and is defined by

$$\neg A = \int_U (1 - \mu_A(u))/u \tag{3.33}$$

The operation of complementation corresponds to negation. Thus, if A is a label for a fuzzy set, then not A would be interpreted as $\neg$ A. (See Example (3.8).)

2. The union of fuzzy sets A and B is denoted by A + B (or, more conventionally, by A $\cup$ B) and is defined by

$$A + B = \int_U (\mu_A(u) \vee \mu_B(u))/u \tag{3.34}$$

The union corresponds to the connective or. Thus, if A and B are labels of fuzzy sets, then A or B would be interpreted as A + B.

3. The intersection of A and B is denoted by A $\cap$ B and is defined by

$$A \cap B = \int_U (\mu_A(u) \wedge \mu_B(u))/u \tag{3.35}$$

The intersection corresponds to the connective and; thus

$$A \text{ and } B = A \cap B \tag{3.36}$$

Comment 3.7 It should be understood that $\vee$ ($\triangleq$ Max) and $\wedge$ ($\triangleq$ Min) are not the only operations in terms of which the union and intersection can be defined. (See [25] and [26] for discussions of this point.) In this connection, it is important to note that when *and* is identified with Min, as in (3.36), it represents a "hard" *and* in the sense that it allows no trade-offs between its operands. By contrast, an *and* which is identified with the arithmetic product, as in (3.37), would act as a "soft" *and*. Which of these two and possibly other definitions is more appropriate depends on the context in which *and* is used.

4. The *product* of A and B is denoted by AB and is defined by

$$AB = \int_U \mu_A(u)\mu_B(u)/u \tag{3.37}$$

Thus, A^α, where α is any positive number, should be interpreted as

$$A^\alpha = \int_U (\mu_A(u))^\alpha/u \tag{3.38}$$

Similarly, if α is any nonnegative real number such that $\alpha \operatorname{Sup}_u \mu_A(u) \leq 1$, then

$$\alpha A = \int_U \alpha\mu_A(u)/u \tag{3.39}$$

As a special case of (3.38), the operation of *concentration* is defined as

$$\text{CON}(A) = A^2 \tag{3.40}$$

while that of *dilation* is expressed by

$$\text{DIL}(A) = A^{0.5} \tag{3.41}$$

As will be seen in Sec. 6, the operations of concentration and dilation are useful in the representation of linguistic hedges.

Example 3.8 If

$$U = 1 + 2 + \dots + 10$$

$$A = 0.8/3 + 1/5 + 0.6/6 \qquad (3.42)$$

$$B = 0.7/3 + 1/4 + 0.5/6$$

then

$$\neg A = 1/1 + 1/2 + 0.2/3 + 1/4 + 0.4/6 + 1/7 + 1/8 + 1/9 + 1/10$$

$$A + B = 0.8/3 + 1/4 + 1/5 + 0.6/6 \qquad (3.43)$$

$$A \cap B = 0.7/3 + 0.5/6$$

$$AB = 0.56/3 + 0.3/6$$

$$A^2 = 0.64/3 + 1/5 + 0.36/6$$

$$0.4A = 0.32/3 + 0.4/5 + 0.24/6$$

$$CON(B) = 0.49/3 + 1/4 + 0.25/6$$

$$DIL(B) = 0.84/3 + 1/4 + 0.7/6$$

5. If $A_1,\dots,A_n$ are fuzzy subsets of U, and $w_1,\dots,w_n$ are non-negative weights adding up to unity, then a convex combination of A_1, $\dots,A_n$ is a fuzzy set A whose membership function is expressed by

$$\mu_A = w_1\mu_{A_1} + \dots + w_n\mu_{A_n} \qquad (3.44)$$

where + denotes the arithmetic sum. The concept of a convex combination is useful in the representation of linguistic hedges such as essentially, typically, etc. which modify the weights associated with the components of a fuzzy set [27].

6. If $A_1,\dots,A_n$ are fuzzy subsets of $U_1,\dots,U_n$, respectively, the

cartesian product of $A_1, \ldots, A_n$ is denoted by $A_1 \times \ldots \times A_n$ and is defined as a fuzzy subset of $U_1 \times \ldots \times U_n$ whose membership function is expressed by

$$\mu_{A_1 \times \ldots \times A_n}(u_1, \ldots, u_n) = \mu_{A_1}(u_1) \wedge \ldots \wedge \mu_{A_n}(u_n) \quad (3.45)$$

Thus, we can write (see (3.52))

$$A_1 \times \ldots \times A_n = \int_{U_1 \times \ldots \times U_n} (\mu_{A_1}(u_1) \wedge \ldots \wedge \mu_{A_n}(u_n)) / (u_1, \ldots, u_n) \quad (3.46)$$

Example 3.9 If $U_1 = U_2 = 3 + 5 + 7$, $A_1 = 0.5/3 + 1/5 + 0.6/7$ and $A_2 = 1/3 + 0.6/5$, then

$$\begin{aligned} A_1 \times A_2 &= 0.5/(3,3) + 1/(5,3) + 0.6/(7,3) \\ &\quad + 0.5/(3,5) + 0.6/(5,5) + 0.6/(7,5) \end{aligned} \quad (3.47)$$

7. The operation of fuzzification has, in general, the effect of transforming a nonfuzzy set into a fuzzy set or increasing the fuzziness of a fuzzy set. Thus, a fuzzifier F applied to a fuzzy subset A of U yields a fuzzy subset F(A;K) which is expressed by

$$F(A;K) = \int_U \mu_A(u) K(u) \quad (3.48)$$

where the fuzzy set K(u) is the kernel of F, that is, the result of applying F to a singleton 1/u:

$$K(u) = F(1/u;K); \quad (3.49)$$

$\mu_A(u)K(u)$ represents the product (in the sense of (3.39)) of a scalar $\mu_A(u)$ and the fuzzy set $K(u)$; and $\int$ is the union of the family of fuzzy sets $\mu_A(u)K(u)$, $u \in U$. In effect, (3.48) is analogous to the integral representation of a linear operator, with K(u) being the counterpart of the impulse response.

Example 3.10 Assume that U, A and K(u) are defined by

$$U = 1 + 2 + 3 + 4 \tag{3.50}$$

$$A = 0.8/1 + 0.6/2$$

$$K(1) = 1/1 + 0.4/2$$

$$K(2) = 1/2 + 0.4/1 + 0.4/3$$

Then

$$F(A;K) = 0.8(1/1 + 0.4/2) + 0.6(1/2 + 0.4/1 + 0.4/3)$$

$$= 0.8/1 + 0.6/2 + 0.24/3$$

The operation of fuzzification plays an important role in the definition of linguistic hedges such as more or less, slightly, somewhat, much, etc. For example, if A $\triangleq$ positive is the label for the nonfuzzy class of positive numbers, then slightly positive is a label for a fuzzy subset of the real line whose membership function is of the form shown in Fig. 3.1. In this case, slightly is a fuzzifier which transforms positive into slightly positive. However, it is not always possible to express the effect of a fuzzifier in the form (3.48), and slightly is a case in point. A more detailed discussion of this and related issues may be found in [27].

Fuzzy Relations

If U is the cartesian product of n universes of discourse $U_1,\ldots,U_n$, then an n-ary fuzzy relation, R, in U is a fuzzy subset of U. As in (3.20), R may be expressed as the union of its constituent fuzzy singletons $\mu_R(u_1,\ldots,u_n)/(u_1,\ldots,u_n)$, i.e.

$$R = \int_{U_1 \times \ldots \times U_n} \mu_R(u_1,\ldots,u_n)/(u_1,\ldots,u_n) \qquad (3.52)$$

where μ_R is the membership function of R.

Common examples of (binary) fuzzy relations are: <u>much greater than</u>, <u>resembles</u>, <u>is relevant to</u>, <u>is close to</u>, etc. For example, if $U_1 = U_2 = (-\infty,\infty)$, the relation <u>is close to</u> may be defined by

$$\underline{\text{is close to}} \triangleq \int_{U_1 \times U_2} e^{-a|u_1-u_2|}/(u_1,u_2) \qquad (3.53)$$

where a is a scale factor. Similarly, if $U_1 = U_2 = 1 + 2 + 3 + 4$ then the relation <u>much greater than</u> may be defined by the relation matrix

R	1	2	3	4
1	0	0.3	0.8	1
2	0	0	0	0.8
3	0	0	0	0.3
4	0	0	0	0

(3.54)

in which the (i,j)th element is the value of $\mu_R(u_1,u_2)$ for the ith value of u_1 and jth value of u_2.

If R is a relation from U to V (or, equivalently, a relation in U × V) and S is a relation from V to W, then the composition of R and S is a fuzzy relation from U to W denoted by $R \circ S$ and defined by[4]

$$R \circ S = \int_{U \times W} \vee_v(\mu_R(u,v) \wedge \mu_S(u,w))/(u,w) \qquad (3.55)$$

[4]Equation (3.55) defines the max-min composition of R and V. Max-product composition is defined similarly, except that $\wedge$ is replaced by the arithmetic product. A more detailed discussion of these compositions may be found in [24].

If U, V and W are finite sets, then the relation matrix for R∘S is the max-min product[5] of the relation matrices for R and S. For example, the max-min product of the relation matrices on the left-hand side of (3.56) is given by the right-hand member of (3.56)

$$\overset{R}{\begin{bmatrix} 0.3 & 0.8 \\ 0.6 & 0.9 \end{bmatrix}} \circ \overset{S}{\begin{bmatrix} 0.5 & 0.9 \\ 0.4 & 1 \end{bmatrix}} = \overset{R\circ S}{\begin{bmatrix} 0.4 & 0.8 \\ 0.5 & 0.9 \end{bmatrix}} \tag{3.56}$$

<u>Projections and Cylindrical Fuzzy Sets</u>

If R is an n-ary fuzzy relation in $U_1 \times \ldots \times U_n$, then its <u>projection</u> <u>(shadow)</u> on $U_{i_1} \times \ldots \times U_{i_k}$ is a k-ary fuzzy relation R_q in U which is defined by (compare with (2.12))

$$\begin{aligned} R_q &\triangleq \text{Proj } R \text{ on } U_{i_1} \times \ldots \times U_{i_k} \qquad (3.57) \\ &\triangleq P_q R \\ &\triangleq \int_{U_{i_1} \times \ldots \times U_{i_k}} (\vee_{u_{(q')}} \mu_R(u_1,\ldots,u_n))/(u_{i_1},\ldots,u_{i_k}) \end{aligned}$$

where q is the index sequence $(i_1,\ldots,i_k)$; $u_{(q)} \triangleq (u_{i_1},\ldots,u_{i_k})$; q' is the complement of q; and $\vee_{u_{(q')}}$ is the supremum of $\mu_R(u_1,\ldots,u_n)$ over the u's which are in $u_{(q')}$. It should be noted that when R is a nonfuzzy relation, (3.57) reduced to (2.9).

[5]In the max-min matrix product, the operations of addition and multiplication are replaced by $\vee$ and $\wedge$, respectively.

Example 3.11 For the fuzzy relation defined by the relation matrix (3.54), we have

$$R_1 = 1/2 + 0.8/2 + 0.3/3$$

and

$$R_2 = 0.3/2 + 0.8/3 + 1/4$$

It is clear that distinct fuzzy relations in $U_1 \times \ldots \times U_n$ can have identical projections on $U_{i_1} \times \ldots \times U_{i_k}$. However, given a fuzzy relation R_q in $U_{i_1} \times \ldots \times U_{i_k}$, there exists a unique largest[6] relation $\bar{R}_q$ in $U_1 \times \ldots \times U_n$ whose projection on $U_{i_1} \times \ldots \times U_{i_k}$ is R_q. In consequence of (3.57), the membership function of $\bar{R}_q$ is given by

$$\mu_{\bar{R}_q}(u_1,\ldots,u_n) = \mu_{R_q}(u_{i_1},\ldots,u_{i_k}) \tag{3.58}$$

with the understanding that (3.58) holds for all $u_1,\ldots,u_n$ such that the $i_1,\ldots,i_k$ arguments in $\mu_{\bar{R}_q}$ are equal, respectively, to the first, second, ..., kth arguments in μ_{R_q}. This implies that the value of $\mu_{\bar{R}_q}$ at the point $(u_1,\ldots,u_n)$ is the same as that at the point $(u'_1,\ldots,u'_n)$ provided that $u_{i_1} = u'_{i_1},\ldots,u_{i_k} = u'_{i_k}$. For this reason, $\bar{R}_q$ will be referred to as the cylindrical extension of R_q, with R_q constituting the base of $\bar{R}_q$. (See Fig. 3.2.)

Suppose that R is an n-ary relation in $U_1 \times \ldots \times U_n$, R_q is its projection on $U_{i_1} \times \ldots \times U_{i_k}$, and $\bar{R}_q$ is the cylindrical extension of R_q. Since $\bar{R}_q$ is the largest relation in $U_1 \times \ldots \times U_n$ whose projection on $U_{i_1} \times \ldots \times U_{i_k}$ is R_q, it follows that R_q satisfies the containment

[6]That is, a relation which contains all other relations whose projection on $U_{i_1} \times \ldots \times U_{i_k}$ is R_q.

relation

$$R \subset \overline{R}_q \tag{3.59}$$

for all q, and hence

$$R \subset \overline{R}_{q_1} \cap \overline{R}_{q_2} \cap \ldots \cap \overline{R}_{q_r} \tag{3.60}$$

for arbitrary $q_1, \ldots, q_r$ (index subsequences of $(1,2,\ldots,n)$).

In particular, if we set $q_1 = 1, \ldots, q_r = n$, then (3.60) reduces to

$$R \subset \overline{R}_1 \cap \overline{R}_2 \cap \ldots \cap \overline{R}_n \tag{3.61}$$

where $R_1, \ldots, R_n$ are the projections of R on $U_1, \ldots, U_n$, respectively, and $\overline{R}_1, \ldots \overline{R}_n$ are their cylindrical extensions. But, from the definition of the cartesian product (see (3.45)) it follows that

$$\overline{R}_1 \cap \ldots \cap \overline{R}_n = R_1 \times \ldots \times R_n \tag{3.62}$$

which leads us to the

Proposition 3.12 If R is an n-ary fuzzy relation in $U_1 \times \ldots \times U_n$ and $R_1, \ldots, R_n$ are its projections on $U_1, \ldots, U_n$, then (see Fig. 3.3 for illustration)

$$R \subset R_1 \times \ldots \times R_n \tag{3.63}$$

The concept of a cylindrical extension can also be used to provide an intuitively appealing interpretation of the composition of fuzzy relations. Thus, suppose that R and S are binary fuzzy relations in $U_1 \times U_2$ and $U_2 \times U_3$, respectively. Let $\overline{R}$ and $\overline{S}$ be the cylindrical extensions of R and S in $U_1 \times U_2 \times U_3$. Then, from the definition of $R \circ S$ (see (3.55)) it follows that

$$R \circ S = \text{Proj } \overline{R} \cap \overline{S} \text{ on } U_1 \times U_3 \tag{3.64}$$

If R and S are such that

$$\text{Proj } R \text{ on } U_2 = \text{Proj } S \text{ on } U_2 \tag{3.65}$$

then $\overline{R} \cap \overline{S}$ becomes the join[7] of R and S. A basic property of the join of R and S may be stated as the

Proposition 3.13 If R and S are fuzzy relations in $U_1 \times U_2$ and $U_2 \times U_3$, respectively, and $\overline{R} \cap \overline{S}$ is the join of R and S, then

$$R = \text{Proj } \overline{R} \cap \overline{S} \text{ on } U_1 \times U_2 \tag{3.66}$$

and

$$S = \text{Proj } \overline{R} \cap \overline{S} \text{ on } U_2 \times U_3 \tag{3.67}$$

Thus, R and S can be retrieved from the join of R and S.

Proof. Let μ_R and μ_S denote the membership functions of R and S, respectively. Then the right-hand members of (3.66) and (3.67) translate into

$$\vee_{u_3}(\mu_R(u_1,u_2) \wedge \mu_S(u_2,u_3)) \tag{3.68}$$

and

[7]The concept of the join of nonfuzzy relations was introduced by E. F. Codd in [28].

$$\vee_{u_1}(\mu_R(u_1,u_2)\wedge \mu_s(u_2,u_3)) \tag{3.69}$$

In virtue of the distributivity and commutativity of $\vee$ and $\wedge$, (3.68) and (3.69) may be rewritten as

$$\mu_R(u_1,u_2) \wedge (\vee_{u_3}\mu_s(u_2,u_3)) \tag{3.70}$$

and

$$\mu_s(u_2,u_3) \wedge (\vee_{u_1}\mu_R(u_1,u_2)) \tag{3.71}$$

Furthermore, the definition of the join implies (3.65) and hence that

$$\vee_{u_1}\mu_R(u_1,u_2) = \vee_{u_3}\mu_s(u_2,u_3) \tag{3.72}$$

From this equality and the definition of $\vee$ it follows that

$$\mu_R(u_1,u_2) \leq \vee_{u_1}\mu_R(u_1,u_2) = \vee_{u_3}\mu_S(u_2,u_3) \tag{3.73}$$

and

$$\mu_S(u_2,u_3) \leq \vee_{u_3}\mu_S(u_2,u_3) = \vee_{u_1}\mu_R(u_1,u_2) \tag{3.74}$$

Consequently

$$\mu_R(u_1,u_2) \wedge (\vee_{u_3}\mu_S(u_2,u_3)) = \mu_R(u_1,u_2) \tag{3.75}$$

and

$$\mu_S(u_2,u_3) \wedge (\vee_{u_1}\mu_R(u_1,u_3) = \mu_S(u_2,u_3) \tag{3.76}$$

which translate into (3.66) and (3.67). Q.E.D.

A basic property of projections which we shall have an occasion to use in Sec. 4 is the following.

Proposition 3.14 If R is a normal relation (see (3.23)), then so is every projection of R.

Proof. Let R be an n-ary relation in $U_1 \times \ldots \times U_n$, and let R_q be its projection (shadow) on $U_{i_1} \times \ldots \times U_{i_k}$, with $q = (i_1,\ldots,i_k)$. Since R is normal, we have by (3.23),

$$\bigvee_{(u_1,\ldots,u_n)} \mu_R(u_1,\ldots,u_n) = 1 \tag{3.77}$$

or more compactly

$$\bigvee_u \mu_R(u) = 1$$

On the other hand, by the definition of R_q (see 3.57))

$$\mu_{R_q}(u_{i_1},\ldots,u_{i_k}) = \bigvee_{(u_{j_1},\ldots,u_{j_n})} \mu_R(u_1,\ldots,u_n)$$

or

$$\mu_{R_q}(u_{(q)}) = \bigvee_{u_{(q')}} \mu_R(u)$$

and hence the height of R_q is given by

$$\bigvee_{u_{(q)}} \mu_{R_q}(u_{(q)}) = \bigvee_{u_{(q)}} \bigvee_{u_{(q')}} \mu_R(u) \tag{3.78}$$

$$= \bigvee_u \mu_R(u)$$

$$= 1 \qquad \text{Q.E.D.}$$

The Extension Principle

The extension principle for fuzzy sets is in essence a basic identity which allows the domain of the definition of a mapping or a relation to be extended from points in U to fuzzy subsets of U. More specifically, suppose that f is a mapping from U to V and A is a fuzzy subset of U expressed as

$$A = \mu_1 u_1 + \ldots + \mu_n u_n . \tag{3.79}$$

Then, the extension principle asserts that[8]

$$f(A) = f(\mu_1 u_1 + \ldots + \mu_n u_n) \equiv \mu_1 f(u_1) + \ldots + \mu_n f(u_n) \tag{3.80}$$

Thus, the image of A under f can be deduced from the knowledge of the images of $u_1, \ldots, u_n$ under f.

<u>Example 3.15</u> Let

$$U = 1 + 2 + .. + 10$$

and let f be the operation of squaring. Let <u>small</u> be a fuzzy subset of U defined by

$$\underline{\text{small}} = 1/1 + 1/2 + 0.8/3 + 0.6/4 + 0.4/5 \tag{3.81}$$

Then, in consequence of (3.80), we have[9]

$$\underline{\text{small}}^2 = 1/1 + 1/4 + 0.8/9 + 0.6/16 + 0.4/25 \tag{3.82}$$

If the support of A is a continuum, that is

---

[8] The extension principle is implicit in a result given in [29]. In probability theory, the extension principle is analogous to the expression for the probability distribution induced by a mapping [30]. In the special case of intervals, the results of applying the extension principle reduce to those of interval analysis [31].

[9] Note that this definition of <u>small</u>^{2} differs from that of (3.38).

$$A = \int_U \mu_A(u)/u \tag{3.83}$$

then the statement of the extension principle assumes the following form

$$f(A) = f\left(\int_U \mu_A(u)/u\right) \equiv \int_V \mu_A(u)/f(u) \tag{3.84}$$

with the understanding that f(u) is a point in V and $\mu_A(u)$ is its grade of membership in f(A), which is a fuzzy subset of V.

In some applications it is convenient to use a modified form of the extension principle which follows from (3.84) by decomposing A into its constituent level-sets rather than its fuzzy singletons (see the resolution identity (3.28)). Thus, on writing

$$A = \int_0^1 \alpha A_\alpha \tag{3.85}$$

where A_α is an α-level-set of A, the statement of the extension principle assumes the form

$$f(A) = f\left(\int_0^1 \alpha A_\alpha\right) \equiv \int_0^1 \alpha f(A_\alpha) \tag{3.86}$$

when the support of A is a continuum, and

$$f(A) = f\left(\sum_\alpha \alpha A_\alpha\right) = \sum_\alpha \alpha f(A_\alpha) \tag{3.87}$$

when either the support of A is a countable set or the distinct level-sets of A form a countable collection.

Comment 3.16 Written in the form (3.84), the extension principle extends the domain of definition of f from points in U to fuzzy subsets of U. By contrast, (3.86) extends the domain of definition of f from nonfuzzy subsets of U to fuzzy subsets of U. It should be clear, however, that (3.84) and (3.86) are equivalent, since (3.86) results from (3.84) by a regrouping of terms in the representation of A.

Comment 3.17 The extension principle is analogous to the superposition principle for linear systems. Under the latter principle, if L is a linear system and $u_1,\ldots,u_n$ are inputs to L, then the response of L to any linear combination

$$u = w_1u_1 + \ldots + w_nu_n \tag{3.88}$$

where the w_i are constant coefficients, is given by

$$L(u) = L(w_1u_1 + \ldots + w_nu_n) = w_1L(u_1) + \ldots + w_nL(u_n). \tag{3.89}$$

The important point of difference between (3.89) and (3.80) is that in (3.80) + is the union rather than the arithmetic sum and f is not restricted to linear mappings.

Comment 3.18 It should be noted that when $A = u_1 + \ldots + u_n$, the result of applying the extension principle is analogous to that of forming the n-fold cartesian product of the algebraic system (U,f) with itself. (An extension of the multiplication table is shown in Table 3.4.)

In many applications of the extension principle, one encounters the following problem. We have an n-ary function, f, which is a mapping from a cartesian product $U_1 \times \ldots \times U_n$ to a space V, and a fuzzy set (relation)

A in $U_1 \times \ldots \times U_n$ which is characterized by a membership function $\mu_A(u_1,\ldots,u_n)$, with u_i, $i = 1,\ldots,n$, denoting a generic point in U_i. A direct application of the extension principle (3.84) to this case yields

$$f(A) = f\left(\int_{U_1 \times \ldots \times U_n} \mu_A(u_1,\ldots,u_n)/(u_1,\ldots,u_n)\right) \tag{3.90}$$

$$= \int_V \mu_A(u_1,\ldots,u_n)/f(u_1,\ldots,u_n)$$

However, in many instances what we know is not A but its projections $A_1,\ldots,A_n$ on $U_1,\ldots,U_n$, respectively (see (3.57)). The question that arises, then, is: What expression for μ_A should be used in (3.90)?

In such cases, unless otherwise specified we shall assume that the membership function of A is expressed by

$$\mu_A(u_1,\ldots,u_n) = \mu_{A_1}(u_1) \wedge \mu_{A_2}(u_2) \wedge \ldots \wedge \mu_{A_n}(u_n) \tag{3.91}$$

where μ_{A_i}, $i = 1,\ldots,n$, is the membership function of A_i. In view of (3.45), this is equivalent to assuming that A is the cartesian product of its projections, i.e.,

$$A = A_1 \times \ldots \times A_n$$

which in turn implies that A is the largest set whose projections on $U_1,\ldots,U_n$ are $A_1,\ldots,A_n$, respectively. (See (3.63).)

<u>Example 3.19</u> Suppose that, as in Example (3.15),

$$U_1 = U_2 = 1 + 2 + 3 + \ldots + 10$$

and

$$A_1 = \underset{\sim}{2} \triangleq \underline{\text{approximately}}\ 2 = 1/2 + 0.6/1 + 0.8/3 \tag{3.92}$$

$$A_2 = \underset{\sim}{6} \triangleq \underline{\text{approximately}}\ 6 = 1/6 + 0.8/5 + 0.7/7 \tag{3.93}$$

and

$$f(u_1,u_2) = u_1 \times u_2 = \text{arithmetic product of } u_1 \text{ and } u_2$$

Using (3.91) and applying the extension principle as expressed by (3.90) to this case, we have

$$\underset{\sim}{2} \times \underset{\sim}{6} = (1/2 + 0.6/1 + 0.8/3) \times (1/6 + 0.8/5 + 0.7/7) \tag{3.94}$$

$$= 1/12 + 0.8/10 + 0.7/14 +$$

$$0.6/6 + 0.6/5 + 0.6/7 +$$

$$0.8/18 + 0.8/15 + 0.7/21$$

$$= 0.6/5 + 0.6/6 + 0.6/7 + 0.8/10 + 1/12 +$$

$$0.7/14 + 0.8/15 + 0.8/18 + 0.7/21$$

Thus, the arithmetic product of the fuzzy numbers approximately 2 and approximately 6 is a fuzzy number given by (3.94).

More generally, let * be a binary operation defined on $U \times V$ with values in W. Thus, if $u \in U$ and $v \in V$, then

$$w = u * v, \qquad w \in W$$

Now suppose that A and B are fuzzy subsets of U and V, respectively, with

$$A = \mu_1 u_1 + \ldots + \mu_n u_n \tag{3.95}$$

and

$$B = \nu_1 v_1 + \ldots + \nu_m v_m$$

By using the extension principle under the assumption (3.91), the operation * may be extended to fuzzy subsets of U and V by the defining relation

$$\begin{aligned} A * B &= (\sum_i \mu_i u_i) * (\sum_j \nu_j v_j) \\ &= \sum_{i,j} (\mu_i \wedge \nu_j)(u_i * v_j) \end{aligned} \tag{3.96}$$

It is easy to verify that for the case where $A = \underset{\sim}{2}$, $B = \underset{\sim}{6}$ and $* = \times$, as in Example 3.19, the application of (3.96) yields the expression for $\underset{\sim}{2} \times \underset{\sim}{6}$.

Comment 3.20 It is important to note that the validity of (3.97) depends in an essential way on the assumption (3.91), that is

$$\mu_{(A,B)}(u,v) = \mu_A(u) \wedge \mu_B(v)$$

The implication of this assumption is that u and v are noninteractive in the sense of Definition 2.9. Thus, if there is a constraint on (u,v) which is expressed as a relation R with a membership function μ_R, then the expression for A * B becomes

$$\begin{aligned} A * B &= \left((\sum_i \mu_i u_i) * (\sum_j \nu_j v_j)\right) \cap R \\ &= \sum_{i,j} (\mu_i \wedge \nu_j \wedge \mu_R(u_i, v_j))(u_i * v_j) \end{aligned} \tag{3.97}$$

Note that if R is a nonfuzzy relation, then the right-hand member of (3.97) will contain only those terms which satisfy the constraint R.

A simple illustration of a situation in which u and v are interactive is provided by the expression

$$w = z \times (x+y) \tag{3.98}$$

in which $+ \triangleq$ arithmetic sum and $\times \triangleq$ arithmetic product. If x, y and z are noninteractive, then we can apply the extension principle in the form (3.96) to the computation of $A \times (B + C)$, where A, B and C are fuzzy subsets of the real line. On the other hand, if (3.98) is rewritten as

$$w = z \times x \quad + z \times y$$

then the terms $z \times x$ and $z \times y$ are interactive by virtue of the common factor z, and hence

$$A \times (B + C) \neq A \times B + A \times C \tag{3.99}$$

A significant conclusion that can be drawn from this observation is that the product of fuzzy numbers is not distributive if it is computed by the use of (3.96). To obtain equality in (3.99), we may apply the unrestricted form of the extension principle (3.96) to the left-hand member of (3.99), and must apply the restricted form (3.97) to its right-hand member.

<u>Remark 3.21</u> The extension principle can be applied not only to functions, but also to relations or, equivalently, to predicates. We shall not discuss this subject here, since the application of the extension principle to relations does not play a significant role in the present paper.

Fuzzy Sets With Fuzzy Membership Functions

Our consideration of fuzzy sets with fuzzy membership functions is motivated by the close association which exists between the concept of a linguistic truth with truth-values such as true, quite true, very true, more or less true, etc., on the one hand, and fuzzy sets in which the grades of membership are specified in linguistic terms such as low, medium, high, very low, not low and not high, etc., on the other.

Thus, suppose that A is a fuzzy subset of a universe of discourse U, and the values of the membership function, μ_A, of A are allowed to be fuzzy subsets of the interval [0,1]. To differentiate such fuzzy sets from those considered previously, we shall refer to them as fuzzy sets of *type* 2, with the fuzzy sets whose membership functions are mappings from U to [0,1] classified as *type* 1. More generally:

Definition 3.22 A fuzzy set is of type n, n = 2,3,..., if its membership function ranges over fuzzy sets of type n-1. The membership function of a fuzzy set of type 1 ranges over the interval [0,1].

To define such operations as complementation, union, intersection, etc. for fuzzy sets of type 2, it is natural to make use of the extension principle. It is convenient, however, to accomplish this in two stages: first, by extending the type 1 definitions to fuzzy sets with interval-valued membership functions; and second, generalizing from intervals to fuzzy sets[10] by the use of the level-set form of the extension principle (see (3.86)). In what follows, we shall illustrate this technique by

[10] We are tacitly assuming that the fuzzy sets in question are convex, that is, have intervals as level-sets (see [29]). Only minor modifications are needed when the sets are not convex.

extending to fuzzy sets of type 2 the concept of intersection - which is defined for fuzzy sets of type 1 by (3.35).

Our point of departure is the expression for the membership function of the intersection of A and B, where A and B are fuzzy subsets of type 1 of U:

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

Now if $\mu_A(u)$ and $\mu_B(u)$ are intervals in [0,1] rather than points in [0,1], that is, for a fixed u

$$\mu_A(u) = [a_1, a_2]$$

$$\mu_B(u) = [b_1, b_2]$$

where a_1, a_2, b_1 and b_2 depend on u, then the application of the extension principle (3.86) to the function $\wedge$ (Min) yields

$$[a_1, a_2] \wedge [b_1, b_2] = [a_1 \wedge b_1, a_2 \wedge b_2] \qquad (3.100)$$

Thus, if A and B have interval-valued membership functions as shown in Fig. 3.5, then their intersection is an interval-valued curve whose value for each u is given by (3.100).

Next, let us consider the case where, for each u, $\mu_A(u)$ and $\mu_B(u)$ are fuzzy subsets of the interval [0,1]. For simplicity, we shall assume that these subsets are <u>convex</u>, that is, have intervals as level-sets. In other words, we shall assume that, for each α in (0,1], the α-level sets of μ_A and μ_B are interval-valued membership functions. (See Fig. 3.6.)

By applying the level-set form of the extension principle (3.86) to the α-level sets of μ_A and μ_B we are led to the following definition of the intersection of fuzzy sets of type 2.

Definition 3.23 Let A and B be fuzzy subsets of type 2 of U such that, for each $u \in U$, $\mu_A(u)$ and $\mu_B(u)$ are convex fuzzy subsets of type 1 of [0,1], which implies that, for each α in (0,1], the α-level sets of the fuzzy membership functions μ_A and μ_B are interval-valued membership functions μ_A^α and μ_B^α.

Let the α-level-set of the fuzzy membership function of the intersection of A and B be denoted by $\mu^\alpha_{A \cap B}$, with the α-level-sets μ_A^α and μ_B^β defined for each u by

$$\mu_A^\alpha \triangleq \{v \mid \nu_A(v) \geq \alpha\} \tag{3.101}$$

$$\mu_B^\alpha \triangleq \{v \mid \nu_B(v) \geq \alpha\} \tag{3.102}$$

where $\nu_A(v)$ denotes the grade of membership of a point v, $v \in [0,1]$, in the fuzzy set $\mu_A(u)$, and likewise for μ_B. Then, for each u,

$$\mu^\alpha_{A\cap B} = \mu_A^\alpha \wedge \mu_B^\alpha \tag{3.103}$$

In other words, the α-level-set of the fuzzy membership function of the intersection of A and B is the minimum (in the sense of (3.100)) of the α-level-sets of the fuzzy membership functions of A and B. Thus, using the resolution identity (3.28), we can express $\mu_{A\cap B}$ as

$$\mu_{A\cap B} = \int_0^1 \alpha(\mu_A^\alpha \wedge \mu_B^\alpha) \tag{3.104}$$

For the case where μ_A and μ_B have finite supports, that is, μ_A and

μ_B are of the form

$$\mu_A = \alpha_1 v_1 + \dots \alpha_n v_n , \quad v_i \in [0,1], \quad i = 1,\dots,n \tag{3.105}$$

and

$$\mu_B = \beta_1 w_1 + \dots + \beta_m w_m , \quad w_j \in [0,1], \quad j = 1,\dots,m \tag{3.106}$$

where α_i and β_j are the grades of membership of v_i and w_j in μ_A and μ_B, respectively, the expression for $\mu_{A \cap B}$ can readily be derived by employing the extension principle in the form (3.96). Thus, by applying (3.96) to the operation $\wedge$ ($\triangleq$ Min), we obtain at once

$$\begin{aligned} \mu_{A\cap B} &= \mu_A \wedge \mu_B & (3.107) \\ &= (\alpha_1 v_1 + \dots + \alpha_n v_n) \wedge (\beta_1 w_1 + \dots + \beta_m w_m) \\ &= \sum_{i,j} (\alpha_i \wedge \beta_j)(v_i \wedge w_j) \end{aligned}$$

as the desired expression for $\mu_{A\cap B}$.[11]

Example 3.24 As a simple illustration of (3.104), suppose that at a point u the grades of membership of u in A and B are labeled as high and medium, respectively, with high and medium defined as fuzzy subsets of V = 0 + 0.1 + 0.2 + ... + 1 by the expressions

$$\underline{high} = 0.8/0.8 + 0.8/0.9 + 1/1 \tag{3.108}$$

$$\underline{medium} = 0.6/0.4 + 1/0.5 + 0.6/0.6 \tag{3.109}$$

The level sets of high and medium are expressed by

[11]Actually, Definition 3.23 can be deduced from (3.90).

$$\underline{\text{high}}_{0.6} = 0.8 + 0.9 + 1$$

$$\underline{\text{high}}_{0.8} = 0.9 + 1$$

$$\underline{\text{high}}_{1} = 1$$

$$\underline{\text{medium}}_{0.6} = 0.4 + 0.5 + 0.6$$

$$\underline{\text{medium}}_{1} = 0.5$$

and consequently the α-level-sets of the intersection are given by

$$\mu^{0.6}_{A \cap B}(u) = \underline{\text{high}}_{0.6} \wedge \underline{\text{medium}}_{0.6} \tag{3.110}$$

$$= (0.8 + 0.9 + 1) \wedge (0.4 + 0.5 + 0.6)$$

$$= 0.4 + 0.5 + 0.6$$

$$\mu^{0.8}_{A \cap B}(u) = \underline{\text{high}}_{0.8} \wedge \underline{\text{medium}}_{0.8} \tag{3.111}$$

$$= (0.9 + 1) \wedge 0.5$$

$$= 0.5$$

and

$$\mu^{1}_{A \cap B}(u) = \underline{\text{high}}_{1} \wedge \underline{\text{medium}}_{1} \tag{3.112}$$

$$= 1 \wedge 0.5$$

$$= 0.5$$

Combining (3.110), (3.111) and (3.112), the fuzzy set representing the grade of membership of u in the intersection of A and B is found to be

$$\mu_{A \cap B}(u) = 0.6/(0.4 + 0.5 + 0.6) + 1/0.5 \tag{3.113}$$

$$= \underline{\text{medium}}$$

which is equivalent to the statement

$$\underline{\text{high}} \wedge \underline{\text{medium}} = \underline{\text{medium}} \tag{3.114}$$

The same result can be obtained more expeditiously by the use of (3.107). Thus, we have

$$\text{high} \wedge \text{medium} = (0.8/0.8 + 0.8/0.9 + 1/1) \wedge (0.6/0.4 + 1/0.5 + 0.6/0.6)$$

$$= 0.6/0.4 + 1/0.5 + 0.6/0.6 \tag{3.115}$$

$$= \underline{\text{medium}}$$

In a similar fashion, we can extend to fuzzy sets of type 2 the operations of complementation, union, concentration, etc. This will be done in Sec. 6, in conjunction with our discussion of a fuzzy logic in which the truth-values are linguistic in nature.

Remark 3.25 The results derived in Example 3.24 may be viewed as an instance of a general conclusion that can be drawn from (3.100) concerning an extension of the inequality $\leq$ from real numbers to fuzzy subsets of the real line. Specifically, in the case of real numbers a, b, we have the equivalence

$$a \leq b \Leftrightarrow a \wedge b = a \tag{3.116}$$

Using this as a basis for the extension of $\leq$ to intervals, we have in virtue of (3.100),

$$[a_1,a_2] \leq [b_1,b_2] \;\Leftrightarrow\; a_1 \leq b_1 \text{ and } a_2 \leq b_2. \qquad (3.117)$$

This, in turn, leads us to the following definition.

Definition 3.26. Let A and B be convex fuzzy subsets of the real line, and let A_α and B_α denote the α-level-sets of A and B, respectively. Then an extension of the inequality $\leq$ to convex fuzzy subsets of the real line is expressed by[11]

$$A \leq B \;\Leftrightarrow\; A \wedge B = A \qquad (3.118)$$

$$\Leftrightarrow\; A_\alpha \wedge B_\alpha = A_\alpha \quad \text{for all } \alpha \text{ in } [0,1] \qquad (3.119)$$

where $A_\alpha \wedge B_\alpha$ is defined by (3.100).

In the case of Example 3.24, it is easy to verify by inspection that

$$\underline{\text{medium}}_\alpha \leq \underline{\text{high}}_\alpha \text{ for all } \alpha \qquad (3.120)$$

in the sense of (3.119), and hence we can conclude at once that

$$\underline{\text{medium}} \wedge \underline{\text{high}} = \underline{\text{medium}} \qquad (3.121)$$

which is in agreement with (3.114).

[11] It can be readily be verified that $\leq$ as defined by (3.117) constitutes a partial ordering.

4. The Concept of a Fuzzy Variable

We are now in a position to generalize the concepts introduced in Sec. 2 to what might be called *fuzzy* variables. For our purposes, it will be convenient to formalize the concept of a fuzzy variable in a way that parallels the characterization of a nonfuzzy variable as expressed by Definition 2.1. Specifically:

Definition 4.1 A fuzzy variable is characterized by a triple $(X,U,R(X;u))$, in which X is the name of the variable; U is a universe of discourse (finite or infinite set); u is a generic name for the elements of U; and $R(X;u)$ is a fuzzy subset of U which represents a fuzzy *restriction* on the values of u imposed by X. (As in the case of nonfuzzy variables, $R(X;u)$ will usually be abbreviated to $R(X)$ or $R(u)$ or $R(x)$, where x denotes a generic name for the values of X, and $R(X;u)$ will be referred to as the restriction *on* u or the restriction *imposed by* X.) The nonrestricted nonfuzzy variable u constitutes the *base variable* for X.

The *assignment equation* for X has the form

$$x = u: \quad R(X) \tag{4.1}$$

and represents an assignment of a value u to x subject to the restriction $R(X)$.

The degree to which this equation is satisfied will be referred to as the *compatibility of* u *with* $R(X)$ and will be denoted by $c(u)$. By definition,

$$c(u) = \mu_{R(X)}(u), \quad u \in U \tag{4.2}$$

where $\mu_{R(X)}(u)$ is the grade of membership of u in the restriction R(X).

Comment 4.2 It is important to observe that the compatibility of u is not the same as the probability of u. Thus, the compatibility of u with R(X) is merely a measure of the degree to which u satisfies the restriction R(X) and has no relation to how probable or improbable u happens to be.

Comment 4.3 In terms of the valise analogy (see Comment 2.4), a fuzzy variable may be likened to a tagged valise with soft sides, with X representing the name on the tag, U corresponding to a list of objects which can be put in a valise, and R(X) representing a sublist of U in which each object u is associated with a number c(u) representing the degree of ease with which u can be fitted in valise X. (Fig. 4.1.)

In order to simplify the notation it is convenient to use the same symbol for both X and x, relying on the context for disambiguation. We do this in the following example.

Example 4.4 Consider a fuzzy variable named budget, with U = $[0,\infty)$ and R(X) defined by (see Fig. 4.2)

$$R(\underline{budget}) = \int_0^{1000} 1/u + \int_{1000}^{\infty} \left(1 + \left(\frac{u-1000}{200}\right)^2\right)^{-1} / u \qquad (4.3)$$

Then, in the assignment equation

$$\underline{budget} = 1100: \quad R(\underline{budget}) \qquad (4.4)$$

the compatibility of 1100 with the restriction imposed by budget is

$$c(1100) = \mu_{R(\underline{budget})}(1100) \qquad (4.5)$$
$$= 0.80$$

As in the case of nonfuzzy variables, if $X_1,\ldots,X_n$ are fuzzy variables in $U_1,\ldots,U_n$, respectively, then $X \triangleq (X_1,\ldots,X_n)$ is an <u>n-ary composite (joint) variable</u> in $U = U_1 \times \ldots \times U_n$. Correspondingly, in the <u>n-ary assignment equation</u>

$$(x_1,\ldots,x_n) = (u_1,\ldots,u_n): \quad R(X_1,\ldots,X_n) \qquad (4.6)$$

x_i, $i = 1,\ldots,n$, is a generic name for the values of X_i; u_i is a generic name for the elements of U_i; and $R(X) \triangleq R(X_1,\ldots,X_n)$ is an n-ary fuzzy relation in U which represents the <u>restriction</u> imposed by $X \triangleq (X_1,\ldots,X_n)$. The <u>compatibility of</u> $(u_1,\ldots,u_n)$ <u>with</u> $R(X_1,\ldots,X_n)$ is defined by

$$c(u_1,\ldots,u_n) = \mu_{R(X)}(u_1,\ldots,u_n) \qquad (4.7)$$

where $\mu_{R(X)}$ is the membership function of the restriction on $u \triangleq (u_1,\ldots, u_n)$.

<u>Example 4.5</u> Suppose that $U_1 = U_2 = (-\infty,\infty)$; $X_1 \triangleq$ <u>horizontal proximity</u>; $X_2 \triangleq$ <u>vertical proximity</u>; and the restriction on u is expressed by

$$R(X) = \int_{U_1 \times U_2} \left(1 + u_1^2 + u_2^2\right)^{-1} / (u_1,u_2) \qquad (4.8)$$

Then the compatibility of the value $u = (2,1)$ in the assignment equation

$$(x_1,x_2) = (2,1): \ R(X) \qquad (4.9)$$

is given by

$$c(2,1) = \mu_{(R(X)}(2,1) \qquad (4.10)$$

$$= 0.16$$

<u>Comment 4.6</u> In terms of the valise analogy (see Comment 4.3), an n-ary

composite fuzzy variable may be likened to a soft valise named X with n compartments named $X_1,\ldots,X_n$. The compatibility function $c(u_1,\ldots,u_n)$ represents the degree of ease with which objects $u_1,\ldots,u_n$ can be put into respective compartments $X_1,\ldots,X_n$ simultaneously. (Fig. 4.3.)

A basic question that arises in connection with an n-ary assignment equation relates to its decomposition into a sequence of n unary assignment equations, as in (2.21). In the case of fuzzy variables, the process of decomposition is somewhat more involved, and we shall take it up after defining marginal and conditioned restrictions.

Marginal and Conditioned Restrictions

In Sec. 2, the concepts of marginal and conditioned restrictions were intentionally defined in such a way as to make them easy to extend to fuzzy restrictions. Thus, in the more general context of fuzzy variables, these concepts can be formulated in almost exactly the same terms as in Sec. 2. This is what we shall do in the sequel.

Note 4.7 As we have seen in our earlier discussion of the notions of marginal and conditioned restrictions in Sec. 2, it is convenient to simplify the representation of n-tuples by employing the following notation.

Let

$$q \triangleq (i_1,\ldots,i_k) \tag{4.11}$$

be an ordered subsequence of the index sequence $(1,\ldots,n)$. E.g., for $n = 7$, $q = (2,4,5)$.

The ordered complement of q is denoted by

$$q' = (j_1,\ldots,j_m) \tag{4.12}$$

E.g., for $q = (2,4,5)$, $q' = (1,3,6,7)$.

A k-tuple of variables such as $(v_{i_1},\ldots,v_{i_k})$ is denoted by $v_{(q)}$. Thus

$$v_{(q)} \triangleq (v_{i_1},\ldots,v_{i_k}) \tag{4.13}$$

and similarly

$$v_{(q')} \triangleq (v_{j_1},\ldots,v_{j_m}) \tag{4.14}$$

For example, if

$$v_{(q)} = (v_2,v_4,v_5)$$

then

$$v_{(q')} = (v_1,v_3,v_6,v_7)$$

If $k = n$, we shall write more simply

$$v = (v_1,\ldots,v_n) \tag{4.15}$$

This notation will be used in the following without further explanation.

<u>Definition 4.8</u> An n-ary restriction $R(X_1,\ldots,X_n)$ in $U_1 \times \ldots \times U_n$ induces a k-ary <u>marginal restriction</u> $R(X_{i_1},\ldots,X_{i_k})$ which is defined as the projection (shadow) of $R(X_1,\ldots,X_n)$ on $U_{i_1} \times \ldots \times U_{i_k}$. Thus, using the definition of projection (see (3.57)) and employing the notation of Note 4.7, we can express the membership function of the marginal restriction $R(X_{i_1},\ldots,X_{i_k})$ as

$$\mu_{R(X_{(q)})}(u_{(q)}) = \bigvee_{u_{(q')}} \mu_{R(X)}(u) \tag{4.16}$$

Example 4.9 For the fuzzy binary variable defined in Example 4.5, we have

$$R_1 \triangleq R(X_1)$$

$$R_2 \triangleq R(X_2)$$

$$\mu_{R_1}(u_1) = \vee_{u_2}(1 + u_1^2 + u_2^2)^{-1}$$

$$= (1 + u_1^2)^{-1}$$

$$\mu_{R_2} = \mu_{R_1}$$

Example 4.10 Assume that

$$U_1 = U_2 = U_3 = 0 + 1 + 2$$

and $R(X_1,X_2,X_3)$ is a ternary fuzzy relation in $U_1 \times U_2 \times U_3$ expressed by

$$R(X_1,X_2,X_3) = 0.8/(0,0,0) + 0.6/(0,0,1) + 0.2/(0,1,0) \quad (4.17)$$

$$+ 1/(1,0,2) + 0.7/(1,1,0) + 0.4/(0,1,1)$$

$$+ 0.9/(1,2,0) + 0.4/(2,1,1) + 0.8/(1,1,2)$$

Applying (4.16) to (4.17), we obtain

$$R(X_1,X_2) = 0.8/(0,0) + 0.4/(0,1) + 1/(1,0) \quad (4.18)$$

$$+ 0.8/(1,1) + 0.9/(1,2) + 0.4/(2,1)$$

and

$$R(X_1) = 0.8/0 + 1/1 + 0.4/2 \quad (4.19)$$

$$R(X_2) = 1/0 + 0.8/1 + 0.9/2$$

Definition 4.11 Let $R(X_1,\ldots,X_n)$ be a restriction on $(u_1,\ldots,u_n)$ and let $u^o_{i_1},\ldots,u^o_{i_k}$, be particular values of $u_{i_1},\ldots,u_{i_k}$, respectively. If in the membership function of $R(X_1,\ldots,X_n)$, the values of $u_{i_1},\ldots,u_{i_k}$ are set equal to $u^o_{i_1},\ldots,u^o_{i_k}$, then the resulting function of the arguments $u_{j_1},\ldots,u_{j_m}$, where the index sequence $q' = (j_1,\ldots,j_m)$ is complementary to $q = (i_1,\ldots,i_k)$, is defined to be the membership function of a conditioned restriction $R(X_{j_1},\ldots,X_{j_m}|u^o_{i_1},\ldots,u^o_{i_k})$ or, more simply, $R(X_{(q')}|u^o_{(q)})$.

Thus

$$\mu_{R(X_{j_1},\ldots,X_{j_m}|u^o_{i_1},\ldots,u^o_{i_k})}(u_{j_1},\ldots,u_{j_m}) = \mu_{R(X_1,\ldots,X_n)}(u_1,\ldots,u_n| u_{i_1} = u^o_{i_1},\ldots, u_{i_k} = u^o_{i_k})$$

or more compactly

$$\mu_{R(X_{(q')}|u^o_{(q)})}(u_{(q')}) = \mu_{R(X)}(u|u_{(q)} = u^o_{(q)}) \tag{4.20}$$

The simplicity of the relation between conditioned and unconditioned restrictions becomes more transparent if the u^o_i are written without the superscript. Then, (4.20) becomes

$$\mu_{R(X_{j_1},\ldots,X_{j_m}|u_{i_1},\ldots,u_{i_k})}(u_{j_1},\ldots,u_{j_m}) \triangleq \mu_{R(X_1,\ldots,X_n)}(u_1,\ldots,u_n)$$

or more compactly

$$\mu_{R(X_{(q')}|u_{(q)})}(u_{(q')}) \triangleq \mu_{R(X)}(u) \tag{4.21}$$

Note 4.12 In some instances, it is preferable to use an alternative notation for conditioned restrictions. For example, if $n = 4$, $q = (1,3)$ and $q' = (2,4)$, it may be simpler to write $R(u_1^o, X_2, u_3^o, X_4)$ for $R(X_2, X_4 | u_1^o, u_3^o)$. This is particularly true when numerical values are used in place of the subscripted arguments, e.g., 5 and 2 in place of u_1^o and u_3^o. In such cases, in order to avoid ambuiguity we shall write explicitly $R(X_2, X_4 | u_1^o = 5,\ u_3^o = 2)$ or more simply $R(5, X_2, 2, X_4)$.

<u>Example 4.13</u> In Example 4.10, we have

$$R(X_1, X_2, 0) = 0.8/(0,0) + 0.2/(0,1) + 0.7/(1,1) + 0.9/(1,2) \tag{4.22}$$

$$R(X_1, X_2, 1) = 0.6/(0,0) + 0.4/(0,1) + 0.4/(2,1)$$

$$R(X_1, X_2, 2) = 1/(1,0) + 0.8/(1,1)$$

and, using (4.16)

$$R(X_1, 0) = 0.8/0 + 1/1 \tag{4.23}$$

$$R(X_1, 1) = 0.4/0 + 0.8/1 + 0.4/2$$

$$R(X_1, 2) = 0.9/1$$

It is useful to observe that an immediate consequence of the definitions of marginal and conditioned restrictions is the following

<u>Proposition 4.14</u> Let $R(X_{j_1}, \ldots, X_{j_m})$ be a marginal restriction induced by $R(X_1, \ldots, X_n)$, and let $R(X_{j_1}, \ldots, X_{j_m} | u_{i_1}, \ldots, u_{i_k})$ or, more simply, $R(X_{(q')} | u_{(q)})$ be a restriction conditioned on $u_{i_1}, \ldots, u_{i_k}$, with $q = (i_1, \ldots, i_k)$ and $q' = (j_1, \ldots, j_m)$ being complementary index sequences.

Then, in consequence of (4.16), (4.21) and the definition of the union (see (3.34)), we can assert that

$$R(X_{(q')}) = \sum_{u_{(q)}} R(X_{(q')} | u_{(q)}) \tag{4.24}$$

where $\sum_{u_{(q)}}$ stands for the union (rather than the arithmetic sum) over the $u_{(q)}$.

Example 4.15 With reference to Examples 4.9 and 4.12, it is easy to verify that

$$R(X_1,X_2) = R(X_1,X_2,0) + R(X_1,X_2,1) + R(X_1,X_2,2)$$

and

$$R(X_1) = R(X_1,0) + R(X_1,1) + R(X_1,2)$$

Separability and Noninteraction

Definition 4.16 A n-ary restriction $R(X_1,\ldots,X_n)$ is separable iff it can be expressed as the cartesian product of unary restrictions

$$R(X_1,\ldots,X_n) = R(X_1) \times \ldots \times R(X_n) \tag{4.25}$$

or, equivalently, as the intersection of cylindrical extensions (see (3.62))

$$R(X_1,\ldots,X_n) = \bar{R}(X_1) \cap \ldots \cap \bar{R}(X_n) \tag{4.26}$$

It should be noted that, if $R(X_1,\ldots,X_n)$ is normal, then so are its marginal restrictions (see Proposition 3.14). It follows, then, that the $R(X_i)$ in (4.25) are marginal restrictions induced by $R(X_1,\ldots,X_n)$. For, (4.25) implies that

$$\mu_{R(X_1,\ldots,X_n)}(u_1,\ldots,u_n) = \mu_{R(X_1)}(u_1) \wedge \ldots \wedge \mu_{R(X_n)}(u_n) \tag{4.27}$$

and hence by (3.57)

$$P_i R(X_1,\ldots,X_n) = R(X_i), \quad i = 1,\ldots,n. \tag{4.28}$$

Unless stated to the contrary, we shall assume henceforth that $R(X_1,\ldots,X_n)$ is normal.

Example 4.17 The relation matrix of the restriction shown below can be expressed as the max-min dyadic product of a column vector (a unary relation) and a row vector (a unary relation). This implies that the restriction in question is separable

$$\begin{bmatrix} 0.3 & 0.8 & 0.8 & 0.1 \\ 0.3 & 0.8 & 1 & 0.1 \\ 0.2 & 0.2 & 0.2 & 0.1 \\ 0.3 & 0.6 & 0.6 & 0.1 \end{bmatrix} = \begin{bmatrix} 0.8 \\ 1 \\ 0.2 \\ 0.6 \end{bmatrix} \begin{bmatrix} 0.3 & 0.8 & 1 & 0.1 \end{bmatrix}$$

Example 4.18 The restrictions defined in Examples 4.8 and 4.9 are not separable.

An immediate consequence of separability is the following

Proposition 4.19 If $R(X_1,\ldots,X_n)$ is separable, so is every marginal restriction induced by $R(X_1,\ldots,X_n)$.

Also, in consequence of (4.25), we can assert the

Proposition 4.20 The separable restriction $R(X_1) \times \ldots \times R(X_n)$ is the largest restriction with marginal restrictions $R(X_1),\ldots,R(X_n)$.

The concept of separability is closely related to that of noninteraction of fuzzy variables. More specifically:

Definition 4.21 The fuzzy variables $X_1,\ldots,X_n$ are said to be noninteractive iff the restriction $R(X_1,\ldots,X_n)$ is separable.

It will be recalled that, in the case of nonfuzzy variables, the justification for characterizing $X_1,\ldots,X_n$ as noninteractive is that if (see (2.18))

$$R(X_1,\ldots,X_n) = R(X_1) \times \ldots \times R(X_n) \tag{4.29}$$

then the n-ary assignment equation

$$(x_1,\ldots,x_n) = (u_1,\ldots,u_n): \; R(X_1,\ldots,X_n) \tag{4.30}$$

can be decomposed into a sequence of n unary assignment equations

$$\begin{array}{l} x_1 = u_1: \; R(X_1) \\ \text{--------------} \\ x_n = u_n: \; R(X_n) \end{array} \tag{4.31}$$

In the case of fuzzy variables, a basic consequence of noninteraction - from which (2.19) follows as a special case - is expressed by the following

Proposition 4.22 If the fuzzy variables $X_1,\ldots,X_n$ are noninteractive, then the n-ary assignment equation (4.30) can be decomposed into a sequence of n unary assignment equations (4.31), with the understanding that if $c(u_1,\ldots,u_n)$ is the compatibility of $(u_1,\ldots,u_n)$ with $R(X_1,\ldots,X_n)$, and $c_i(u_i)$, $i = 1,\ldots,n$, is the compatibility of u_i with $R(X_i)$, then

$$c(u_1,\ldots,u_n) = c_1(u_1) \wedge \ldots \wedge c_n(u_n) \tag{4.32}$$

Proof. By the definitions of compatibility, noninteraction and separability, we have at once

$$c(u_1,\ldots,u_m) = \mu_{R(X_1,\ldots,X_n)}(u_1,\ldots,u_n) \tag{4.33}$$

$$= \mu_{R(X_1)}(u_1) \wedge \cdots \wedge \mu_{R(X_n)}(u_n)$$

$$= c_1(u_1) \wedge \cdots \wedge c_n(u_n) \qquad \text{Q.E.D.}$$

Comment 4.23 Pursuing the valise analogy further (see Comment 4.6), noninteractive fuzzy variables $X_1,\ldots,X_n$ may be likened to n separate soft valises with name-tags $X_1,\ldots,X_n$. The restriction associated with valise X_i is characterized by the compatibility function $c(u_i)$. Then, the overall compatibility function for the valises $X_1,\ldots,X_n$ is given by (4.32). (Fig. 4.4.)

Comment 4.24 In terms of the base variables of $X_1,\ldots,X_n$ (see Definition 4.1), noninteraction implies that there are no constraints which jointly involve $u_1,\ldots,u_n$, where u_i is the base variable for X_i, $i = 1,\ldots,n$. For example, if the u_i are constrained by

$$u_1 + \ldots + u_n = 1$$

then $X_1,\ldots,X_n$ are interactive, i.e., are not noninteractive. (See Comment 3.20.)

If $X_1,\ldots,X_n$ are interactive, it is still possible to decompose an n-ary assignment equation into a sequence of n unary assignment equations. However, the restriction on u_i will, in general, depend on the values assigned to $u_1,\ldots,u_{i-1}$. Thus, the n assignment equations will have the following form (see also (2.21))

$$x_1 = u_1: \quad R(X_1) \tag{4.34}$$

$$x_2 = u_2: \quad R(X_2 | u_1)$$

$$x_3 = u_3: \quad R(X_3 | u_1, u_2)$$

$$x_n = u_n: \quad R(X_n | u_1, \ldots, u_{n-1})$$

where $R(X_i | u_1, \ldots, u_{i-1})$ denotes the restriction on u_i conditioned on $u_1, \ldots, u_{i-1}$ (see Definition 4.11).

Example 4.25 Taking Example 4.10, assume that $u_1 = 1$, $u_2 = 2$ and $u_3 = 0$. Then

$$R(X_1) = 0.8/0 + 1/1 + 0.4/2 \tag{4.35}$$

$$R(X_2 | u_1 = 1) = 1/0 + 0.8/1 + 0.9/2$$

and

$$R(X_3 | u_1 = 1, u_2 = 2) = 0.9/0$$

so that

$$c_1(1) = 1 \tag{4.36}$$

$$c_2(2) = 0.9$$

and

$$c_3(0) = 0.9$$

As in the case of (4.31), the justification for (4.34) is provided by the following

Proposition 4.26 If $X_1, \ldots, X_n$ are interactive fuzzy variables subject to the restriction $R(X_1,\ldots,X_n)$, and $c_i(u_i)$, $i = 1,\ldots,n$, is the compatibility of u_i with the conditioned restriction $R(X_i|u_1,\ldots,u_{i-1})$ in (4.34), then

$$c(u_1,\ldots,u_n) = c_1(u_1) \wedge \ldots \wedge c_n(u_n) \tag{4.37}$$

where $c(u_1,\ldots,u_n)$ is the compatibility of $(u_1,\ldots,u_n)$ with $R(X_1,\ldots,X_n)$.

Proof By the definition of a conditioned restriction (see (4.20)), we have, for all i, $1 \leq i \leq n$

$$\mu_{R(X_i|u_1,\ldots,u_{i-1})}(u_i) = \mu_{R(X_1,\ldots,X_i)}(u_1,\ldots,u_i) \tag{4.38}$$

On the other hand, the definition of a marginal restriction (see (4.16)) implies that, for all i and all $u_1,\ldots,u_i$, we have

$$\mu_{R(X_1,\ldots,X_i)}(u_1,\ldots,u_i) \geq \mu_{R(X_1,\ldots,X_{i+1})}(u_1,\ldots,u_{i+1}) \tag{4.39}$$

and hence that

$$\mu_{R(X_{i+1}|u_1,\ldots,u_i)}(u_{i+1}) \wedge \mu_{R(X_i|u_i,\ldots,u_{i-1})}(u_i) = \mu_{R(X_{i+1}|u_1,\ldots,u_i)}(u_{i+1}) \tag{4.40}$$

Combining (4.40) with the defining equation

$$c_i(u_i) = \mu_{R(X_i|u_1,\ldots,u_{i-1})}(u_i) \tag{4.41}$$

we derive

$$c(u_1,\ldots,u_n) = c_1(u_1) \wedge \ldots \wedge c_n(u_n) \qquad \text{Q.E.D.} \tag{4.42}$$

This concludes our discussion of some of the properties of fuzzy variables which are relevant to the concept of a linguistic variable. In the following section, we shall formalize the concept of a linguistic variable and explore some of its implications.

5. The Concept of a Linguistic Variable

In our informal discussion of the concept of a linguistic variable in Sec. 1, we have stated that a linguistic variable differs from a numerical variable in that its values are not numbers but words or sentences in a natural or artificial language. Since words, in general, are less precise than numbers, the concept of a linguistic variable serves the purpose of providing a means of approximate characterization of phenomena which are too complex or too ill-defined to be amenable to description in conventional quantitative terms. More specifically, the fuzzy sets which represent the restrictions associated with the values of a linguistic variable may be viewed as summaries of various subclasses of elements in a universe of discourse. This, of course, is analogous to the role played by words and sentences in a natural language. For example, the adjective <u>handsome</u> is a summary of a complex of characteristics of the appearance of an individual. It may also be viewed as a label for a fuzzy set which represents a restriction imposed by a fuzzy variable named <u>handsome</u>. From this point of view, then, the terms <u>very handsome</u>, <u>not handsome</u>, <u>extremely handsome</u>, <u>quite handsome</u>, etc., are names of fuzzy sets which result from operating on the fuzzy set named <u>handsome</u> with the modifiers named <u>very</u>, <u>not</u>, <u>extremely</u>, <u>quite</u>, etc. In effect, these fuzzy sets, together with the fuzzy set labeled <u>handsome</u>, play the role of values of the linguistic variable <u>Appearance</u>.

An important facet of the concept of a linguistic variable is that it is a variable of a higher order than a fuzzy variable, in the sense that a linguistic variable takes fuzzy variables as its values. For example, the values of a linguistic variable named <u>Age</u> might be: <u>young</u>, not young, old, very old, not young and not old, quite old, etc., each of which is the name of a fuzzy variable. If X is the name of such a fuzzy variable, the restriction imposed by X may be interpreted as the meaning of X. Thus, if the restriction imposed by the fuzzy variable named old is a fuzzy subset of U = [0,100] defined by

$$R(\underline{old}) = \int_{50}^{100} \left(1 + \left(\frac{u-50}{5}\right)^{-2}\right)^{-1} / u \quad , \quad u \in U \tag{5.1}$$

then the fuzzy set represented by R(old) may be taken to be the meaning of old. (Fig. 5.1.)

Another important facet of the concept of a linguistic variable is that, in general, a linguistic variable is associated with two rules: (1) a syntactic rule, which may have the form of a grammar for generating the names of the values of the variable; and (2) a semantic rule which defines an algorithmic procedure for computing the meaning of each value. These rules constitute an essential part of the characterization of a structured linguistic variable.[1]

Since a linguistic variable is a variable of a higher order than a fuzzy variable, its characterization is necessarily more complex than that expressed by Definition 4.1. More specifically, we have

Definition 5.1 A linguistic variable is characterized by a quintuple $(\mathcal{X}, T(\mathcal{X}), U, G, M)$ in which $\mathcal{X}$ is the name of the variable; $T(\mathcal{X})$ (or simply T)

[1]It is primarily the semantic rule that distinguishes a linguistic variable from the more conventional concept of a syntactic variable.

denotes the term-set of $\mathcal{X}$, that is, the set of names of linguistic values of $\mathcal{X}$, with each value being a fuzzy variable denoted generically by X and ranging over a universe of discourse U which is associated with the base variable u; G is a syntactic rule (which usually has the form of a grammar) for generating the names, X, of values of $\mathcal{X}$; and M is a semantic rule for associating with each X its meaning, M(X), which is a fuzzy subset of U. A particular X, that is, a name generated by G, is called a term. A term consisting of a word or words which function as a unit (i.e., always occur together) is called an atomic term. A term which contains one or more atomic terms is a composite term. A concatenation of components of a composite term is a subterm. If X_1, X_2, ... are terms in T, then T may be expressed as the union

$$T = X_1 + X_2 + \dots \tag{5.2}$$

Where necessary to place in evidence that T is generated by a grammar G, T will be written as T(G).

The meaning, M(X), of a term X is defined to be the restriction, R(X), on the base variable u which is imposed by the fuzzy variable named X. Thus

$$M(X) \triangleq R(X) \tag{5.3}$$

with the understanding that R(X) -- and hence M(X) -- may be viewed as a fuzzy subset of U carrying the name X. The connection between $\mathcal{X}$, the linguistic value X and the base variable u is illustrated in Fig. 1.3.

Note 5.2 In order to avoid a profusion of symbols, it is expedient to assign more than one meaning to some of the symbols occurring in Definition 5.1, relying on the context for disambiguation. Specifically:

a) We shall frequently employ the symbol $\mathcal{X}$ to denote both the name of the variable and the generic name of its values. Likewise, X will be used to denote both the generic name of the values of the variable and the name of the variable itself.

b) The same symbol will be used to denote a set and the name of that set. Thus, the symbols X, M(X) and R(X) will be used interchangeably, although strictly speaking X- as the name of M(X) (or R(X))- is distinct from M(X). In other words, when we say that a term X (e.g. young) is a value of $\mathcal{X}$ (e.g., Age), it should be understood that the actual value is M(X) and that X is merely the name of the value.

Example 5.3 Consider a linguistic variable named Age, i.e., $\mathcal{X}$ = Age, with U = [0,100]. A linguistic value of Age might be named old, with old being an atomic term. Another value might be named very old, in which case very old is a composite term which contains old as an atomic component and has very and old as subterms. The value of Age named more or less young is a composite term which contains young as an atomic term and in which more or less is a subterm. The term-set associated with Age may be expressed as

$$T(\underline{Age}) = \underline{old} + \underline{very}\ \underline{old} + \underline{not}\ \underline{old} + \underline{more}\ \underline{or}\ \underline{less}\ \underline{young} + \underline{quite}\ \underline{young} + \underline{not}\ \underline{very}\ \underline{old}\ \underline{and}\ \underline{not}\ \underline{very}\ \underline{young} + \ldots \tag{5.4}$$

in which each term is the name of a fuzzy variable in the universe of discourse U = [0,100]. The restriction imposed by a term, say R(old), constitutes the meaning of old. Thus, if R(old) is defined by (5.1), then the meaning of the linguistic value <u>old</u> is given by

$$M(\underline{old}) = \int_{50}^{100} \left(1 + \left(\frac{u-50}{5}\right)^{-2}\right)^{-1} / u \tag{5.5}$$

or more simply (see Note 5.2)

$$\underline{old} = \int_{50}^{100} \left(1 + \left(\frac{u-50}{5}\right)^{-2}\right)^{-1} / u \tag{5.6}$$

Similarly, the meaning of a linguistic value such as <u>very</u> <u>old</u> may be expressed as (see Fig. 5.1)

$$M(\underline{very}\ \underline{old}) = \underline{very}\ \underline{old} = \int_{50}^{100} \left(1 + \left(\frac{u-50}{5}\right)^{-2}\right)^{-2} / u \tag{5.7}$$

The assignment equation in the case of a linguistic variable assumes the form

$$\begin{aligned} X &= \text{term in } T(\mathcal{X}) \\ &= \text{name generated by } G \end{aligned} \tag{5.8}$$

which implies that the meaning assigned to X is expressed by

$$M(X) = R(\text{term in } T(\mathcal{X})) \tag{5.9}$$

In other words, the meaning of X is given by the application of the semantic rule M to the value assigned to X by the right-hand member of (5.8). Furthermore, as defined by (5.3), M(X) is identical to the restriction imposed by X.

Comment 5.4 In accordance with Note 5.2a, the assignment equation will usually be written as

$$\mathcal{X} = \text{name in } T(\mathcal{X}) \tag{5.10}$$

rather than in the form (5.8). For example, if $\mathcal{X}$ = Age, and old is a term in $T(\mathcal{X})$, we shall write

$$\underline{\text{Age}} = \underline{\text{old}} \tag{5.11}$$

with the understanding that old is a restriction on the values of u defined by (5.1), which is assigned by (5.11) to the linguistic variable named Age. It is important to note that the equality symbol in (5.10) does not represent a symmetric relation -- as it does in the case of arithmetic equality. Thus, it would not be meaningful to write (5.11) as

$$\underline{\text{old}} = \underline{\text{Age}}$$

To illustrate the concept of a linguistic variable, we shall consider first a very elementary example in which $T(\mathcal{X})$ contains just a few terms and the syntactic and semantic rules are trivially simple.

Example 5.5. Consider a linguistic variable named Number which is associated with the finite term-set

$$T(\underline{\text{Number}}) = \underline{\text{few}} + \underline{\text{several}} + \underline{\text{many}} \tag{5.12}$$

in which each term represents a restriction on the values of u in the universe of discourse

$$U = 1 + 2 + 3 + \ldots + 10 \tag{5.13}$$

These restrictions are assumed to be fuzzy subsets of U which are defined as follows

$$\underline{few} = 0.4/1 + 0.8/2 + 1/3 + 0.4/4 \tag{5.14}$$

$$\underline{several} = 0.5/3 + 0.8/4 + 1/5 + 1/6 + 0.8/7 + 0.5/8 \tag{5.15}$$

$$\underline{many} = 0.4/6 + 0.6/7 + 9.8/8 + 0.9/9 + 1/10 \tag{5.16}$$

Thus

$$R(\underline{few}) = M(\underline{few}) = 0.4/1 + 0.8/2 + 1/3 + 0.4/4 \tag{5.17}$$

and likewise for the other terms in T. The implication of (5.17) is that few is the name of a fuzzy variable which is a value of the linguistic variable Number. The meaning of few - which is the same as the restriction on few -is a fuzzy subset of U which is defined by the right-hand number of (5.17).

To assign a value such as few to the linguistic variable Number, we write

$$\underline{Number} = \underline{few} \tag{5.18}$$

with the understanding that what we actually assign to Number is a fuzzy variable named few.

Example 5.6. In this case, we assume that we are dealing with a composite linguistic variable[2] named $(\mathcal{X},\mathcal{Y})$ which is associated with the base variable (u,v) ranging over the universe of discourse U x V, where

$$U \times V = (1 + 2 + 3 + 4) \times (1 + 2 + 3 + 4) \tag{5.19}$$

[2]Composite linguistic variables will be discussed in greater detail in Sec. 6 in connection with linguistic truth variables.

$$= (1,1) + (1,2) + (1,3) + (1,4) + \qquad (5.20)$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$+ (4,1) + (4,2) + (4,3) + (4,4)$$

with the understanding that

$$i \times j = (i,j) \quad , \quad i,j = 1,2,3,4. \qquad (5.21)$$

Furthermore, we assume that the term-set of $(\mathcal{X},\mathcal{Y})$ comprises just two terms:

$$T = \underline{\text{approximately}}\ \underline{\text{equal}} + \underline{\text{more}}\ \underline{\text{or}}\ \underline{\text{less}}\ \underline{\text{equal}} \qquad (5.22)$$

where <u>approximately</u> <u>equal</u> and <u>more</u> <u>or</u> <u>less</u> <u>equal</u> are names of binary fuzzy relations defined by the relation matrices

$$\underline{\text{approximately}}\ \underline{\text{equal}} = \begin{bmatrix} 1 & 0.6 & 0.4 & 0.2 \\ 0.6 & 1 & 0.6 & 0.4 \\ 0.4 & 0.6 & 1 & 0.6 \\ 0.2 & 0.4 & 0.6 & 1 \end{bmatrix} \qquad (5.23)$$

and

$$\underline{\text{more}}\ \underline{\text{or}}\ \underline{\text{less}}\ \underline{\text{equal}} = \begin{bmatrix} 1 & 0.8 & 0.6 & 0.4 \\ 0.8 & 1 & 0.8 & 0.6 \\ 0.6 & 0.8 & 1 & 0.8 \\ 0.4 & 0.6 & 0.8 & 1 \end{bmatrix} \qquad (5.24)$$

In these relation matrices, the (i,j)th entry represents the compatibility of the pair (i,j) with the restriction in question. For example, the

(2,3) entry in approximately equal - which is 0.6 - is the compatibility of the ordered pair (2,3) with the binary restriction named approximately equal.

To assign a value, say approximately equal, to $(\mathcal{X},\mathcal{Y})$, we write

$$(\mathcal{X},\mathcal{Y}) = \underline{\text{approximately}}\ \underline{\text{equal}} \tag{5.25}$$

where, as in (5.18), it is understood that what we assign to $(\mathcal{X},\mathcal{Y})$ is a binary fuzzy relation named approximately equal, which is a binary restriction on the values of (u,v) in the universe of discourse (5.20).

Comment 5.7 In terms of the valise analogy (see Comment 4.3), a linguistic variable as defined by Definition 5.1 may be likened to a hard valise into which we can put soft valises, as illustrated in Fig. 5.2. A soft valise corresponds to a fuzzy variable which is assigned as a linguistic value to $\mathcal{X}$, with X playing the role of the name-tag of the soft valise.

Structured Linguistic Variables

In both of the above examples the term-set contains only a small number of terms, so that it is practicable to list the elements of $T(\mathcal{X})$ and set up a direct association between each element and its meaning. In the more general case, however, the number of elements in $T(\mathcal{X})$ may be infinite, necessitating the use of an algorithm, rather than a table look-up procedure, for generating the elements of $T(\mathcal{X})$ as well as for computing their meaning.

A linguistic variable $\mathcal{X}$ will be said to be structured if its term-set, $T(\mathcal{X})$, and the function, M, which associates a meaning with each term in the term-set, can be characterized algorithmically. In this sense, the syntactic and semantic rules associated with a structured linguistic variable may be viewed as algorithmic procedures for generating the elements of $T(\mathcal{X})$ and computing the meaning of each term in $T(\mathcal{X})$, respectively. Unless stated to the contrary, we shall assume henceforth that the linguistic variables we deal with are structured.

<u>Example 5.8</u>. As a very simple illustration of the role played by the syntactic and semantic rules in the case of a structured linguistic variable, we shall consider a variable named <u>Age</u> whose terms are exemplified by: <u>old</u>, <u>very</u> <u>old</u>, <u>very</u> <u>very</u> <u>old</u>, <u>very</u> <u>very</u> <u>very</u> <u>old</u>, etc. Thus, the term set of <u>Age</u> can be written as

$$T(\underline{Age}) = \underline{old} + \underline{very}\ \underline{old} + \underline{very}\ \underline{very}\ \underline{old} + \dots \tag{5.26}$$

In this simple case, it is clear by inspection that every term in T(<u>Age</u>) is of the form <u>old</u> or <u>very</u> <u>very</u> ... <u>very</u> <u>old</u>. To deduce this rule in a more general way, we proceed as follows.

Let xy denote the concatenation of character strings x and y, e.g., x = <u>very</u>, y = <u>old</u>, xy = <u>very</u> <u>old</u>. If A and B are sets of strings, e.g.,

$$A = x_1 + x_2 + \dots \tag{5.27}$$

$$B = y_1 + y_2 + \dots \tag{5.28}$$

where x_i and y_j are character strings, then the concatenation of A and B is denoted by AB and is defined as the set of strings

$$AB = (x_1 + x_2 + \ldots)\ (y_1 + y_2 + \ldots) \tag{5.29}$$

$$= \sum_{i,j} x_i y_j$$

For example, if A = very and B = old + very old, then

$$\underline{\text{very}}\ (\underline{\text{old}} + \underline{\text{very}}\ \underline{\text{old}}) = \underline{\text{very}}\ \underline{\text{old}} + \underline{\text{very}}\ \underline{\text{very}}\ \underline{\text{old}} \tag{5.30}$$

Using this notation, the given expression for T(Age), or simply T, may be taken to be the solution of the equation[3]

$$T = \underline{\text{old}} + \underline{\text{very}}\ T \tag{5.31}$$

which, in words, means that every term in T is of the form old or very followed by some term in T.

Equation (5.31) can be solved by iteration, using the recursion equation

$$T^{i+1} = \underline{\text{old}} + \underline{\text{very}}\ T^i \quad , \quad i = 0,1,2,\ldots \tag{5.32}$$

with the initial value of T^i being the empty set θ. Thus

$$\begin{aligned} T^0 &= \theta \\ T^1 &= \underline{\text{old}} \\ T^2 &= \underline{\text{old}} + \underline{\text{very}}\ \underline{\text{old}} \\ T^3 &= \underline{\text{old}} + \underline{\text{very}}\ \underline{\text{old}} + \underline{\text{very}}\ \underline{\text{very}}\ \underline{\text{old}} \\ &- - - - - - - - - - - - - - - - - - \end{aligned} \tag{5.33}$$

[3] As is well-known in the theory of regular expressions (see [32]), the solution of (5.31) can be expressed as

$$T = (\lambda + \underline{\text{very}} + \underline{\text{very}}^2 + \ldots)\ \underline{\text{old}}$$

where λ is the null string. This expression for T is equivalent to that of (5.34).

and the solution of (5.31) is given by

$$T = T^{\infty} = \underline{old} + \underline{very}\ \underline{old} + \underline{very}\ \underline{very}\ \underline{old} + \underline{very}\ \underline{very}\ \underline{very}\ \underline{old} + \ldots \tag{5.34}$$

For the example under consideration, the syntactic rule, then, is expressed by (5.31) and its solution (5.34). Equivalently, the syntactic rule can be characterized by the production system

$$T \rightarrow \underline{old} \tag{5.35}$$

$$T \rightarrow \underline{very}\ T \tag{5.36}$$

for which (5.31) plays the role of an algebraic representation.[4] In this case, a term in T can be generated through a standard derivation procedure ([36],]37]) involving a successive application of the rewriting rules (5.35) and (5.36) starting with the symbol T. Thus, if T is rewritten as very T and then T in very T is rewritten as old, we obtain the term very old. In a similar fashion, the term very very very old can be obtained from T by the derivation chain

$$T \rightarrow \underline{very}\ T \rightarrow \underline{very}\ \underline{very}\ T \rightarrow \underline{very}\ \underline{very}\ \underline{very}\ T \rightarrow \underline{very}\ \underline{very}\ \underline{very}\ \underline{old} \tag{5.37}$$

Turning to the semantic rule for Age, we note that to compute the meaning of a term such as very ... very old we need to know the meaning of old and the meaning of very. The term old plays the role of a primary term, that is, a term whose meaning must be specified as an initial datum in order to provide a basis for the computation of the meaning of composite terms in T. As for the term very, it acts as a

[4] A discussion of the algebraic representation of context-free grammars may be found in [33], [34] and [35]. Algebraic treatment of fuzzy languages is discussed in [6] and [58].

linguistic hedge, that is, as a modifier of the meaning of its operand. If - as very simple approximation - we assume that very acts as a concentrator (see(3.40)), then

$$\underline{\text{very}}\ \underline{\text{old}} = \text{CON}(\underline{\text{old}}) = \underline{\text{old}}^2 \qquad (5.38)$$

Consequently, the semantic rule for Age may be expressed as

$$M(\underline{\text{very}}\ \ldots\ \underline{\text{very}}\ \underline{\text{old}}) = \underline{\text{old}}^{2n} \qquad (5.39)$$

where n is the number of occurrences of very in the term very...very old and M(very...very old) is the meaning of very...very old. Furthermore, if the primary term old is defined as

$$\underline{\text{old}} = \int_{50}^{100} \left(1 + \left(\frac{u-50}{5}\right)^{-2}\right)^{-1} /u \qquad (5.40)$$

then

$$M(\underline{\text{very}}\ \ldots\ \underline{\text{very}}\ \underline{\text{old}}) = \int_{50}^{100} \left(1 + \left(\frac{u-50}{5}\right)^{-2}\right)^{-2n} /u \quad , \quad n = 1,2,\ldots \qquad (5.41)$$

This equation provides an explicit semantic rule for the computation of the meaning of composite terms generated by (5.31), from the knowledge of the meaning of the primary term old and the hedge very.

Boolean Linguistic Variables

The linguistic variable considered in Example 5.8 is a special case of what might be called a Boolean linguistic variable. Typically, such a variable involves a finite number of primary terms, a finite number of hedges, the connectives _and_ and _or_, and the negation _not_. For example, the term-set of a Boolean linguistic variable _Age_ might be:

$$\underline{Age} = \underline{young} + \underline{old} + \underline{not}\ \underline{young} + \underline{not}\ \underline{old} + \underline{very}\ \underline{young} + \underline{very}\ \underline{very}\ \underline{young} \tag{5.42}$$

$$+ \underline{not}\ \underline{very}\ \underline{young}\ \underline{and}\ \underline{not}\ \underline{very}\ \underline{old} + \underline{quite}\ \underline{young} + \underline{more}\ \underline{or}\ \underline{less}\ \underline{old}$$

$$+ \underline{extremely}\ \underline{old} + \ldots.$$

More formally, a Boolean linguistic variable may be defined recursively as follows.

Definition 5.9. A _Boolean linguistic variable_ is a linguistic variable whose terms, X, are Boolean expressions in variables of the form X_p, hX_p, X or hX, where h is a linguistic hedge, X_p is a primary term and hX is the name of a fuzzy set resulting from acting with h on X.

As an illustration, in the case of the linguistic variable _Age_ whose term-set is defined by (5.42), the term _not very young and not very old_ is of the form (5.9) with $h \triangleq$ _very_, $X \triangleq$ young and $X \triangleq$ old. Similarly, in the case of the term _very very young_, $h \triangleq$ _very very_ and $X \triangleq$ _young_.

Boolean linguistic variables are particularly convenient to deal with because much of our experience in the manipulation and evaluation of Boolean expressions is transferable to variables of this type. To illustrate this point, we shall consider a simple example which involves two primary terms and a single hedge.

Example 5.10. Let _Age_ be a Boolean linguistic variable with the term-set

$$T(\underline{Age}) = \underline{young} + \underline{not}\ \underline{young} + \underline{old} + \underline{not}\ \underline{old} + \underline{very}\ \underline{young}$$
$$+ \underline{not}\ \underline{young}\ \underline{and}\ \underline{not}\ \underline{old} + \underline{young}\ \underline{or}\ \underline{old} + \underline{young}\ \underline{or}\ (\underline{not} \tag{5.43}$$
$$\underline{very}\ \underline{young}\ \underline{and}\ \underline{not}\ \underline{very}\ \underline{old}) + \ldots$$

If we identify and with intersection, or with union, not with complementation and very with concentration (see (5.40)), the meaning of a typical value of Age can be written down by inspection. For example

$$\begin{aligned}
M(\underline{not}\ \underline{young}) &= \neg\,\underline{young} \\
M(\underline{not}\ \underline{very}\ \underline{young}) &= \neg\,(\underline{young}^2) \\
M(\underline{not}\ \underline{very}\ \underline{young}\ \underline{and}\ \underline{not}\ \underline{very}\ \underline{old}) &= \neg\,(\underline{young}^2) \cap \neg\,(\underline{old}^2) \\
M(\underline{young}\ \underline{or}\ \underline{old}) &= \underline{young} \cup \underline{old}
\end{aligned} \tag{5.44}$$

In effect, these equations express the meaning of a composite term as a function of the meaning of its constituent primary terms. Thus, if young and old are defined as

$$\underline{young} = \int_0^{25} 1/u + \int_{25}^{100} \left(1 - \left(\frac{u-25}{5}\right)^2\right)^{-1} /u \tag{5.45}$$

$$\underline{old} = \int_{50}^{100} \left(1 - \left(\frac{u-50}{5}\right)^{-2}\right)^{-1} /u \tag{5.46}$$

then (see Fig. 5.3)

$$M(\underline{young}\ \underline{or}\ \underline{old}) = \int_0^{25} 1/u + \int_{25}^{50} \left(1 + \left(\frac{u-25}{5}\right)^2\right)^{-1} /u \tag{5.47}$$
$$+ \int_{50}^{100} \left(1 + \left(\frac{u-25}{5}\right)^2\right)^{-1} \vee \left(1 + \left(\frac{u-50}{5}\right)^{-2}\right)^{-1} /u$$

The linguistic variable considered in the above example involves just one type of hedge, namely, very. More generally, a Boolean linguistic variable may involve a finite number of hedges, as in (5.42). The procedure for computing the meaning of a composite term remains the same, however, once the operations corresponding to the hedges are defined.

The question of what constitutes an appropriate representation for a particular hedge, e.g., more or less or quite or essentially is by no means a simple one.[5] To illustrate the point, in some contexts the effect of the hedge more or less may be approximated by (see (3.41))

$$M(\underline{\text{more}}\ \underline{\text{or}}\ \underline{\text{less}}\ X) = DIL(X) = X^{0.5} \tag{5.48}$$

For example, if X = old, and old is defined by (5.46), then

$$\underline{\text{more}}\ \underline{\text{or}}\ \underline{\text{less}}\ \underline{\text{old}} = \int_{50}^{100} \left(1 + \left(\frac{u-50}{5}\right)^{-2}\right)^{-0.5} \Big/ u \tag{5.49}$$

In many instances, however, more or less acts as a fuzzifier in the sense of (3.48), rather than as a dilator. As an illustration, suppose that the meaning of a primary term recent is specified as

$$\underline{\text{recent}} = 1/1974 + 0.8/1973 + 0.7/1972 \tag{5.50}$$

and that more or less recent is defined as the result of acting with a fuzzifier F on recent, i.e.,

---

[5]A more detailed discussion of linguistic hedges from a fuzzy-set-theoretic point of view may be found in [27] and [38]. The idea of treating various types of linguistic hedges as operators on fuzzy sets originated in the course of the author's collaboration with Professor G. Lakoff.

$$\underline{\text{more}}\ \underline{\text{or}}\ \underline{\text{less}}\ \underline{\text{recent}} = F(\underline{\text{recent}}; K) \tag{5.51}$$

where the kernel K of F is defined by

$$\begin{aligned} K(1974) &= 1/1974 + 0.9/1973 \\ K(1973) &= 1/1973 + 0.9/1972 \\ K(1972) &= 1/1972 + 0.8/1971 \end{aligned} \tag{5.52}$$

On substituting the values of K into (3.48), we obtain the meaning of <u>more</u> <u>or</u> <u>less</u> <u>recent</u>, i.e.,

$$\underline{\text{more}}\ \underline{\text{or}}\ \underline{\text{less}}\ \underline{\text{recent}} = 1/1974 + 0.9/1973 + 0.72/1972 + 0.56/1971 \tag{5.53}$$

On the other hand, if the hedge <u>more</u> <u>or</u> <u>less</u> were assumed to be a dilator, then we would have

$$\begin{aligned} \underline{\text{more}}\ \underline{\text{or}}\ \underline{\text{less}}\ \underline{\text{recent}} &= (1/1974 + 0.8/1973 + 0.7/1972)^{0.5} \\ &= 1/1974 + 0.9/1973 + 0.84/1972 \end{aligned} \tag{5.54}$$

which differs from (5.53) mainly in the absence of the term 0.56/1971. Thus, if this term were of importance in the definition of <u>more</u> <u>or</u> <u>less</u> <u>recent</u>, then the approximation to <u>more</u> <u>or</u> <u>less</u> by a dilator would not be a good one.

In Example 5.10, we have deduced the semantic rule by inspection, talking advantage of our familiarity with the evaluation of Boolean expressions. To illustrate a more general technique, we shall consider the same linguistic variable as in Example 5.10, but use a method [39] which is an adaptation of the approach employed by Knuth in [40] to define the semantics of context-free languages.

Example 5.11. It can readily be verified that the term-set of Example 5.10 is generated by a context-free grammar $G = (V_T, V_N, T, P)$ in which the non-terminals (syntactic categories) are denoted by T,A,B,C,D, i.e.,

$$V_N = T + A + B + C + D + E \tag{5.55}$$

while the set of terminals (components of terms in T) is expressed by

$$V_T = \underline{\text{young}} + \underline{\text{old}} + \underline{\text{very}} + \underline{\text{not}} + \underline{\text{and}} + \underline{\text{or}} + (\ +\ ) \tag{5.56}$$

and the production system, P, is given by

$$\begin{array}{ll} T \rightarrow A & C \rightarrow D \\ T \rightarrow T\ \underline{\text{or}}\ A & C \rightarrow E \\ A \rightarrow B & D \rightarrow \text{very}\ D \\ A \rightarrow A\ \underline{\text{and}}\ B & E \rightarrow \text{very}\ E \\ B \rightarrow C & D \rightarrow \underline{\text{young}} \\ B \rightarrow \text{not}\ C & E \rightarrow \underline{\text{old}} \\ C \rightarrow (T) & \end{array} \tag{5.57}$$

The production system, P, can also be represented in an algebraic form as the set of equations (see Footnote 3)

$$\begin{aligned} T &= A + T\ \underline{\text{or}}\ A \\ A &= B + A\ \underline{\text{and}}\ B \\ B &= C + \underline{\text{not}}\ C \\ C &= (T) + D + E \\ D &= \underline{\text{very}}\ D + \text{young} \\ E &= \underline{\text{very}}\ E + \underline{\text{old}} \end{aligned} \tag{5.58}$$

The solution of this set of equations for T yields the term set T as expressed by (5.43). Similarly, the solutions for A,B,C,D, and E yield sets of terms which constitute the syntactic categories denoted by A,B,C,D, and E, respectively. The solution of (5.58) can be obtained iteratively, as in (5.32), by using the recursion equation

$$(T,A,B,C,D,E)^{i+1} = f((T,A,B,C,D,E)^{i}) \quad , \quad i=0,1,2,\ldots \tag{5.59}$$

with

$$(T,A,B,C,D,E)^{o} = (\theta,\ldots,\theta)$$

where (T,A,B,C,D,E) is a 6-tuple whose components are the nonterminals in (5.58); f is the mapping defined by the system of equations (5.58); θ is the empty set; and $(T,A,B,C,D,E)^{i}$ is the ith iterate of (T,A,B,C,D,E). The solution of (5.58), which is the fixed point of f, is given by $(T,A,B,C,D,E)^{\infty}$. However, it is true for all i that

$$(T,A,B,C,D,E)^{i} \subset (T,A,B,C,D,E) \tag{5.60}$$

which means that every component in the 6-tuple on the left of (5.60) is a subset of the corresponding component on the right of (5.60). The implication of (5.60), then, is that we generate more and more terms in each of the syntactic categories T,A,B,C,D,E as we iterate (5.59) on i.

In a more conventional fashion, a term in T, say <u>not very young and not very old</u>, is generated by G through a succession of substitutions (derivations) involving the productions in P, with each derivation chain starting with T and terminating on a term generated by G. For example, the derivation chain for the term <u>not very young and not very old</u> is

(see also Example 5.8)

$$T \rightarrow A \rightarrow A \underline{\text{and}} B \rightarrow B \underline{\text{and}} B \rightarrow \underline{\text{not}} C \underline{\text{and}} B \rightarrow \underline{\text{not}} D \underline{\text{and}} B \rightarrow \underline{\text{not}}\ \underline{\text{very}} D \underline{\text{and}} B \rightarrow \underline{\text{not}}\ \underline{\text{very}}\ \underline{\text{young}}\ \underline{\text{and}} B \rightarrow \underline{\text{not}}\ \underline{\text{very}}\ \underline{\text{young}}\ \underline{\text{and}}\ \underline{\text{not}} C \rightarrow \underline{\text{not}}\ \underline{\text{very}}\ \underline{\text{young}}\ \underline{\text{and}}\ \underline{\text{not}} E \rightarrow \underline{\text{not}}\ \underline{\text{very}}\ \underline{\text{young}}\ \underline{\text{and}}\ \underline{\text{not}}\ \underline{\text{very}} E \rightarrow \underline{\text{not}}\ \underline{\text{very}}\ \underline{\text{young}}\ \underline{\text{and}}\ \underline{\text{not}}\ \underline{\text{very}}\ \underline{\text{old}} \tag{5.61}$$

This derivation chain can be deduced from the syntax (parse) tree shown in Fig. 5.4, which exhibits the phrase structure of the term <u>not</u> <u>very</u> <u>young</u> <u>and</u> <u>not</u> <u>very</u> <u>old</u> in terms of the syntactic categories T,A,B,C,D,E. In effect,this procedure for generating the terms in T by the use of the grammar G constitutes the syntactic rule for the variable <u>Age</u>.

The semantic rule for <u>Age</u> is <u>induced</u> by the syntactic rule described above in the sense that the meaning of a term in T is determined, in part, by its syntax tree. Specifically, each production in (5.57) is associated with a relation between the fuzzy sets labeled by the corresponding terminal and nonterminal symbols. The resulting dual system of productions and associated equations has the appearance shown below, with the subscripts L and R serving to differentiate between the symbols on the left-and right-hand sides of a production ($+ \triangleq$ union)

$$T \rightarrow A \qquad \Rightarrow T_L = A_R \tag{5.62}$$

$$T \rightarrow T \ \underline{\text{or}}\ A \quad \Rightarrow T_L = T_R + A_R \tag{5.63}$$

$$A \rightarrow B \qquad \Rightarrow A_L = B_R \tag{5.64}$$

$$A \rightarrow A \ \underline{\text{and}}\ B \Rightarrow A_L = A_R \cap B_R \tag{5.65}$$

$$B \rightarrow C \qquad \Rightarrow B_L = C_R \tag{5.66}$$

$$B \rightarrow \underline{\text{not}}\ C \quad \Rightarrow B_L = C_R \tag{5.67}$$

$$C \rightarrow (T) \qquad \Rightarrow C_L = T_R \tag{5.68}$$

$$C \rightarrow D \qquad \Rightarrow C_L = D_R \tag{5.69}$$

$$C \to E \quad \Rightarrow C_L = E_R \tag{5.70}$$

$$D \to \underline{\text{very}}\ D \Rightarrow D_L = (D_R)^2 \tag{5.71}$$

$$E \to \underline{\text{very}}\ E \Rightarrow E_L = (E_R)^2 \tag{5.72}$$

$$D \to \underline{\text{young}} \quad \Rightarrow D_L = \underline{\text{young}} \tag{5.73}$$

$$E \to \underline{\text{old}} \quad \Rightarrow E_L = \underline{\text{old}} \tag{5.74}$$

This dual system is employed in the following manner to compute the meaning of a composite term in T.

1. The term in question, e.g., <u>not</u> <u>very</u> <u>young</u> <u>and</u> <u>not</u> <u>very</u> <u>old</u> is parsed by the use of an appropriate parsing algorithm for G [37], yielding a syntax tree such as shown in Fig. 5.4. The leaves of this syntax tree are (a) primary terms whose meaning is specified a priori; (b) names of modifiers (i.e., hedges, connectives, negation, etc.): and (c) markers such as parentheses which serve as aids to parsing.

2. Starting from the bottom, the primary terms are assigned their meaning and, using the equations of (5.62), the meaning of nonterminals connected to the leaves is computed. Then, the subtrees which have these nonterminals as their roots are deleted, leaving the nonterminals in question as the leaves of the pruned tree. This process is repeated until the meaning of the term associated with the root of the syntax tree is computed.

In applying this procedure to the syntax tree shown in Fig. 5.5 , we first assign to <u>young</u> and <u>old</u> the meanings expressed by (5.45) and (5.46). Then, using (5.73) and (5.74) we find

$$D_7 = \underline{\text{young}} \tag{5.75}$$

and

$$E_{11} = \underline{old} \tag{5.76}$$

Next, using (5.71) and (5.72), we obtain

$$D_6 = E_7^2 = \underline{young}^2 \tag{5.77}$$

and

$$E_{10} = E_{11}^2 = \underline{old}^2 \tag{5.78}$$

Continuing in this manner, we obtain

$$C_5 = D_6 = \underline{young}^2 \tag{5.79}$$

$$C_9 = D_{10} = \underline{old}^2 \tag{5.80}$$

$$B_4 = \neg C_5 = \neg(\underline{young}^2) \tag{5.81}$$

$$B_8 = \neg C_9 = \neg(\underline{old}^2) \tag{5.82}$$

$$A_3 = B_4 = \neg(\underline{young}^2) \tag{5.83}$$

$$A_2 = A_3 \cap B_8 = \neg(\underline{young}^2) \cap \neg(\underline{old}^2) \tag{5.84}$$

and hence

$$\underline{not}\ \underline{very}\ \underline{young}\ \underline{and}\ \underline{not}\ \underline{very}\ \underline{old} = \neg(\underline{young}^2) \cap \neg(\underline{old}^2)$$

which agrees with the expression which we had obtained previously by inspection (see (5.44)).

The basic idea behind the procedure described above is to relate the meaning of a composite term to that of its constituent primary terms by means of a system of equations which are determined by the grammar which generates the terms in T. In the case of the Boolean linguistic variable of Example 5.10, this can be done by inspection. More generally, the nature of the hedges in the linguistic variable and its grammar G might be such as to make the computation of the meaning of its values a nontrivial problem.

Graphical Representation of a Linguistic Variable

A linguistic variable may be represented in a graphical form which is similar to that of an object in the Vienna definition language [41], [42], [43]. Specifically, a variable, $\mathcal{X}$, is represented as a branch (see Fig. 5.6) whose root is labeled $\mathcal{X}$ and whose edges are labeled with the names of the values of $\mathcal{X}$, i.e., $X_1, X_2, \ldots$. The object attached to the edge labeled X_i is the meaning of X_i. For example, in the case of the variable named Age, the edges might be labeled young, old, not young, etc., and the meaning of each such label can be represented as the graph of the membership function of the fuzzy set which is the meaning of the label in question (Fig. 5.7). It is important to note that, in the case of a structured linguistic variable, both the labels of the edges and the objects attached to them are generated algorithmically by the syntactic and semantic rules which are associated with the variable.

More generally, the graph of a linguistic variable may have the form of a tree rather than a single branch (see Fig. 5.8). In the case of a tree, it is understood that the name of a value of the variable is the concatenation of the names associated with an upward path from the leaf to the root. For example, in the tree of Fig. 5.8, the composite name associated with the path leading from node 3 to the root is very tall. quite fat. extremely intelligent.

This concludes our discussion of some of the basic aspects of the concept of a linguistic variable. In the following sections, we shall focus our attention on some of the applications of this concept.

6. Linguistic Truth Variables and Fuzzy Logic

In everyday discourse, we frequently characterize the degree of truth of a statement by expressions such as _very true_, _quite true_, _more or less true_, _essentially true_, _false_, _completely false_, etc. The similarity between these expressions and the values of a linguistic variable suggests that in situations in which the truth or falsity of an assertion is not well defined, it may be appropriate to treat _Truth_ as a linguistic variable for which _true_ and _false_ are merely two of the primary terms in its term-set rather than a pair of extreme points in the universe of truth-values. Such a variable and its values will be called a _linguistic truth variable_ and _linguistic truth-values_, respectively.

Treating truth as a linguistic variable leads to a _fuzzy linguistic logic_, or simply _fuzzy logic_, which is quite different from the conventional two-valued or even n-valued logic. This fuzzy logic provides a basis for what might be called _approximate reasoning_, that is, a mode of reasoning in which the truth-values and the rules of inference are fuzzy rather than precise. In many ways, approximate reasoning is akin to the reasoning used by humans in ill-defined or unquantifiable situations. Indeed, it may well be true that much - perhaps most - of human reasoning is approximate rather than precise in nature.

In the sequel, the term _proposition_ will be employed to denote statements of the form "u is A," where u is a name of an object and A is the name of a possibly fuzzy subset of a universe of discourse U, e.g., "John is young," "X is small," "apple is red," etc. If A is interpreted as a fuzzy predicate,[1] then the statement "u is A" may be paraphrased as

[1] More precisely, a fuzzy predicate may be viewed as the equivalent of the membership function of a fuzzy set. To simplify our terminology, both A and μ_A will be referred to as a fuzzy predicate.

"u has property A." Equivalently, "u is A" may be interpreted as an assignment equation in which a fuzzy set named A is assigned as a value to a linguistic variable which denotes an attribute of u, e.g.,

John is young $\leftrightarrow$ Age(John) = young

X is small $\leftrightarrow$ Magnitude (X) = small

apple is red $\leftrightarrow$ Color (apple) = red

A proposition such as "u is A" will be assumed to be associated with two fuzzy subsets: (i) The meaning of A, M(A), which is a fuzzy subset of U named A; and (ii) the truth-value of "u is A," or simply truth-value of A, which is denoted by v(A) and is defined to be a possibly fuzzy subset of a universe of truth-values V. In the case of two-valued logic, V = T + F (T $\triangleq$ true, F $\triangleq$ false). In what follows, unless stated to the contrary, it will be assumed that V = [0,1].

A truth-value which is a point in [0,1], e.g. v(A) = 0.8, will be referred to as a numerical truth-value. The numerical truth-values play the role of the values of the base variable for the linguistic variable Truth. The linguistic values of Truth will be referred to as linguistic truth-values. More specifically, we shall assume that Truth is the name of a Boolean linguistic variable in which the primary term is true, with false defined not as the negation of true,[2] but as its mirror image with respect to the point 0.5 in [0,1]. Typically, the term-set of Truth will be assumed to be the following

[2]As will be seen later (6.11), the definition of false as the mirror image of true is a consequence of defining false as the truth-value of not A under the assumption that the truth-value of A is true.

$$T(\underline{\text{Truth}}) = \underline{\text{true}} + \underline{\text{not}}\ \underline{\text{true}} + \underline{\text{very}}\ \underline{\text{true}} + \underline{\text{more}}\ \underline{\text{or}}\ \underline{\text{less}}\ \underline{\text{true}} + \underline{\text{very}}\ \underline{\text{very}}\ \underline{\text{true}} + \underline{\text{essentially}}\ \underline{\text{true}} + \underline{\text{very}}\ (\underline{\text{not}}\ \underline{\text{true}}) + \underline{\text{not}}\ \underline{\text{very}}\ \underline{\text{true}} + \dots + \underline{\text{false}} + \underline{\text{not}}\ \underline{\text{false}} + \underline{\text{very}}\ \underline{\text{false}} + \dots + \dots\ \underline{\text{not}}\ \underline{\text{very}}\ \underline{\text{true}}\ \underline{\text{and}}\ \underline{\text{not}}\ \underline{\text{very}}\ \underline{\text{false}} + \dots \tag{6.1}$$

in which the terms are the names of the truth-values.

The meaning of the primary term true is assumed to be a fuzzy subset of the interval $V = [0,1]$ characterized by a membership function of the form shown in Fig. 6.1. More precisely, true should be regarded as the name of a fuzzy variable whose restriction is the fuzzy set depicted in Fig. 6.1.

A possible approximation to the membership function of true is provided by the expression

$$\begin{aligned} \mu_{\underline{\text{true}}}(v) &= 0 \quad \text{for } 0 \leq v \leq a \\ &= 2\left(\frac{v-a}{1-a}\right)^2 \quad \text{for } a \leq v \leq \frac{a+1}{2} \\ &= 1 - \left(\frac{v-1}{1-a}\right)^2 \quad \text{for } \frac{a+1}{2} \leq v \leq 1 \end{aligned} \tag{6.2}$$

which has $v = \frac{1+a}{2}$ as its crossover point. (Note that the support of true is the interval $[a,1]$). Correspondingly, for false, we have (see Fig. 6.1)

$$\mu_{\underline{\text{false}}}(v) = \mu_{\underline{\text{true}}}(1-v) \quad , \quad 0 \leq v \leq 1.$$

In some instances it is simpler to assume that true is a subset of the finite universe of truth-values

$$V = 0 + 0.1 + 0.2 + \ldots + 0.9 + 1 \tag{6.3}$$

rather than of the unit interval $V = [0,1]$. With this assumption, true may be defined as, say,

$$\underline{\text{true}} = 0.5/0.7 + 0.7/0.8 + 0.9/0.9 + 1/1$$

where the pair 0.5/0.7, for example, means that the compatibility of the truth-value 0.7 with true is 0.5.

In what follows, our main concern will be with relations of the general form

$$\begin{aligned} &v(u \text{ is: linguistic value of a Boolean linguistic variable } \mathcal{X}) = \\ &\text{linguistic value of a Boolean linguistic truth-variable } \mathcal{T} \end{aligned} \tag{6.4}$$

as in

v(John is tall and dark and handsome) = not very true and not very false

where tall and dark and handsome is a linguistic value of a variable named $\mathcal{X} \triangleq$ Appearance, and not very true and not very false is that of a linguistic truth variable $\mathcal{T}$. In abbreviated form, (6.4) will usually be written as

$$v(X) = T$$

where X is a linguistic value of $\mathcal{X}$ and T is that of $\mathcal{T}$.

Now suppose that X_1, X_2 and $X_1 * X_2$, where $*$ is a binary connective, are linguistic values of $\mathcal{X}$ with respective truth-values $v(X_1)$, $v(X_2)$ and $v(X_1*X_2)$. A basic question that arises in this connection is whether or not it is possible to express $v(X_1*X_2)$ as a function of $v(X_1)$ and $v(X_2)$, that is, write

$$v(X_1 * X_2) = v(X_1) *' v(X_2) \tag{6.5}$$

where $*'$ is a binary connective associated with the linguistic truth variable $\mathcal{T}$.[3] It is this question that provides the motivation for the following discussion.

Logical Connectives in Fuzzy Logic

To construct a basis for fuzzy logic it is necessary to extend the meaning of such logical operations as negation, disjunction, conjunction and implication to operands which have linguistic rather than numerical truth-values. In other words, given propositions A and B we have to be able to compute the truth-value of, say, A and B from the knowledge of the linguistic truth-values of A and B.

In considering this problem it is helpful to observe that, if A is a fuzzy subset of a universe of discourse U and $u \in U$, then the two statements

a) The grade of membership of u in the fuzzy set A is $\mu_A(u)$ (6.6)

b) The truth-value of the fuzzy predicate A is $\mu_A(u)$

are equivalent. Thus, the question "What is the truth-value of A and B given the linguistic truth-values of A and B?" is similar to the question to which we had addressed ourselves in Sec. 3, namely, "What

[3] From an algebraic point of view, v may be regarded as a homomorphic mapping from $T(\mathcal{X})$, the term-set of $\mathcal{X}$, to $T(\mathcal{T})$, the term-set of $\mathcal{T}$, with $*'$ representing the operation in $T(\mathcal{T})$ induced by $*$.

is the grade of membership of u in $A \cap B$ given the fuzzy grades of membership of u in A and B?"

To answer the latter question we made use of the extension principle. The same procedure will be followed to extend the meaning of <u>not</u>, <u>and</u>, <u>or</u> and <u>implies</u> to linguistic truth-values.

Specifically, if $v(A)$ is a point in $V = [0,1]$ representing the truth-value of the proposition "u is A," (or simply A) where u is an element of a universe of discourse U, then truth value of <u>not</u> A (or $\neg A$) is given by

$$v(\underline{not}\ A) = 1 - v(A) \tag{6.7}$$

Now suppose that $v(A)$ is not a point in [0,1] but a fuzzy subset of [0,1] expressed as

$$v(A) = \mu_1/v_1 + \ldots + \mu_n/v_n \tag{6.8}$$

where the v_i are points in [0,1] and the μ_i are their grades of membership in $v(A)$. Then, by applying the extension principle (3.80) to (6.7), we obtain the expression for $v(\underline{not}\ A)$ as a fuzzy subset of [0,1], i.e.,

$$v(\underline{not}\ A) = \mu_1/(1-v_1) + \ldots + \mu_n/(1-v_n) \tag{6.9}$$

In particular, if the truth-value of A is true, i.e.,

$$v(A) = \underline{true} \tag{6.10}$$

then the truth-value <u>false</u> may be defined as

$$\underline{false} \triangleq v(\underline{not}\ A) \tag{6.11}$$

For example, if

$$\underline{\text{true}} = 0.5/7 + 0.7/0.8 + 0.9/0.9 + 1/1 \tag{6.12}$$

then the truth-value of <u>not</u> A is given by

$$\underline{\text{false}} = v(\underline{\text{not}}\ A) = 0.5/0.3 + 0.7/0.2 + 0.9/0.1 + 1/0$$

<u>Comment 6.1</u> It should be noted that if

$$\underline{\text{true}} = \mu_1/v_1 + \ldots + \mu_n/v_n \tag{6.13}$$

then by (3.33)

$$\underline{\text{not}}\ \underline{\text{true}} = (1-\mu_1)/v_1 + \ldots + (1-\mu_n)/v_n \tag{6.14}$$

By contrast, if

$$\begin{aligned} v(A) &= \underline{\text{true}} \\ &= \mu_1/v_1 + \ldots + \mu_n/v_n \end{aligned} \tag{6.15}$$

then

$$\begin{aligned} \underline{\text{false}} &= v(\underline{\text{not}}\ A) \\ &= \mu_1/(1-v_1) + \ldots + \mu_n/(1-v_n) \end{aligned} \tag{6.16}$$

The same applies to hedges. For example, by the definition of <u>very</u> (see (5.38))

$$\underline{\text{very}}\ \underline{\text{true}} = \mu_1^2/v_1 + \ldots + \mu_n^2/v_n \tag{6.17}$$

On the other hand, the truth-value of <u>very</u> A is expressed by

$$v(\underline{\text{very}}\ A) = \mu_1/v_1^2 + \ldots + \mu_n/v_n^2 \tag{6.18}$$

Turning our attention to binary connectives, let $v(A)$ and $v(B)$ be the linguistic truth-values of propositions A and B, respectively. To simplify the notation, we shall adopt the convention of writing - as in the case where $v(A)$ and $v(B)$ are points in [0,1]:

$$v(A) \wedge v(B) \quad \text{for} \quad v(A \text{ \underline{and} } B) \tag{6.19}$$

$$v(A) \vee v(B) \quad \text{for} \quad v(A \text{ \underline{or} } B) \tag{6.20}$$

$$v(A) \Rightarrow v(B) \quad \text{for} \quad v(A \Rightarrow B) \tag{6.21}$$

and

$$\neg v(A) \quad \text{for} \quad v(\text{\underline{not} } A) \tag{6.22}$$

with the understanding that $\wedge$, $\vee$ and $\neg$ reduce to Min (conjunction), Max (disjunction) and 1- operations when $v(A)$ and $v(B)$ are points in [0,1].

Now if $v(A)$ and $v(B)$ are linguistic truth-values expressed as

$$v(A) = \alpha_1/v_1 + \dots + \alpha_n/v_n \tag{6.23}$$

$$v(B) = \beta_1/w_1 + \dots + \beta_m/w_m \tag{6.24}$$

where the v_i and w_j are points in [0,1] and the α_i and β_j are their respective grades of membership in A and B, then by applying the extension principle to $v(A \text{ \underline{and} } B)$, we obtain

$$\begin{aligned} v(A \text{ \underline{and} } B) &= v(A) \wedge v(B) \\ &= (\alpha_1/v_1+\dots+\alpha_n/v_n) \wedge (\beta_1/w_1+\dots+\beta_m/w_m) \\ &= \sum_{i,j} (\alpha_i \wedge \beta_j)/(v_i \wedge w_j) \end{aligned} \tag{6.25}$$

Thus, the truth-value of A and B is a fuzzy subset of [0,1] whose support is comprised of the points $v_i \wedge w_j$, $i = 1, \ldots, n$, $j = 1, \ldots, m$, with respective grades of membership $(\alpha_i \wedge \beta_j)$. Note that (6.25) is equivalent to the expression (3.107) for the membership function of the intersection of fuzzy sets having fuzzy membership functions.

Example 6.2 Suppose that

$$\begin{aligned} v(A) &= \underline{\text{true}} \\ &= 0.5/0.7 + 0.7/0.8 + 0.9/0.9 + 1/1 \end{aligned} \tag{6.26}$$

and

$$\begin{aligned} v(B) &= \underline{\text{not}}\ \underline{\text{true}} \\ &= 1/0 + 1/0.1 + 1/0.2 + 1/0.3 + 1/0.4 + 1/0.5 + 1/0.6 \\ &\quad + 0.5/0.7 + 0.3/0.8 + 0.1/0.9 \end{aligned} \tag{6.27}$$

Then, the use of (6.25) leads to

$$\begin{aligned} v(A\ \underline{\text{and}}\ B) &= \underline{\text{true}} \wedge \underline{\text{not}}\ \underline{\text{true}} \\ &= 1/(0 + 0.1 + 0.2 + 0.3 + 0.4 + 0.5 + 0.6) + 0.5/0.7 \\ &\quad + 0.3/0.8 + 0.1/0.9 \\ &= \underline{\text{not}}\ \underline{\text{true}} \end{aligned} \tag{6.28}$$

In a similar fashion, for the truth-value of A or B, we obtain

$$\begin{aligned} v(A\ \underline{\text{or}}\ B) &= v(A) \vee v(B) \\ &= (\alpha_1/v_1 + \ldots + \alpha_n/v_n) \vee (\beta_1/w_1 + \ldots + \beta_m/w_m) \\ &= \sum_{i,j} (\alpha_i \wedge \beta_j)/(v_i \vee w_j) \end{aligned} \tag{6.29}$$

The truth-value of $A \Rightarrow B$ depends on the manner in which the connective $\Rightarrow$ is defined for numerical truth-values. Thus, if we define (see (8.24))

$$v(A \Rightarrow B) = \neg v(A) \vee v(A) \wedge v(B) \tag{6.30}$$

for the case where v(A) and v(B) are points in [0,1], then the application of the extension principle yields (see Comment 3.20)

$$\begin{aligned} v(A \Rightarrow B) &= ((\alpha_1/v_1+\ldots+\alpha_n/v_n) \Rightarrow (\beta_1/w_1+\ldots+\beta_m/w_m)) \\ &= \sum_{i,j} (\alpha_i \wedge \beta_j)/(1-v_i) \vee (v_i \wedge w_j) \end{aligned} \tag{6.31}$$

for the case where v(A) and v(B) are fuzzy subsets of [0,1].

<u>Comment 6.3</u> It is important to have a clear understanding of the difference between <u>and</u> in, say, <u>true</u> <u>and</u> <u>not</u> <u>true</u>, and $\wedge$ in <u>true</u> $\wedge$ <u>not</u> <u>true</u>. In the former, our concern is with the meaning of the term <u>true</u> <u>and</u> <u>not</u> <u>true</u>, and <u>and</u> is defined by the realtion

$$M(\underline{\text{true}}\ \underline{\text{and}}\ \underline{\text{not}}\ \underline{\text{true}}) = M(\underline{\text{true}}) \cap M(\underline{\text{not}}\ \underline{\text{true}}) \tag{6.32}$$

where M is the function mapping a term into its meaning (see Definition 5.1). By contrast, in the case of <u>true</u> $\wedge$ <u>not</u> <u>true</u> we are concerned with the truth-value of <u>true</u> $\wedge$ <u>not</u> <u>true</u>, which is derived from the equivalence (see (6.19))

$$v(A\ \underline{\text{and}}\ B) = v(A) \wedge v(B) \tag{6.33}$$

Thus, in (6.32) $\cap$ is the operation of intersection of fuzzy sets, whereas in (6.33), $\wedge$ is that of conjunction. To illustrate the difference by a simple example, let $V = 0 + 0.1 + \ldots + 1$, and let P and Q be fuzzy subsets of V defined by

$$P = 0.5/0.3 + 0.8/0.7 + 0.6/1 \tag{6.34}$$

$$Q = 0.1/0.3 + 0.6/0.7 + 1/1 \tag{6.35}$$

Then

$$P \cap Q = 0.1/0.3 + 0.6/0.7 + 0.6/1 \tag{6.36}$$

whereas

$$P \wedge Q = 0.5/0.3 + 0.8/0.7 + 0.6/1. \tag{6.37}$$

Note that the same issue arises in the case of _not_ and $\neg$, as pointed out in Comment 6.1.

Comment 6.4 It should be noted that in applying the extension principle (3.96) to the computation of v(A _and_ B), v(A _or_ B) and v(A $\Rightarrow$ B), we are tacitly assuming that v(A) and v(B) are noninteractive fuzzy variables in the sense of Comment 3.20. If v(A) and v(B) are interactive, then it is necessary to apply the extension principle as expressed by (3.97) rather than (3.96). It is of interest to observe that the issue of possible interaction between v(A) and v(B) arises even when v(A) and v(B) are points in [0,1] rather than fuzzy variables.

Comment 6.5 By employing the extension principle to define the operations $\wedge$, $\vee$, $\neg$ and $\Rightarrow$ on linguistic truth-values, we are in effect treating fuzzy logic as an extension of multivalued logic. In the same sense, the classical three-valued logic may be viewed as an extension of two-valued logic (see 6.64 _et seq_.).

The expressions for v(_not_ A), v(A _and_ B), v(A _or_ B) and v(A $\Rightarrow$ B) given above become more transparent if we first decompose v(A) and v(B) into level-sets and then apply the level-set form of the extension principle (see (3.86)) to the operations $\neg$, $\wedge$, v and $\Rightarrow$. In this way, we are led to a simple graphical rule for computing the truth-values in question (see Fig. 6.2). Specifically, let the intervals $[a_1,a_2]$ and $[b_1,b_2]$ be the α-level-sets for v(A) and v(B). Then, by using the extensions of the operations $\neg$, $\wedge$ and $\vee$ to intervals, namely (see (3.100))

$$\neg[a_1,a_2] = [1 - a_2, 1 - a_1] \tag{6.38}$$

$$[a_1,a_2] \wedge [b_1,b_2] = [a_1 \wedge b_1, a_2 \wedge b_2] \tag{6.39}$$

$$[a_1,a_2] \vee [b_1,b_2] = [a_1 \vee b_1, a_2 \vee b_2] \tag{6.40}$$

we can find by inspection the α-level-sets for v(<u>not</u> A), v(A <u>and</u> B) and v(A <u>or</u> B). Having found these level-sets, v(<u>not</u> A), v(A <u>and</u> B) and v(A <u>or</u> B) can readily be determined by varying α from 0 to 1.

As a simple illustration, consider the determination of the conjunction of linguistic truth-values v(A) $\triangleq$ <u>true</u> and v(B) $\triangleq$ <u>false</u>, with the membership functions of <u>true</u> and <u>false</u> having the form shown in Fig. 6.1.

We observe that, for all values of α,

$$[a_1,a_2] \wedge [b_1,b_2] = [b_1,b_2] \tag{6.41}$$

which implies that (see (3.118))

$$[b_1,b_2] \leq [a_1,a_2] \tag{6.42}$$

Consequently, merely on the basis of the form of the membership functions of <u>true</u> and <u>false</u>, we can conclude that

$$\underline{\text{true}} \wedge \underline{\text{false}} = \underline{\text{false}} \tag{6.43}$$

which is consistent with (6.25).

Truth Tables and Linguistic Approximation

In two-valued, three-valued and, more generally, n-valued logics the binary connectives $\wedge$, $\vee$ and $\Rightarrow$ are usually defined by a tabulation of the truth-values of A and B, A or B and A $\Rightarrow$ B in terms of the truth-values of A and B.

Since in a fuzzy logic the number of truth-values is, in general, infinite, $\wedge$, $\vee$ and $\Rightarrow$ cannot be defined by tabulation. However, it may be desirable to tabulate say, $\wedge$, for a finite set of truth-values of interest, e.g., true, not true, false, very true, very (not true), more or less true, etc. In such a table, for an entry in ith row, say not true, and an entry in jth column, say more or less true, the (i,j)th entry would be

$$\text{(i,j)th entry} = \text{ith row entry } (\triangleq \text{not true}) \wedge \text{jth column entry } (\triangleq \text{more or less true}) \tag{6.44}$$

Given the definition of the primary term true and the definitions of the modifiers not and more or less, we can compute the right-hand member of (6.44), that is,

$$\text{not true} \wedge \text{more or less true} \tag{6.45}$$

by using (6.25). However, the problem is that in most instances the result of the computation would be a fuzzy subset of the universe of truth-values which may not correspond to any of the truth-values in the term-set of Truth. Thus, if we wish to have a truth table in which the entries are linguistic, we must be content with an approximation to the exact truth-value of (ith row entry $\wedge$ jth column entry). Such an approximation will be referred to as a linguistic approximation. (See Fig. 1.5.)

As an illustration, suppose that the universe of truth-values is expressed as

$$V = 0 + 0.1 + 0.2 + \dots + 1 \tag{6.46}$$

and that

$$\underline{true} = 0.7/0.8 + 1/0.9 + 1/1 \tag{6.47}$$

$$\underline{more}\ \underline{or}\ \underline{less}\ \underline{true} = 0.5/0.6 + 0.7/0.7 + 1/0.8 + 1/0.9 + 1/1 \tag{6.48}$$

and

$$\underline{almost}\ \underline{true} = 0.6/0.8 + 1/0.9 + 0.6/1 \tag{6.49}$$

In the truth-table for $\vee$, assume that the ith row entry is more or less true and the jth column entry is almost true. Then, for the (i,j)th entry in the table, we have

$$\begin{aligned} \underline{more}\ \underline{or}\ \underline{less}\ \underline{true} \vee \underline{almost}\ \underline{true} &= (0.5/0.6 + 0.7/0.7 + 1/0.8 + 1/0.9 \\ &\quad + 1/1) \vee (0.6/0.8 + 1/0.9 + 0.6/1) \\ &= 0.6/0.8 + 1/0.9 + 1/1 \end{aligned} \tag{6.50}$$

Now, we observe the right-hand member of (6.50) is approximately equal to true as defined by (6.47). Consequently, in the truth table for $\vee$, a linguistic approximation to the (i,j)th entry would be true.

The Truth-Values *Unknown* and *Undefined*

Among the truth-values that can be associated with the linguistic variable *Truth*, there are two that warrant special attention, namely, the empty set, θ, and the unit interval [0,1]-which correspond to the the least and greatest elements (under set inclusion) of the lattice of fuzzy subsets of [0,1]. The importance of these particular truth-values stems from their interpretability as the truth-values *undefined* and *unknown*,[4] respectively. For convenience we shall denote these truth-values by θ and ?, with the understanding that θ and ? are defined by

$$\theta \triangleq \int_0^1 0/v \tag{6.51}$$

and

$$\begin{aligned} ? &\triangleq V = \text{universe of truth-values} \\ &= [0,1] \\ &= \int_0^1 1/w \end{aligned} \tag{6.52}$$

Interpreted as grades of membership, *undefined* and *unknown* enter also in the representation of fuzzy sets of type 1. For such sets, the grade of membership of a point u in U may have one of three possible forms: (i) a number in the interval [0,1]; (ii) θ (*undefined*); and (iii) ? (*unknown*). As a simple example, let

[4] The concept of *unknown* is related to that of *don't care* in the context of switching circuits [44]. Another related concept is that of *possible* in modal logic [45].

$$U = a + b + c + d + e \tag{6.53}$$

and consider a fuzzy subset of U represented as

$$A = 0.1\ a + 0.9\ b + ?c + \theta d \tag{6.54}$$

In this case, the grade of membership of c in A is <u>unknown</u> and that of d is <u>undefined</u>. More generally, we may have

$$A = 0.1\ a + 0.9\ b + 0.8\ ?\ c + \theta\ d \tag{6.55}$$

meaning that the grade of membership of c in A is partially unknown, with 0.8 ? c interpreted as

$$0.8\ ?c = \left(\int_0^1 0.8/v\right) /c \tag{6.56}$$

It is important to have a clear understanding of the difference between 0 and θ. When we say that the grade of membership of a point u in A is θ, what we mean is that the membership function $\mu_A : U \rightarrow [0,1]$ is undefined at u. For example, suppose that U is the set of real numbers and μ_A is a function defined on integers, with $\mu_A(u) = 1$ if n is an even integer and $\mu_A(u) = 0$ if u is an odd integer. Then the grade of membership of u = 1.5 in A is θ rather than 0. On the other hand, if μ_A were defined on real numbers and $\mu_A(u) = 1$ iff n is even, then the grade of membership of 1.5 in A would be 0.

Since we know how to compute the truth-values of A <u>and</u> B, A <u>or</u> B and <u>not</u> B given the linguistic truth-values of A and B, it is a simple matter to compute v(A <u>and</u> B), v(A <u>or</u> B) and v(<u>not</u> B) when v(B) = ? .

Thus, suppose that

$$v(A) = \int_0^1 \mu(v)/v \tag{6.57}$$

and

$$v(B) = ? = \int_0^1 1/w \tag{6.58}$$

By applying the extension principle, as in (6.25), we obtain

$$\begin{aligned} v(A) \wedge ? &= \int_0^1 \mu(v)/v \wedge \int_0^1 1/w \\ &= \int_0^1 \int_0^1 \mu(v)/(v \wedge w) \end{aligned} \tag{6.59}$$

where

$$\int_0^1 \int_0^1 \triangleq \int_{[0,1] \times [0,1]} \tag{6.60}$$

and which upon simplification reduces to

$$v(A) \wedge ? = \int_0^1 (\vee_{[w,1]} \mu(v))/w \tag{6.61}$$

In other words, the truth-value of A and B, where v(B) = unknown, is a fuzzy subset of [0,1] in which the grade of membership of a point w is given by the supremum of $\mu(v)$ (membership function of A) over the interval [w,1].

In a similar fashion, the truth-value of A <u>or</u> B is found to be expressed by

$$v(A \text{ } \underline{or} \text{ } B) = \int_0^1 \int_0^1 \mu(v)/(v \vee w) \tag{6.62}$$

$$= \int_0^1 (\vee_{[0,w]} \mu(v))/w$$

It should be noted that both (6.61) and (6.62) can readily be obtained by the graphical procedure described earlier (see (6.38) et seq.). An example illustrating its application is shown in Fig. 6.4.

Turning to the case where v(B) = θ, we find

$$v(A) \wedge \theta = \int_0^1 \int_0^1 0/(v \wedge w) \tag{6.63}$$

$$= \int_0^1 0/w$$

$$= \theta$$

and likewise for $v(A) \vee \theta$.

It is instructive to examine what happens to the above relations when we apply them to the special case of two-valued logic, that is, to the case where the universe V is of the form

$$V = 0 + 1 \tag{6.64}$$

or, expressed more conventionally,

$$V = T + F \tag{6.65}$$

where T stands for <u>true</u> and F stands for <u>false</u>. Since ? is V, we can identify the truth-value <u>unknown</u> with <u>true</u> or <u>false</u>, that is,

$$? = T + F \tag{6.66}$$

The resulting logic has four truth-values: θ, T, F and T + F ($\triangleq ?$) and is an extension of two-valued logic in the sense of Comment 6.5.

Since the universe of truth-values has only two elements, it is expedient to derive the truth tables for $\vee$, $\wedge$ and $\Rightarrow$ in this four-valued logic directly rather than through specialization of the general formulae (6.25), (6.29) and (6.31). Thus, by applying the extension principle to $\wedge$, we find at once

$$T \wedge \theta = \theta \tag{6.67}$$

$$\begin{aligned} T \wedge (T + F) &= T \wedge T + T \wedge F \\ &= T + F \end{aligned} \tag{6.68}$$

$$\begin{aligned} F \wedge (T + F) &= F \wedge T + F \wedge F \\ &= F + F \\ &= F \end{aligned} \tag{6.69}$$

$$\begin{aligned} (T + F) \wedge (T + F) &= T \wedge T + T \wedge F + F \wedge T + F \wedge F \\ &= T + F + F + F \\ &= T + F \end{aligned} \tag{6.70}$$

and consequently the extended truth-table for $\wedge$ has the form shown in Table 6.5.

$\wedge$	θ	T	F	T + F
θ	θ	θ	θ	θ
T	θ	T	F	T + F
F	θ	F	F	F
T + F	θ	T + F	F	T + F

Table 6.5

which upon suppression of the entry θ reads

$\wedge$	T	F	T + F
T	T	F	T + F
F	F	F	F
T + F	T + F	F	T + F

Table 6.6

Similarly, for the operation $\vee$ we obtain

$\vee$	T	F	T + F
T	T	T	T
F	T	F	T + F
T + F	T	T + F	T

Table 6.7

These tables agree - as they should - with the corresponding truth tables for $\wedge$ and $\vee$ in conventional three-valued logic [46].

The approach employed above provides some insight into the definition of $\Rightarrow$ in two-valued logic - a somewhat controversial issue which motivated the development of modal logic [45], [47]. Specifically, instead of defining $\Rightarrow$ in the conventional fashion, we may define $\Rightarrow$ as a connective in three-valued logic by the partial truth table

$\Rightarrow$	T	F	T + F
T	T	F	
F			T

Table 6.8

which expresses the untuitively reasonable idea that if $A \Rightarrow B$ is <u>true</u> and A is <u>false</u>, then the truth-value of B is <u>unknown</u>. Now we can raise the question: How should the blank entries in the above table be filled in order to yield the entry T in the (2,3) position in Table 6.8 upon the application of the extension principle? Thus, denoting the unknown entries in positions (2,1) and (2,2) by x and y, respectively, we must have

$$\begin{aligned} F \Rightarrow (T + F) &= (F \Rightarrow T) + (F \Rightarrow F) \\ &= x + y \\ &= T \end{aligned} \tag{6.71}$$

which necessitates that

$$x = y = T \tag{6.72}$$

In this way, we are led to the conventional definition of $\Rightarrow$ in two-valued logic, which is expressed by the truth table

$\Rightarrow$	T	F
T	T	F
F	T	T

As the above example demonstrates, the notion of the <u>unknown</u> truth-value in conjunction with the extension principle helps to clarify some of the concepts and relations in the conventional two-valued and three-valued logics. These logics may be viewed, of course, as degenerate cases of a fuzzy logic in which the truth-value unknown is the entire unit interval rather than the set 0 + 1.

Composite Truth Variables and Truth-Value Distributions

In the foregoing discussion, we have limited our attention to linguistic truth variables which are unary variables in the sense of Definition 2.1. In the following, we shall define the concept of a composite truth variable and dwell briefly on some of its implications.

Thus, let

$$\mathcal{T} \triangleq (\mathcal{T}_1, \ldots, \mathcal{T}_n) \tag{6.73}$$

denote an n-ary composite linguistic truth variable in which each $\mathcal{T}_i$, i=1, ..., n, is a unary linguistic truth variable associated with a term-set T_i, a universe of discourse V_i, and a base variable v_i (see Definition 5.1). For simplicity, we shall sometimes employ the symbol $\mathcal{T}_i$ in the dual role of (a) the name of the ith variable in (6.73); and (b) a generic name for the truth-values of $\mathcal{T}_i$. Furthermore, we shall assume that $T_1 = T_2 = \ldots = T_n$ and $V_1 = V_2 = \ldots = V_n = [0,1]$.

Viewed as a composite variable whose component variables $\mathcal{T}_1, \ldots, \mathcal{T}_n$ take values in their respective universes $T_1, \ldots, T_n$, $\mathcal{T}$ is an n-ary nonfuzzy variable (see (2.3) et seq.). Thus, the restriction $R(\mathcal{T})$ imposed by $\mathcal{T}$ is an n-ary nonfuzzy relation in $T_1 \times \ldots \times T_n$ which may be represented as an unordered list of ordered n-tuples of the form

$$\begin{aligned} R(\mathcal{T}) = &(\underline{\text{true}}, \underline{\text{very}}\ \underline{\text{true}}, \underline{\text{false}}, \ldots, \underline{\text{quite}}\ \underline{\text{true}}) \\ &+ (\underline{\text{quite}}\ \underline{\text{true}}, \underline{\text{true}}, \underline{\text{very}}\ \underline{\text{true}}, \ldots, \underline{\text{very}}\ \underline{\text{true}}) \\ &+ (\text{true}, \text{true}, \underline{\text{more}}\ \underline{\text{or}}\ \underline{\text{less}}\ \underline{\text{true}}, \ldots, \text{true}) \\ &+ \ldots \end{aligned} \tag{6.74}$$

The n-tuples in $R(\mathcal{T})$ will be referred to as truth-value assignment lists, since each such n-tuple may be interpreted as an assignment of truth-values to a list of propositions $A_1, \ldots, A_n$, with

$$A \triangleq (A_1, \ldots, A_n) \tag{6.75}$$

representing a composite proposition. For example, if

$$A \triangleq (\text{Scott is } \underline{\text{tall}}, \text{ Pat is } \underline{\text{dark-haired}}, \text{ Tina is } \underline{\text{very pretty}})$$

then a triple in $R(\mathcal{T})$ of the form (very true, true, very true) would represent the following truth-value assignments:

$$v(\text{Scott is } \underline{\text{tall}}) = \underline{\text{very true}} \tag{6.76}$$

$$v(\text{Pat is } \underline{\text{dark-haired}}) = \underline{\text{true}} \tag{6.77}$$

$$v(\text{Tina is } \underline{\text{very pretty}}) = \underline{\text{very true}} \tag{6.78}$$

Based on this interpretation of the n-tuples in $R(\mathcal{T})$, we shall frequently refer to $R(\mathcal{T})$ as a truth-value distribution. Correspondingly, the restriction $R(\mathcal{T}_{i_1}, \ldots, \mathcal{T}_{i_k})$ which is imposed by the k-ary variable $(\mathcal{T}_{i_1}, \ldots, \mathcal{T}_{i_k})$, where $q = (i_1, \ldots, i_k)$ is a subsequence of the index sequence $(1, \ldots, n)$, will be referred to as a marginal truth-value distribution induced by $R(\mathcal{T}_1, \ldots, \mathcal{T}_n)$ (see (2.8)). Then, using the notation employed in Sec. 2 (see also Note 4.7), the relation between $R(\mathcal{T}_{i_1}, \ldots, \mathcal{T}_{i_k})$ and $R(\mathcal{T}_1, \ldots, \mathcal{T}_n)$ may be expressed compactly as

$$R(\mathcal{T}_{(q)}) = P_q\, R(\mathcal{T}) \tag{6.79}$$

where P_q denotes the operation of projection on the cartesian product $T_{i_1} \times \ldots \times T_{i_k}$.

Example 6.6 Suppose that $R(\mathcal{T})$ is expressed by

$$R(\mathcal{T}) \triangleq R(\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3) \tag{6.80}$$

$$= (\underline{true}, \underline{quite}\ \underline{true}, \underline{very}\ \underline{true})$$

$$+ (\underline{very}\ \underline{true}, \underline{true}, \underline{very}\ \underline{very}\ \underline{true})$$

$$+ (\underline{true}, \underline{false}, \underline{quite}\ \underline{true})$$

$$+ (\underline{false}, \underline{false}, \underline{very}\ \underline{true})$$

To obtain $R(\mathcal{T}_1, \mathcal{T}_2)$ we delete the $\mathcal{T}_3$ component in each triple, yielding

$$R(\mathcal{T}_1, \mathcal{T}_2) = (\underline{true}, \underline{quite}\ \underline{true}) \tag{6.81}$$

$$+ (\underline{very}\ \underline{true}, \underline{true})$$

$$+ (\underline{true}, \underline{false})$$

$$+ (\underline{false}, \underline{false})$$

Similarly, by deleting the $\mathcal{T}_2$ components in $R(\mathcal{T}_1, \mathcal{T}_2)$, we obtain

$$R(\mathcal{T}_1) = \underline{true} + \underline{very}\ \underline{true} + \underline{false} \tag{6.82}$$

If we view $\mathcal{T}$ as an n-ary nonfuzzy variable whose values are linguistic truth-values, the definition of noninteraction (Definition 2.9) assumes the following form in the case of linguistic truth variables.

Definition 6.7 The components of an n-ary linguistic truth variable $\mathcal{T} = (\mathcal{T}_1, \ldots, \mathcal{T}_n)$ are λ-noninteractive (λ standing for linguistic) iff the truth-value distribution $R(\mathcal{T}_1, \ldots, \mathcal{T}_n)$ is separable in the sense that

$$R(\mathcal{T}_1, \ldots, \mathcal{T}_n) = R(\mathcal{T}_1) \times \ldots \times R(\mathcal{T}_n) \tag{6.83}$$

The implication of this definition is that, if $\mathcal{T}_1, \ldots, \mathcal{T}_n$ are λ-noninteractive, then the assignment of specific linguistic truth-values to $\mathcal{T}_{i_1},\ldots,\mathcal{T}_{i_k}$ does not affect the truth-values that can be assigned to the complementary components in $(\mathcal{T}_1,\ldots,\mathcal{T}_n)$, $\mathcal{T}_{j_1},\ldots,\mathcal{T}_{j_m}$.

Before proceeding to illustrate the concept of λ-noninteraction by examples, we shall define another type of noninteraction which will be referred to as β-<u>noninteraction</u> (β standing for base variable).

<u>Definition 6.8</u> The components of an n-ary linguistic truth variable $\mathcal{T} = (\mathcal{T}_1,\ldots,\mathcal{T}_n)$ are β-<u>noninteractive</u> iff their respective base variables $v_1,\ldots,v_n$ are noninteractive in the sense of Definition 2.9; that is, the v_i are not jointly constrained.

To illustrate the concepts of noninteraction defined above we shall consider a few simple examples.

<u>Example 6.9</u> For the truth-value distribution of Example 6.6, we have

$$\begin{aligned} R(\mathcal{T}_1) &= \underline{\text{true}} + \underline{\text{very}}\ \underline{\text{true}} + \underline{\text{false}} \\ R(\mathcal{T}_2) &= \underline{\text{quite}}\ \underline{\text{true}} + \underline{\text{true}} + \underline{\text{false}} \\ R(\mathcal{T}_3) &= \underline{\text{very}}\ \underline{\text{true}} + \underline{\text{very}}\ \underline{\text{very}}\ \underline{\text{true}} + \underline{\text{quite}}\ \underline{\text{true}} \end{aligned} \tag{6.84}$$

and thus

$$\begin{aligned} R(\mathcal{T}_1)\times R(\mathcal{T}_2)\times R(\mathcal{T}_3) &= (\underline{\text{true}},\ \underline{\text{quite}}\ \underline{\text{true}},\ \underline{\text{very}}\ \underline{\text{true}}) \\ &+ (\underline{\text{very}}\ \underline{\text{true}},\ \underline{\text{quite}}\ \underline{\text{true}},\ \underline{\text{very}}\ \underline{\text{true}}) \\ &- - - - - - - - - - - - - - - - - - \\ &+ (\underline{\text{false}},\ \underline{\text{false}},\ \underline{\text{quite}}\ \underline{\text{true}}) \\ &\neq R(\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3) \end{aligned} \tag{6.85}$$

which implies that $R(\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3)$ is not separable and hence $\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3$

are λ-interactive.

Example 6.10 Consider a composite proposition of the form (A, not A) and assume for simplicity that $T(\mathcal{T}) = \underline{true} + \underline{false}$. In view of (6.11), if the truth-value of A is true then that of not A is false, and vice-versa. Consequently, the truth-value distribution for the propositions in question must be of the form

$$R(\mathcal{T}_1, \mathcal{T}_2) = (\underline{true}, \underline{false}) + (\underline{false}, \underline{true}) \tag{6.86}$$

which induces

$$R(\mathcal{T}_1) = R(\mathcal{T}_2) = \underline{true} + \underline{false} \tag{6.87}$$

Now

$$\begin{aligned} R(\mathcal{T}_1) \times R(\mathcal{T}_2) &= (\underline{true} + \underline{false}) \times (\underline{true} + \underline{false}) \\ &= (\underline{true}, \underline{true}) + (\underline{true}, \underline{false}) \\ &\quad + (\underline{false}, \underline{true}) + (\underline{false}, \underline{false}) \end{aligned} \tag{6.88}$$

and since

$$R(\mathcal{T}_1, \mathcal{T}_2) \neq R(\mathcal{T}_1) \times R(\mathcal{T}_2)$$

it follows that $\mathcal{T}_1$ and $\mathcal{T}_2$ are λ-interactive.

Example 6.11 The above example can also be used as an illustration of β-interaction. Specifically, regardless of the truth-values assigned to A and not A, it follows from the definition of not (see (3.33)) that the base variables v_1 and v_2 are constrained by the equation

$$v_1 + v_2 = 1 \tag{6.89}$$

In other words, in the case of a composite proposition of the form (A, <u>not</u> A), the sum of the numerical truth-values of A and <u>not</u> A must be unity.

<u>Remark 6.12</u> It should be noted that, in Example 6.11, β-interaction is a consequence of A_2 being related to A_1 by negation. In general, however, $\mathcal{T}_1,\ldots,\mathcal{T}_n$ may be λ-interactive without being β-interactive, and vice-versa.

A useful application of the concept of interaction relates to the truth-value <u>unknown</u> (see (6.52)). Specifically, assuming for simplicity that V = T + F, suppose that

$$A_1 \triangleq \text{Pat lives in Berkeley} \tag{6.90}$$

$$A_2 \triangleq \text{Pat lives in San Francisco} \tag{6.91}$$

with the understanding that one and only one of these statements is true. This implies that, although the truth-values of A_1 and A_2 are <u>unknown</u> ($\triangleq$? = T + F), that is,

$$\begin{aligned} v(A_1) &= T + F \\ v(A_2) &= T + F \end{aligned} \tag{6.92}$$

they are constrained by the relations

$$v(A_1) \vee v(A_2) = T \tag{6.93}$$

$$v(A_1) \wedge v(A_2) = F \tag{6.94}$$

Equivalently, the truth-value distribution associated with (6.90) and (6.91) may be regarded as the solution of the equations

$$v(A_1) \vee v(A_2) = T \tag{6.95}$$

$$v(A_1) \wedge v(A_2) = F \tag{6.96}$$

which is

$$R(\mathcal{T}_1, \mathcal{T}_2) = (T,F) + (F,T) \tag{6.97}$$

Note that (6.97) implies

$$v(A_1) = R(\mathcal{T}_1) = T + F \tag{6.98}$$

and

$$v(A_2) = R(\mathcal{T}_2) = T + F \tag{6.99}$$

in agreement with (6.92). Note also that $\mathcal{T}_1$ and $\mathcal{T}_2$ are β-interactive in the sense of Definition 6.8, with $V = T + F$.

Now if A_1 and A_2 were changed to

$$A_1 \triangleq \text{Pat lived in Berkeley} \tag{6.100}$$

$$A_2 \triangleq \text{Pat lived in San Francisco} \tag{6.101}$$

with the possibility that both A_1 and A_2 could be true, then we would still have

$$v(A_1) = ? = T + F \tag{6.102}$$

$$v(A_2) = ? = T + F \tag{6.103}$$

but the constraint equation would become

$$v(A_1) \vee v(A_2) = T \tag{6.104}$$

In this case, the truth-value distribution is the solution of (6.104), which is given by

$$R(\mathcal{T}_1, \mathcal{T}_2) = (\underline{\text{true}}, \underline{\text{true}}) + (\underline{\text{true}}, \underline{\text{false}}) + (\underline{\text{false}}, \underline{\text{true}}) \tag{6.105}$$

An important observation that should be made in connection with the above examples is that in some cases a truth-value distribution may be given in an implicit from, e.g., as a solution of a set of truth-value equations, rather than as an explicit list of ordered n-tuples of truth-values. In general, this will be the case where linguistic truth-values are assigned not to each A_i in $A = (A_1, \ldots, A_n)$ but to Boolean expressions involving two or more of the components of A.

Another point that should be noted is that truth-value distributions may be nested. As a simple illustration, in the case of a unary proposition we may have a nested sequence of assertions of the form

""""Vera is <u>very very intelligent</u>" is very <u>true</u>" is <u>true</u>."

Restrictions induced by assertions of this type may be computed as follows.

Let the base variable in (6.106) be IQ, and let $R_o(IQ)$ denote the restriction on the IQ of Vera. Then the proposition "Vera is <u>very very intelligent</u>" implies that

$$R_o(IQ) = \underline{\text{very}}\ \underline{\text{very}}\ \underline{\text{intelligent}} \tag{6.107}$$

Now, the proposition ""Vera is <u>very very intelligent</u>" is <u>very true</u>" implies that the grade of membership of Vera in the fuzzy set $R_o(IQ)$ is <u>very true</u> (see (6.6.)). Let $\mu_{\underline{\text{very true}}}$ denote the membership function of very true (see (6.2)) and let μ_{R_o} denote that of $R_o(IQ)$. Regarding μ_{R_o} as a relation from the range of IQ to [0,1], let $\mu_{R_o}^{-1}$ denote the inverse relation from [0,1] to the range of IQ. This relation, then, induces a fuzzy set $R_1(IQ)$ expressed by

$$R_1(IQ) = \mu_{R_o}^{-1}(\underline{\text{very}}\ \underline{\text{true}}) \tag{6.108}$$

which can be computed by the use of the extension principle in the form (3.80). The fuzzy set $R_1(IQ)$ represents the restriction on IQ induced by the assertion ""Vera is very very intelligent" is very true."

Continuing the same argument, the restriction on IQ induced by the assertion """"Vera is very very intelligent" is very true" is true" may be expressed as

$$R_2(IQ) = \mu_{R_1}^{-1}(\underline{\text{true}}) \tag{6.109}$$

where $\mu_{R_1}^{-1}$ denotes the relation inverse to μ_{R_1}, which is the membership function of $R_1(IQ)$ given by (6.108). In this way, we can compute the restriction induced by a nested sequence of assertions such as that exemplified by (6.106).

The basic idea behind the technique sketched above is that an assertion of the form ""u is A" is T," where A is a fuzzy predicate and T is a linguistic truth-value, modifies the restriction associated with A in accordance with the expression

$$A' = \mu_A^{-1}(T)$$

where μ_A^{-1} is the inverse of the membership function of A, and A' is the restriction induced by the assertion "u is A" is T."

7. Linguistic Probabilities and Averages Over Fuzzy Sets

In the classical approach to probability theory, an event, A, is defined as a member of a σ-field, $\mathcal{A}$, of subsets of a sample space Ω. Thus, if P is a normed measure over a measurable space $(\Omega,\mathcal{A})$, the probability of A is defined as P(A), the measure of A, and is a number in the interval [0,1].

There are many real-world problems in which one or more of the basic assumptions which are implicit in the above definition are violated. First, the event, A, is frequently ill-defined, as in the question, "What is the probability that it will be a warm day tomorrow?" In this instance, the event warm day is a fuzzy event in the sense that there is no sharp dividing line between its occurrence and non-occurrence. As shown in [48], such an event may be characterized as a fuzzy subset, A, of the sample space Ω, with μ_A, the membership function of A, being a measurable function.

Second, even if A is a well-defined nonfuzzy event, its probability, P(A), may be ill-defined. For example, in response to the question, "What is the probability that the Dow Jones average of stock prices will be higher in a month from now," it would be patently unreasonable to give an unequivocal numerical answer, e.g., 0.7. In this instance, a vague response like "quite probable," would be much more commensurate with our lack of understanding of the dynamics of stock prices, and hence a more realistic - if less precise - characterization of the probability in question.

The limitations imposed by the assumption that A is well-defined may be removed, at least in part, by allowing A to be a fuzzy event, as was done in [48]. Another and perhaps more important step that can be taken to widen the applicability of probability theory to ill-defined problems, is to allow P to be a linguistic variable in the sense defined in Section 6. In what follows, we shall outline a way in which this can be done and explore some of the elementary consequences of allowing P to be a linguistic variable.

Linguistic Probabilities

To simplify our exposition, we shall assume that the object of our concern is a variable, X, whose universe of discourse, U, is a finite set

$$U = u_1 + u_2 + \dots + u_n \tag{7.1}$$

Furthermore, we assume that the restriction imposed by X coincides with U. Thus, any point in U can be assigned as a value to X.

With each u_i, $i = 1, \dots, n$, we associate a linguistic probability, $\mathcal{P}_i$, which is a Boolean linguistic variable in the sense of Definition 5.9, with p_i, $0 \leq p_i \leq 1$, representing the base variable for $\mathcal{P}_i$. For concreteness, we shall assume that V, the universe of discourse associated with $\mathcal{P}_i$, is either the unit interval [0,1] or the finite set

$$V = 0 + 0.1 + \dots + 0.9 + 1 \tag{7.2}$$

Using $\mathcal{P}$ as a generic name for the $\mathcal{P}_i$, the term-set for $\mathcal{P}$ will typically be the following.

$T(\mathcal{P})$ = likely + not likely + unlikely + very likely + more or less likely

+ very unlikely + ...

+ probable + improbable + very probable + ...

neither very probable nor very improbable + ...

+ close to 0 + close to 0.1 + ... + close to 1 + ...

+ very close to 0 + very close to 0.1 + ... (7.3)

in which likely, probable and close to play the role of primary terms.

The shape of the membership function of likely will be assumed to be like that of true (see (6.2)), with not likely and unlikely defined by

$$\mu_{\underline{\text{not likely}}}(p) = 1 - \mu_{\underline{\text{likely}}}(p) \tag{7.4}$$

and

$$\mu_{\underline{\text{unlikely}}}(p) = \mu_{\underline{\text{likely}}}(1 - p) \tag{7.5}$$

where p is a generic name for the p_i.

Example 7.1 A graphic example of the meaning attached to the terms likely, not likely, very likely and unlikely is shown in Fig. 7.1. In numerical terms, if the primary term likely is defined as

$$\underline{\text{likely}} = 0.5/0.6 + 0.7/0.7 + 0.9/0.8 + 1/0.9 + 1/1 \tag{7.6}$$

then

$$\underline{\text{not likely}} = 1/(0 + 0.1 + 0.2 + 0.3 + 0.4 + 0.5) + 0.5/0.6 + 0.3/0.7 + 0.1/0.8 \tag{7.7}$$

$$\underline{\text{unlikely}} = 1/0 + 1/0.1 + 0.9/0.2 + 0.7/0.3 + 0.5/0.4 \tag{7.8}$$

and

$$\underline{\text{very likely}} = 0.25/0.6 + 0.49/0.7 + 0.81/0.8 + 1/0.9 + 1/1 \tag{7.9}$$

The term probable will be assumed to be more or less synonymous with likely. The term close to α, where α is a point in $[0,1]$, will be abbreviated as $\underset{\sim}{\alpha}$ or, alternatively, as "α",[1] suggesting that α is a "best example" of the fuzzy set "α". In this sense, then,

$$\underline{\text{likely}} \triangleq \underline{\text{close}}\ \underline{\text{to}}\ 1 \triangleq \text{"1"} \tag{7.10}$$

$$\underline{\text{unlikely}} \triangleq \underline{\text{close}}\ \underline{\text{to}}\ 0 \triangleq \text{"0"} \tag{7.11}$$

and $$\underline{\text{close}}\ \underline{\text{to}}\ 0.8 \triangleq \text{"0.8"} = 0.6/0.7 + 1/0.8 + 0.6/0.9 \tag{7.12}$$

from which it follows that

$$\begin{aligned}\underline{\text{very}}\ \underline{\text{close}}\ \underline{\text{to}}\ 0.8 &= \text{very "0.8"}\\ &= (\text{"0.8"})^2 \quad \text{(in the sense of (5.38))}\\ &= 0.36/0.7 + 1/0.8 + 0.36/0.9\end{aligned}$$

A particular term in $T(\mathcal{P})$ will be denoted by T_j or T_{ji}, in case a double subscript notation is needed. Thus, if T_4 = very likely then T_{43} would indicate that very likely is assigned as a value to the linguistic variable $\mathcal{P}_3$.

The n-ary linguistic variable $(\mathcal{P}_1, \ldots, \mathcal{P}_n)$ constitutes a linguistic probability assignment list associated with X. A variable X which is associated with a linguistic probability assignment list will be referred to as a linguistic random variable. By analogy with linguistic truth-value distributions (see (6.74)), a collection of probability assignment

---

[1] The symbol "α" will be employed in place of $\underset{\sim}{\alpha}$ when the constraints imposed by type-setting dictate its use.

lists will be referred to as a <u>linguistic probability distribution</u>.

The assignment of a probability-value T_j to P_i may be expressed as

$$P_i = T_j \tag{7.13}$$

where P_i is used in a dual role as a generic name for the fuzzy variables which comprise $\mathcal{P}_i$. For example, we may write

$$\begin{aligned} P_3 &= T_4 \\ &= \underline{\text{very likely}} \end{aligned} \tag{7.14}$$

in which case <u>very likely</u> will be identified as T_{43} (i.e., T_4 assigned to P_3).

An important characteristic of the linguistic probabilities $P_1, \ldots, P_n$ is that they are β-interactive in the sense of Definition 6.8. The interaction between the P_i is a consequence of the constraint ($+ \triangleq$ arithmetic sum)

$$p_1 + p_2 + \ldots + p_n = 1 \tag{7.15}$$

in which the p_i are the base variables (i.e., numerical probabilities) associated with the P_i.

More concretely, let $R(p_1 + \ldots + p_n = 1)$ denote the nonfuzzy n-ary relation in $[0,1] \times \ldots \times [0,1]$ representing (7.15). Furthermore, let $R(P_i)$ denote the restriction on the values of p_i. Then, the restriction imposed by the n-ary fuzzy variable $(P_1, \ldots, P_n)$ may be expressed as

$$R(P_1, \dots, P_n) = R(P_1) \times \dots \times R(P_n) \cap R(p_1 + \dots + p_n = 1) \qquad (7.16)$$

which implies that, apart from the constraint imposed by (7.15), the fuzzy variables $P_1, \dots, P_n$ are noninteractive.

<u>Example 7.2</u> Suppose that

$$P_1 = \underline{likely} \qquad (7.17)$$
$$= 0.5/0.8 + 0.8/0.9 + 1/1$$

and

$$P_2 = \underline{unlikely} \qquad (7.18)$$
$$= 1/0 + 0.8/0.1 + 0.5/0.2$$

Then

$$\begin{aligned} R(P_1) \times R(P_2) &= \underline{likely} \times \underline{unlikely} &(7.19)\\ &= (0.5/0.8 + 0.8/0.9 + 1/1)\times(1/0 + 0.8/0.1 + 0.5/0.2)\\ &= 0.5/(0.8,0) + 0.8/(0.9,0) + 1/(1,0)\\ &\quad + 0.5/(0.8,0.1) + 0.8/(0.9,0.1) + 0.8/(1,0.1)\\ &\quad + 0.5/(0.8,0.2) + 0.5/(0.9,0.2) + 0.5/(1,0.2) \end{aligned}$$

As for $R(p_1 + \dots + p_n = 1)$, it can be expressed as

$$R(p_1 + p_2 = 1) = \sum_k 1/(k,1-k), \qquad k = 0,0.1, \dots, 0.9,1 \quad (7.20)$$

and forming the intersection of (7.19) and (7.20), we obtain

$$R(P_1,P_2) = 1/(1,0) + 0.8/(0.9,1) + 0.5/(0.8,0.2) \qquad (7.21)$$

as the expression for the restriction imposed by (P_1,P_2). Obviously, $R(P_1,P_2)$ is comprised of those terms in $R(P_1) \times R(P_2)$ which satisfy the constraint (7.15).

Remark 7.3 It should be observed that $R(P_1,P_2)$ as expressed by (7.21) is a normal restriction (see (3.23)). This will be the case, more generally, when the P_i are of the form

$$P_i = "q_i" , \qquad i = 1, \ldots , n \tag{7.22}$$

and $q_1 + \ldots + q_n = 1$. Note that in Example 7.2, we have

$$P_1 = "1" \tag{7.23}$$

$$P_2 = "0" \tag{7.24}$$

and

$$1 + 0 = 1 \tag{7.25}$$

Computation With Linguistic Probabilities

In many of the applications of probability theory, e.g., in the calculation of means, variances, etc., one encounters linear combinations of the form ($+ \triangleq$ arithmetic sum)

$$z = a_1p_1 + \ldots + a_np_n \tag{7.26}$$

where the a_i are real numbers and the p_i are probability-values in [0,1]. Computation of the value of z given the a_i and the p_i presents no difficulties when the p_i are points in [0,1]. It becomes, however, a non-trivial problem when the probabilities in question are linguistic in nature, that is, when

$$Z = a_1P_1 + \dots + a_nP_n \tag{7.27}$$

where the P_i represent linguistic probabilities with names such as <u>likely</u>, <u>unlikely</u>, <u>very likely</u>, <u>close to</u> α, etc. Correspondingly, Z is not a real number - as it is in (7.26) - but a fuzzy subset of the real line $W \triangleq (-\infty, \infty)$, with the membership function of Z being a function of those of the P_i.

Assuming that the fuzzy variables $P_1, \dots, P_n$ are noninteractive (apart from the constraint expressed by (7.15)), the restriction imposed by $(P_1, \dots, P_n)$ assumes the form (see (7.16))

$$R(P_1, \dots, P_n) = R(P_1) \times \dots \times R(P_n) \cap R(p_1 + \dots + p_n = 1) \tag{7.28}$$

Let $\mu(p_1, \dots, p_n)$ be the membership function of $R(P_1, \dots, P_n)$, and let $\mu_i(p_i)$ be that of $R(P_i)$, $i = 1, \dots, n$. Then, by applying the extension principle (3.90) to (7.26), we can express Z as a fuzzy set ($+ \triangleq$ arithmetic sum)

$$Z = \int_W \mu(p_1, \dots, p_n)/(a_1p_1 + \dots + a_np_n) \tag{7.29}$$

which in view of (7.28) may be written as

$$Z = \int_W \mu_1(p_1) \wedge \dots \wedge \mu_n(p_n)/(a_1p_1 + \dots + a_np_n) \tag{7.30}$$

with the understanding that the p_i in (7.30) are subject to the constraint

$$p_1 + \dots + p_n = 1 \tag{7.31}$$

In this way, we can express a linear combination of linguistic probability-values as a fuzzy subset of the real line.

The expression for Z may be cast into other forms which may be more convenient for computational purposes. Thus, let $\mu(z)$ denote the membership function of Z, with $z \in W$. Then, (7.30) implies that

$$\mu(z) = \vee_{p_1,\ldots,\, p_n} \mu_1(p_1) \wedge \ldots \wedge \mu_n(p_n) \tag{7.32}$$

subject to the constraints

$$z = a_1 p_1 + \ldots + a_n p_n \tag{7.33}$$

$$p_1 + \ldots + p_n = 1 \tag{7.34}$$

In this form, the computation of Z reduces to the solution of a nonlinear programming problem with linear constraints. In more explicit terms, this problem may be expressed as: Maximize z subject to the constraints ($+ \triangleq$ arithmetic sum)

$$\mu_1(p_1) \geq z \tag{7.35}$$

$$\cdots\cdots\cdots$$

$$\mu_n(p_n) \geq z$$

$$z = a_1 p_1 + \ldots + a_n p_n$$

$$p_1 + \ldots + p_n = 1$$

Example 7.4 As a very simple illustration, assume that

$$P_1 = \underline{likely} \tag{7.36}$$

and

$$P_2 = \underline{\text{unlikely}} \tag{7.37}$$

where

$$\underline{\text{likely}} = \int_0^1 \mu_{\underline{\text{likely}}}(p)/p \tag{7.38}$$

and

$$\underline{\text{unlikely}} = \neg\, \underline{\text{likely}} \tag{7.39}$$

Thus (see (7.5))

$$\mu_{\underline{\text{unlikely}}}(p) = \mu_{\underline{\text{likely}}}(1-p)\ , \quad 0 \le p \le 1 \tag{7.40}$$

Suppose that we wish to compute the expectation ($+ \triangleq$ arithmetic sum)

$$Z = a_1\, \underline{\text{likely}} + a_2\, \underline{\text{unlikely}} \tag{7.41}$$

Using (7.32), we have

$$\mu(z) = \vee_{p_1,p_2}\, \mu_{\underline{\text{likely}}}(p_1) \wedge \mu_{\underline{\text{unlikely}}}(p_2) \tag{7.42}$$

subject to the constraints

$$z = a_1p_1 + a_2p_2 \tag{7.43}$$

$$p_1 + p_2 = 1$$

Now in view of (7.40), if $p_1 + p_2 = 1$ then

$$\mu_{\underline{\text{likely}}}(p_1) = \mu_{\underline{\text{unlikely}}}(p_2) \tag{7.44}$$

and hence (7.42) reduces to

$$\mu(z) = \mu_{\underline{\text{likely}}}(p_1) \tag{7.45}$$

$$z = a_1p_1 + a_2(1 - p_1)$$

or, more explicitly,

$$\mu(z) = \mu_{\underline{likely}}\left(\frac{z - a_2}{a_1 - a_2}\right) \tag{7.46}$$

This result implies that the fuzziness in our knowledge of the probability p_1 induces a corresponding fuzziness in the expectation of (see Fig. 7.2)

$$z = a_1p_1 + a_2p_2 \;.$$

If the universe of probability-values is assumed to be $V = 0 + 0.1 + \ldots + 0.9 + 1$, then the expression for Z can be obtained more directly by using the extension principle in the form (3.97). As an illustration, assume that

$$P_1 = \text{"0.3"} = 0.8/0.2 + 1/0.3 + 0.6/0.4 \tag{7.47}$$

$$P_2 = \text{"0.7"} = 0.8/0.6 + 1/0.7 + 0.6/0.8 \tag{7.48}$$

and ($\oplus \triangleq$ arithmetic sum)

$$Z = a_1P_1 \oplus a_2P_2 \tag{7.49}$$

where the symbol $\oplus$ is used to avoid confusion with the union.

On substituting (7.47) and (7.48) in (7.49), we obtain

$$Z = a_1(0.8/0.2 + 1/0.3 + 0.6/0.4) \oplus a_2(0.8/0.6 + 1/0.7 + 0.6/0.8) \tag{7.50}$$

$$= (0.8/0.2a_1 + 1/0.3a_1 + 0.6/0.4a_1) \oplus (0.8/0.6a_2 + 1/0.7a_2 + 0.6/0.8a_2)$$

In expanding the right-hand member of (7.50), we have to take into account the constraint $p_1 + p_2 = 1$, which means that a term of the form

$$\mu_1/p_1a_1 \oplus \mu_2/p_2a_2 \tag{7.51}$$

evaluates to

$$\begin{aligned} \mu_1/p_1a_1 \oplus \mu_2/p_2a_2 &= \mu_1 \wedge \mu_2/(p_1a_1 \oplus p_2a_2) \quad \text{if } p_1 + p_2 = 1 \\ &= 0 \quad \text{otherwise.} \end{aligned} \tag{7.52}$$

In this way, we obtain

$$Z = 1/(0.3a_1 \oplus 0.7a_2) + 0.6/(0.2a_1 \oplus 0.8a_2) + 0.6/(0.4a_1 \oplus 0.6a_2) \tag{7.53}$$

which expresses Z as a fuzzy subset of the real line $W = (-\infty, \infty)$.

Averages Over Fuzzy Sets

Our point of departure in the foregoing discussion was the assumption that with each point u_i of a finite[2] universe of discourse U is associated a linguistic probability-value P_i which is a component of a linguistic probability distribution $(\mathcal{P}_1, \ldots, \mathcal{P}_n)$.

In this context, a fuzzy subset, A, of U plays the role of a fuzzy event. Let $\mu_A(u_i)$ be the grade of membership of u_i in A. Then, if the P_i are conventional numerical probabilities, p_i, $0 \leq p_i \leq 1$, then the probability of A, $P(A)$, is defined as (see [48]) ($+ \triangleq$ arithmetic sum)

$$P(A) = \mu_A(u_1)p_1 + \ldots + \mu_A(u_n)p_n \tag{7.54}$$

---

[2] The assumption that U is a finite set is made solely for the purpose of simplifying our exposition. More generally, U can be a countable set or a continuum.

It is natural to extend this definition to linguistic probabilities by defining the linguistic probability[3] of A as

$$P(A) = \mu_A(u_1)\, P_1 + \ldots + \mu_A(u_n)\, P_n \tag{7.55}$$

with the understanding that the right-hand member of (7.55) is a linear form in the sense of (7.27). In connection with (7.55), it should be noted that the constraint

$$p_1 + \ldots + p_n = 1 \tag{7.56}$$

on the underlying probabilities, together with the fact that

$$0 \leq \mu_A(u_i) \leq 1 \quad , \quad i = 1, \ldots , n$$

insures that $P(A)$ is a fuzzy subset of $[0,1]$.

<u>Example 7.5</u> As a very simple illustration, assume that

$$U = a + b + c \tag{7.57}$$

$$A = 0.4a + b + 0.8c \tag{7.58}$$

$$P_a = \text{"}0.3\text{"} = 0.6/0.2 + 1/0.3 + 0.6/0.4 \tag{7.59}$$

$$P_b = \text{"}0.6\text{"} = 0.6/0.5 + 1/0.6 + 0.6/0.7 \tag{7.60}$$

$$P_c = \text{"}0.1\text{"} = 0.6/0 + 1/0.1 + 0.6/0.2 \tag{7.61}$$

Then ($\oplus \triangleq$ arithmetic sum)

$$\begin{aligned} P(A) = \; & 0.4(0.6/0.2 + 1/0.3 + 0.6/0.4) \oplus (0.6/0.5 + 1/0.6 + 0.6/0.7) \\ & \oplus\, 0.8(0.6/0 + 1/0.1 + 0.6/0.2) \end{aligned} \tag{7.62}$$

---

[3] It should be noted that the computation of the right-hand member of (7.55) defines $P(A)$ as a fuzzy subset of $[0,1]$. In general, a linguistic approximation would be needed to express $P(A)$ as a linguistic probability-value.

subject to the constraint

$$p_1 + p_2 + p_3 = 1 \tag{7.63}$$

Picking those terms in (7.62) which satisfy (7.63), we obtain

$$\begin{aligned} P(A) = {} & 0.6/(0.4 \times 0.2 \oplus 0.6 \oplus 0.8 \times 0.2) \\ & + 0.6/(0.4 \times 0.2 \oplus 0.6 \oplus 0.8 \times 0.1) \\ & + 0.6/(0.4 \times 0.3 \oplus 0.5 \oplus 0.8 \times 0.2) \\ & + 1/(0.4 \times 0.3 \oplus 0.6 \oplus 0.8 \times 0.1) \\ & + 0.6/(0.4 \times 0.3 \oplus 0.7) \\ & + 0.6/(0.4 \times 0.4 \oplus 0.5 \oplus 0.8 \times 0.1) \\ & + 0.6/(0.4 \times 0.4 \oplus 0.6) \end{aligned} \tag{7.64}$$

which reduces to

$$P(A) = 0.6/(0.84 + 0.76 + 0.78 + 0.82 + 0.74) + 1/0.8 \tag{7.65}$$

and which may be roughly approximated as

$$P(A) = \text{"0.8"} \tag{7.66}$$

The linguistic probability of a fuzzy event as expressed by (7.55), may be viewed as a particular instance of a more general concept, namely, the _linguistic average_ or, eqivalently, the _linguistic expectation,_ of a function (defined on U) over a fuzzy subset of U. More specifically, let f be a real-valued function defined on U; let A be a fuzzy subset of U; and let $P_1, \ldots, P_n$ be the linguistic probabilities associated with $u_1, \ldots, u_n$, respectively. Then, the _linguistic average of_ f _over_ A is denoted by Av(f;A) and is defined by ($+ \triangleq$ arithmetic sum)

$$Av(f;A) = f(u_1)\mu_A(u_1)\ P_1 + \dots + f(u_n)\mu_A(u_n)\ P_n \tag{7.67}$$

A concrete example of (7.67) is the following. Assume that individuals named $u_1, \dots, u_n$ are chosen with linguistic probabilities $P_1, \dots, P_n$, with P_i being a restriction on p_i, $i = 1, \dots, n$. Suppose that u_i is fined an amount $f(u_i)$, which is scaled down in proportion to the grade of membership of u_i in a class A. Then, the linguistic average (expected) amount of the fine will be expressed by (7.67).

Comment 7.6 Note that (7.67) is basically a linear combination of the form (7.27), with

$$a_i = f(u_i)\mu_A(u_i) \tag{7.68}$$

Thus, to evaluate (7.67), we can employ the technique described earlier for the computation of linear forms in linguistic probabilities. In particular, it should be noted that, in the special case where $f(u_i) = 1$, the right-hand member of (7.67) becomes

$$\mu_A(u_1)P_1 + \dots + \mu_A(u_n)P_n \tag{7.69}$$

and $Av(f;A)$ reduces to $P(A)$.

In addition to subsuming the expression for $P(A)$, the expression for $Av(f;A)$ subsumes as special cases other types of averages which occur in various applications. Among them there are two that may be regarded as degenerate forms of (7.67) and which are encountered in many problems of practical interest. In what follows, we shall dwell briefly on these averages and, for convenience in exposition, will state their definitions in the form of answers to questions.

Question 7.7 What is the number of elements in a given fuzzy set A? Clearly, this question is not well-posed, since in the case of a fuzzy set the dividing line between membership and nonmembership is not sharp. Nevertheless, the concept of the power of a fuzzy set [49], which is defined as

$$|A| \triangleq \sum_i \mu_A(u_i) \tag{7.70}$$

appears to be a natural generalization of that of the number of elements in A.

As an illustration of $|A|$, suppose that U is the universe of residents in a city, and A is the fuzzy set of the unemployed in that city. If $\mu_A(u_i)$ is interpreted as the grade of membership of an individual, u_i, in the class of the unemployed (e.g., $\mu_A(u_i) = 0.5$ if u_i is working half-time and is looking for a full-time job), then $|A|$ may be interpreted as the number of full-time equivalent unemployed.

Question 7.8 Suppose that f is a real-valued function defined on U. What is the average value of f over a fuzzy subset, A, of U?

Using the same notation as in (7.67), let Av(f;A) denote the average value of f over A. If A were nonfuzzy, Av(f;A) would be expressed by

$$\mathrm{Av}(f;A) = \frac{\sum_{u_i \in A} f(u_i)}{|A|} \tag{7.71}$$

where $\sum_{u_i \in A}$ is the summation over those u_i which are in A, and $|A|$ is the number of the u_i which are in A. To extend (7.71) to fuzzy sets, we note that (7.71) may be rewritten as

$$Av(f;A) = \frac{\sum_{u_i \in U} f(u_i)\mu_A(u_i)}{\sum_{u_i \in U} \mu_A(u_i)} \tag{7.72}$$

where μ_A is the characteristic function of A. Then, we adopt (7.72) as the definition of Av(f;A) for a <u>fuzzy</u> A by interpreting $\mu_A(u_i)$ as the grade of membership of u_i in A. In this way, we arrive at an expression for Av(f;A) which may be viewed as a special case of (7.67).

As an illustration of (7.72), suppose that U is the universe of residents in a city and A is the fuzzy subset of residents who are <u>young</u>. Furthermore, assume that $f(u_i)$ represents the income of u_i. Then, the average income of young residents in the city would be expressed by (7.72).

<u>Comment 7.9</u> Since the expression for $|A|$ is a linear form in the $\mu_A(u_i)$, the power of a fuzzy set of type 2 (see Definition 3.22) can readily be computed by employing the technique which we had used earlier to compute P(A). In the case of Av(f;A), however, we are dealing with a ratio of linear forms, and hence the computation of Av(f;A) for fuzzy sets of type 2 presents a more difficult problem.

In the foregoing discussion, our very limited objective was to indicate that the concept of a linguistic variable provides a basis for defining linguistic probabilities and, in conjunction with the extention principle, may be applied to the computation of linear forms in such probabilities. We shall not dwell further on this subject and, in what follows, will turn our attention to a basic rule of inference in fuzzy logic.

8. Compositional Rule of Inference and Approximate Reasoning

The basic rule of inference in traditional logic is the modus ponens, according to which we can infer the truth of a proposition B from the truth of A and the implication $A \Rightarrow B$. For example, if A is identified with "John is in a hospital," and B with "John is ill," then if it is true that "John is in a hospital," it is also true that "John is ill."

In much of human reasoning, however, modus ponens is employed in an approximate rather than exact form. Thus, typically, we know that A is true and that $A^* \Rightarrow B$, where A^* is, in some sense, an approximation to B. Then, from A and $A^* \Rightarrow B$ we may infer that B is approximately true.

In what follows, we shall outline a way of formalizing approximate reasoning based on the concepts introduced in the preceding sections. However, in a departure from traditional logic, our main tool will not be the modus ponens, but a so-called compositional rule of inference of which modus ponens forms a very special case.

Compositional Rule of Inference

The compositional rule of inference is merely a generalization of the following familiar procedure. Referring to Fig. 8.1, suppose that we have a curve $y = f(x)$ and are given $x = a$. Then from $y = f(x)$ and $x = a$, we can infer $y \triangleq b = f(a)$.

Next, let us generalize the above process by assuming that a is an interval and $f(x)$ is an interval-valued function such as shown in Fig. 8.2. In this instance, to find the interval $y \triangleq b$ which corresponds to the interval a, we first construct a cylindrical set, $\bar{a}$, with base a (see (3.58)) and find its intersection, I, with the interval-valued curve.

Then we project the intersection on the OY axis, yielding the desired y as the interval b.

Going one step further in our chain of generalizations, assume that A is a fuzzy subset of the OX axis and F is a fuzzy relation from OX to OY. Again, forming a cylindrical fuzzy set $\bar{A}$ with base A and intersecting it with the fuzzy relation F (see Fig. 8.3), we obtain a fuzzy set $\bar{A} \cap F$ which is the analog of the point of intersection I in Fig. 8.1. Then, projecting this set on OY, we obtain y as a fuzzy subset of OY. In this way, from $y = f(x)$ and $x \triangleq A$ = fuzzy subset of OX, we infer y as a fuzzy subset, B, of OY.

More specifically, let μ_A, $\mu_{\bar{A}}$, μ_F and μ_B denote the membership functions of A, $\bar{A}$, F and B, respectively. Then, by the definition of $\bar{A}$ (see (3.58))

$$\mu_{\bar{A}}(x,y) = \mu_A(x) \tag{8.1}$$

and consequently

$$\begin{aligned} \mu_{\bar{A} \cap F}(x,y) &= \mu_{\bar{A}}(x,y) \wedge \mu_F(x,y) \\ &= \mu_A(x) \wedge \mu_F(x,y) \end{aligned} \tag{8.2}$$

Projecting $\bar{A} \cap F$ on the OY axis, we obtain from (8.2) and (3.57)

$$\mu_B(y) = \vee_x \mu_A(x) \wedge \mu_F(x,y) \tag{8.3}$$

as the expression for the membership function of the projection (shadow) of $\bar{A} \cap F$ on OY. Comparing this expression with the definition of the composition of A and F (see (3.55)), we see that B may be represented as

$$B = A \circ F \tag{8.4}$$

where $\circ$ denotes the operation of composition. As stated in Sec. 3, this operation reduces to the max-min matrix product when A and F have finite supports.

<u>Example 8.1</u> Suppose that A and F are defined by

$$A = 0.2/1 + 1/2 + 0.3/3 \tag{8.5}$$

and

$$\begin{aligned} F = {} & 0.8/(1,1) + 0.9/(1,2) + 0.2/(1,3) \\ & + 0.6/(2,1) + 1/(2,2) + 0.4/(2,3) \\ & + 0.5/(3,1) + 0.8/(3,2) + 1/(3,3) \end{aligned} \tag{8.6}$$

Expressing A and F in terms of their relation matrices and forming the matrix product (8.4), we obtain

$$\overset{A}{\begin{bmatrix} 0.2 & 1 & 0.3 \end{bmatrix}} \circ \overset{F}{\begin{bmatrix} 0.8 & 0.9 & 0.2 \\ 0.6 & 1 & 0.4 \\ 0.5 & 0.8 & 1 \end{bmatrix}} = \overset{B}{\begin{bmatrix} 0.6 & 1 & 0.4 \end{bmatrix}} \tag{8.7}$$

The foregoing comments and examples serve to motivate the following rule of inference.

<u>Rule 8.2</u> Let U and V be two universes of discourse with base variables u and v, respectively. Let R(u), R(u,v) and R(v) denote restrictions on u, (u,v) and v, respectively, with the understanding that R(u), R(u,v) and R(v) are fuzzy relations in U, U × V and V. Let A and F denote particular fuzzy subsets of U and U × V. Then, the <u>compositional rule of inference</u> asserts that the solution of the <u>relational</u> assignment

equations

$$R(u) = A \tag{8.8}$$

and

$$R(u,v) = F \tag{8.9}$$

is given by

$$R(v) = A \circ F \tag{8.10}$$

where $A \circ F$ is the composition of A and F. In this sense, we can <u>infer</u> $R(v) = A \circ F$ from $R(u) = A$ and $R(u,v) = F$.

As a simple illustration of the use of this rule, assume that

$$U = V = 1 + 2 + 3 + 4 \tag{8.11}$$

$$A = \underline{\text{small}} = 1/1 + 0.6/2 + 0.2/3 \tag{8.12}$$

and

$$\begin{aligned} F &= \underline{\text{approximately}}\ \underline{\text{equal}} \\ &= 1/(1,1) + 1/(2,2) + 1/(3,3) + 1/(4,4) \\ &\quad + 0.5/((1,2) + (2,1) + (2,3) + (3,2) + (3,4) + (4,3)) \end{aligned} \tag{8.13}$$

In other words, A is unary fuzzy relation in U named <u>small</u> and F is a binary fuzzy relation in $U \times V$ named <u>approximately</u> <u>equal</u>.

The relational assignment equations in this case read

$$R(u) = \underline{\text{small}} \tag{8.14}$$

$$R(u,v) = \underline{\text{approximately}}\ \underline{\text{equal}} \tag{8.15}$$

and hence

$$R(v) = \underline{\text{small}} \circ \underline{\text{approximately}}\ \underline{\text{equal}} \tag{8.16}$$

$$= [1 \quad 0.6 \quad 0.2 \quad 0] \circ \begin{bmatrix} 1 & 0.5 & 0 & 0 \\ 0.5 & 1 & 0.5 & 0 \\ 0 & 0.5 & 1 & 0.5 \\ 0 & 0 & 0.5 & 1 \end{bmatrix}$$

$$= [1 \quad 0.6 \quad 0.5 \quad 0.2]$$

which may be approximated by the linguistic term

$$R(v) = \underline{\text{more}}\ \underline{\text{or}}\ \underline{\text{less}}\ \underline{\text{small}} \tag{8.17}$$

if <u>more</u> <u>or</u> <u>less</u> is defined as a fuzzifier (see (3.48)), with

$$K(1) = 1/1 + 0.7/2 \tag{8.18}$$

$$K(2) = 1/2 + 0.7/3$$

$$K(3) = 1/3 + 0.7/4$$

$$K(4) = 1/4$$

Note that the application of this fuzzifier to R(u) yields

$$[1 \quad 0.7 \quad 0.42 \quad 0.14] \tag{8.19}$$

as an approximation to [1 0.6 0.5 0.2].

In summary, then, by using the compositional rule of inference, we have infered from R(u) = <u>small</u> and R(u,v) = <u>approximately</u> <u>equal</u>

$$R(v) = [1 \quad 0.6 \quad 0.5 \quad 0.2] \quad \text{exactly} \tag{8.20}$$

and

$$R(v) = \underline{\text{more}}\ \underline{\text{or}}\ \underline{\text{less}}\ \underline{\text{small}} \quad \text{as a linguistic approximation} \tag{8.21}$$

Stated in English, this approximate inference may be expressed as

u is small premiss (8.22)

u and v are approximately equal	premiss
v is more or less small	approximate conclusion

The general idea behind the method sketched above is the following. Each fact or a premiss is translated into a relational assignment equation involving one or more restrictions on the base variables. These equations are solved for the desired restrictions by the use of the composition of fuzzy relations. The solutions to the equations, then, represent deductions from the given set of premisses.

Modus Ponens as a Special Case of the Compositional Rule of Inference

As we shall see in the sequel, modus ponens may be viewed as a special case of the compositional rule of inference. To establish this connection, we shall first extend the notion of material implication from propositional variables to fuzzy sets.

In traditional logic, the material implication ⇒ is defined as a logical connective for propositional variables. Thus, if A and B are propositional variables, the truth table for A ⇒ B or, equivalently, IF A THEN B, is defined to be (see Table 6.8)

A \ B	T	F
T	T	F
F	T	T

Table 8.4

In much of human discourse, however, the expression IF A THEN B is used in situations in which A and B are fuzzy sets (or fuzzy predicates) rather than propositional variables. For example, in the case of the statement IF John is ill THEN John is cranky, which may be abbreviated as ill $\Rightarrow$ cranky, ill and cranky are, in effect, names of fuzzy sets. The same is true of the statement IF apple is red THEN apple is ripe, where red and ripe play the role of fuzzy sets.

To extend the notion of material implication to fuzzy sets, let U and V be two possibly different universes of discourse and let A, B and C be fuzzy subsets of U, V and V respectively. First, we shall define the meaning of the expression IF A THEN B ELSE C and then will define IF A THEN B as a special case of IF A THEN B ELSE C.

Definition 8.3. The expression IF A THEN B ELSE C is a binary fuzzy relation in $U \times V$ defined by

$$\text{IF A THEN B ELSE C} = A \times B + \neg A \times C \qquad (8.23)$$

That is, if A, B and C are unary fuzzy relations in U, V and V, then IF A THEN B ELSE C is a binary fuzzy relation in $U \times V$ which is the union of the cartesian product of A and B (see (3.45)) and the cartesian product of the negation of A and C.

Now IF A THEN B may be viewed as a special case of IF A THEN B ELSE C which results when C is allowed to be the entire universe V. Thus

$$\begin{aligned} \text{IF A THEN B} &\triangleq \text{IF A THEN B ELSE V} \qquad (8.24) \\ &= A \times B + \neg A \times V \end{aligned}$$

In effect, this amounts to interpreting IF A THEN B as IF A THEN B ELSE don't care.[1]

It is helpful to observe that in terms of the relation matrices of

[1]It is conceivable that a better definition for $A \Rightarrow B$ could be formulated by an explicit use in (8.24) of the truth-value unknown (see (6.52)).

A, B and C, (8.23) may be expressed as the sum of dyadic products involving A and B (and $\neg$A and C) as column and row matrices, respectively. Thus,

$$\text{IF A THEN B ELSE C} = \begin{bmatrix} \\ A \end{bmatrix}[\ B\ ] + \begin{bmatrix} \\ \neg A \end{bmatrix}[\ C\ ] \qquad (8.25)$$

<u>Example 8.4</u> As a simple illustration of (8.23) and (8.24), assume that

$$U = V = 1 + 2 + 3 \qquad (8.26)$$

$$A = \underline{\text{small}} = 1/1 + 0.4/2 \qquad (8.27)$$

$$B = \underline{\text{large}} = 0.4/2 + 1/3 \qquad (8.28)$$

$$C = \underline{\text{not}}\ \underline{\text{large}} = 1/1 + 0.6/2 \qquad (8.29)$$

Then

$$\begin{aligned}\text{IF A THEN B ELSE C} &= (1/1 + 0.4/2)\times(0.4/2 + 1/3)+(0.6/2 + 1/3)\times(1/1 \\ &\quad + 0.6/2) \qquad (8.30)\\ &= 0.4/(1,2) + 1/(1,3) + 0.6/(2,1) + 0.6/(2,2) \\ &\quad + 0.4/(2,3) + 1/(3,1) + 0.6/(3,2)\end{aligned}$$

which, represented as a relation matrix, reads

$$\text{IF A THEN B ELSE C} = \begin{bmatrix} 0 & 0.4 & 1 \\ 0.6 & 0.6 & 0.4 \\ 1 & 0.6 & 0 \end{bmatrix} \qquad (8.31)$$

Similarly

$$\begin{aligned}\text{IF A THEN B} &= (1/1 + 0.4/2)\times(0.4/2 + 1/3)+(0.6/2 + 1/3)\times(1/1 + 1/2 + 1/3) \\ &= 0.4/(1,2) + 1/(1,3) + 0.6/(2,1) + 0.6/(2,2) \\ &\quad + 0.6/(2,3) + 1/(3,1) + 1/(3,2) + 1/(3,3)\end{aligned}$$

or equivalently

$$\text{IF A THEN B} = \begin{bmatrix} 0 & 0.4 & 1 \\ 0.6 & 0.6 & 0.6 \\ 1 & 1 & 1 \end{bmatrix} \tag{8.32}$$

Comment 8.5 It should be noted that in defining IF A THEN B by (8.24) we are tacitly assuming that A and B are noninteractive in the sense that there is no joint constraint involving the base variables u and v. This would not be the case in the nonfuzzy statement IF $u \in A$ THEN $u \in B$, which may be expressed as IF $u \in A$ THEN $v \in B$, subject to the constraint $u = v$. Denoting this constraint by $R(u = v)$, the relation representing the statement in question would b

$$\text{IF } u \in A \text{ THEN } u \in B \triangleq (A \times B + \neg A \times V) \cap (R(u = v)) \tag{3.33}$$

Remark 8.6 In defining $A \Rightarrow B$, we assumed that IF A THEN B is a special case of IF A THEN B ELSE C resulting from setting $C = V$. If we set C equal to θ (empty set) rather than V, the right-hand member of (8.23) reduces to the cartesian product $A \times B$ – which may be interpreted as A COUPLED WITH B (rather than A CAUSES B.) Thus, by definition

$$\text{A COUPLED WITH B} \triangleq A \times B \tag{8.34}$$

and hence

$$A \Rightarrow B \triangleq \text{A COUPLED WITH B plus} \neg \text{A COUPLED WITH V} \tag{8.35}$$

More generally, an expression of the form

$$A_1 \times B_1 + \ldots + A_n \times B_n \tag{8.36}$$

would be expressed in words as

$$A_1 \text{ COUPLED WITH } B_1 \text{ plus } \dots \text{ plus } A_n \text{ COUPLED WITH } B_n \tag{8.37}$$

It should be noted that expressions such as (8.37) may be employed to represent a fuzzy graph as a union of fuzzy points (see Fig. 8.4). For example, a fuzzy graph G may be represented as

$$G = \text{"}u_1\text{"} \times \text{"}v_1\text{"} + \text{"}u_2\text{"} \times \text{"}v_2\text{"} + \dots + \text{"}u_n\text{"} \times \text{"}v_n\text{"} \tag{8.38}$$

where the u_i and v_i are points in U and V, respectively, and "u_i" and "v_i," $i = 1,\dots,n$, represent fuzzy sets named <u>close</u> <u>to</u> u_i and <u>close</u> <u>to</u> v_i (see (7.12)).

<u>Comment 8.7</u> The connection between (8.24) and the conventional definition of material implication becomes clearer by noting that

$$\neg A \times B \subset \neg A \times V \tag{8.39}$$

and hence that (8.24) may be rewritten as

$$\begin{aligned} \text{IF } A \text{ THEN } B &= A \times B + \neg A \times B + \neg A \times V \\ &= (A + \neg A) \times B + \neg A \times V \end{aligned} \tag{8.40}$$

Now, if A is a nonfuzzy subset of U, then

$$A + \neg A = U \tag{8.41}$$

and hence IF A THEN B reduces to

$$\text{IF } A \text{ THEN } B = U \times B + \neg A \times V \tag{8.42}$$

which is similar in form to the familiar expression for $A \Rightarrow B$ in the case of propositional variables, namely

$$A \Rightarrow B \equiv \neg A \vee B \tag{8.43}$$

Turning to the connection between <u>modus ponens</u> and the compositional rule of inference, we first define a <u>generalized modus ponens</u> as follows.

<u>Definition 8.8</u> Let A_1, A_2 and B be fuzzy subsets of U, U and V, respectively. Assume that A_1 is assigned to the restriction R(u), and the relation $A_2 \Rightarrow B$ (defined by (3.24)) is assigned to the restriction R(u,v). Thus

$$R(u) = A_1 \tag{8.44}$$

$$R(u,v) = A_2 \Rightarrow B \tag{8.45}$$

As was shown earlier, these relational assignment equations may be solved for the restriction on v, yielding

$$R(v) = A_1 \circ (A_2 \Rightarrow B) \tag{8.46}$$

An expression for this conclusion in the form

A_1	premiss	(8.47)
$A_2 \Rightarrow B$	implication	(8.48)
$A_1 \circ (A_2 \Rightarrow B)$	conclusion	(8.49)

constitutes the statement of the <u>generalized modus ponens</u>.[2]

[2] The generalized <u>modus ponens</u> as defined here is unrelated to probabilistic rules of inference. A discussion of such rules and related issues may be found in [50].

<u>Comment 8.9</u> The above statement differs from the traditional <u>modus ponens</u> in two respects: First, A_1, A_2 and B are allowed to be fuzzy sets, and second, A_1 need not be identical with A_2. To check on what happens when $A_1 = A_2 = A$ and A is nonfuzzy, we substitute the expression for $A_2 \Rightarrow B$ in (8.46), yielding

$$A \circ (A \Rightarrow B) = A \circ (A \times B + \neg A \times V) \qquad (8.50)$$

$$= A_r\, A_c\, B_r + A_r (\neg\, A_c) V_r$$

where r and c stand for <u>row</u> and <u>column</u>, respectively; A_r and A_c denote the relation matrices for A expressed as a row matrix and a column matrix, respectively; and the matrix product is understood to be taken in the max-min sense.

Now, since A is nonfuzzy

$$A_r (\neg\, A_c) = o \qquad (8.51)$$

and so long as A is normal (see (3.23))

$$A_r A_c = 1 \qquad (8.52)$$

Consequently

$$A \circ (A \Rightarrow B) = B \qquad (8.53)$$

which agrees with the conclusion yielded by <u>modus</u> <u>ponens</u>.

<u>Example 8.10</u> As a simple illustration of (8.49), assume that

$$U = V = 1 + 2 + 3 \qquad (8.54)$$

$$A_2 = \underline{\text{small}} = 1/1 + 0.4/2 \qquad (8.55)$$

$$A_1 = \underline{\text{more}}\ \underline{\text{or}}\ \underline{\text{less}}\ \underline{\text{small}} = 1/1 + 0.4/2 + 0.2/3 \tag{8.56}$$

and

$$B = \underline{\text{large}} = 0.4/2 + 1/3 \tag{8.57}$$

Then

$$\underline{\text{small}} \circ \underline{\text{large}} = \begin{bmatrix} 0 & 0.4 & 1 \\ 0.6 & 0.6 & 0.6 \\ 1 & 1 & 1 \end{bmatrix} \tag{8.58}$$

and

$$\underline{\text{more}}\ \underline{\text{or}}\ \underline{\text{less}}\ \underline{\text{small}} \circ (\underline{\text{small}} \Rightarrow \underline{\text{large}}) = [1 \quad 0.4 \quad 0.2] \circ \begin{bmatrix} 0 & 0.4 & 1 \\ 0.6 & 0.6 & 0.6 \\ 1 & 1 & 1 \end{bmatrix} \tag{8.59}$$

$$= [0.4 \quad 0.4 \quad 1]$$

which may be roughly approximated as more or less large. Thus, in the case under consideration, the generalized modus ponens yields

u is more or less small	premiss	(8.60)
IF u is small THEN v is large	implication	
———————————————		
v is more or less large	approximate conclusion	

Comment 8.11 Because of the way in which $A \Rightarrow B$ is defined, namely,

$$A \Rightarrow B = A \times B + \neg A \times V$$

the grade of membership of a point (u,v) will be high in $A \Rightarrow B$ if the grade of membership of u is low in A. This gives rise to an overlap between the terms $A \times B$ and $\neg A \times V$ when A is fuzzy, with the result that (see (8.50)), the inference drawn from A and $A \Rightarrow B$ is not B but[3]

$$A \circ (A \Rightarrow B) = B + A \circ (\neg A \times V) \tag{8.61}$$

where the difference term $A \circ (\neg A \times V)$ represents the effect of the overlap.

To avoid this phenomenon it may be necessary to define $A \Rightarrow B$ in a way that differentiates between the numerical truth-values in [0,1] and the truth-value <u>unknown</u> (see (6.52)). Also, it should be noted that for A COUPLED WITH B (see (8.34)), we do have

$$A \circ (A \text{ COUPLED WITH } B) = B \tag{8.62}$$

so long as A is a normal fuzzy set.

Fuzzy Theorems

By a fuzzy theorem or an assertion we mean a statement, generally of the form IF A THEN B, whose truth-value is <u>true</u> in an approximate sense and which can be inferred from a set of axioms by the use of approximate reasoning, e.g., by repeated application of the generalized <u>modus ponens</u> or similar rules.

As an informal illustration of the concept of a fuzzy theorem, let us consider the theorem in elementary geometry which asserts that if M_1, M_2 and M_3 are the midpoints of the sides of a triangle (see Fig. 8.5), then the lines AM_1, BM_2, and CM_3 intersect at a point.

A fuzzified version of this theorem may be stated as follows.

<u>Fuzzy Theorem 8.12</u> Let AB, BC and CA be approximate straight lines which

---

[3] We assume that A is normal, so that $A_r A_c = 1$.

form an approximate equilateral triangle with vertices A, B, C (see Fig. 8.6). Let M_1, M_2 and M_3 be approximate midpoints of the sides BC, CA and AB, respectively. Then the approximate straight lines AM_1, BM_2 and CM_3 form an approximate triangle $T_1T_2T_3$ which is more or less (more or less small) in relation to ABC.

Before we can proceed to "prove" this fuzzy theorem, we must make more specific the sense in which the terms approximate straight line, approximate midpoint, etc. should be understood. To this end, let us agree that by an approximate straight line AB we mean a curve passing through A and B such that the distance of any point on the curve from the straight line AB is small in relation to the length of AB. With reference to Fig. 8.7, this implies that we are assigning a linguistic value small to the distance d, with the understanding that d is interpreted as a fuzzy variable.

Let $(AB)^o$ denote the straight line AB. Then, by an approximate midpoint of AB we mean a point on AB whose distance from $M_1{}^o$, the midpoint of $(AB)^o$, is small.

Turning to the statement of the fuzzy theorem, let 0 be the intersection of the straight lines $(AM_1{}^o)^o$ and $(BM_2{}^o)^o$ (Fig. 8.8). Since M_1 is assumed to be an approximate midpoint of BC, the distance of M_1 from $M_1{}^o$ is small. Consequently, the distance of any point on $(AM_1)^o$ from $(AM_1{}^o)^o$ is small. Furthermore, since the distance of any point on AM_1 from $(AM_1)^o$ is small, it follows that the distance of any point on AM_1 from $(AM_1{}^o)^o$ is more or less small.

The same argument applies to the distance of points on BM_2 from $(BM_2{}^o)^o$. Then, taking into consideration that the angle between $(AM_1)^o$ and $(BM_2)^o$ is approximately 120°, the distance between an intersection of AM_1

and BM_2 and 0 is (more or less)^2 small (that is, more or less (more or less small.) From this it follows that the distance of any vertex of the triangle T_1 T_2 T_3 from 0 is (more or less)^2 small. It is in this sense that the triangle T_1 T_2 T_3 is (more or less)^2 small in relation to ABC.

The reasoning used above is both approximate and qualitative in nature. It uses as its point of departure the fact that AM_1, BM_2 and CM_3 intersect at 0, and employs what, in effect, are qualitative continuity arguments. Clearly, the "proof" would be longer and more involved if we had to start from the basic axioms of Euclidean geometry rather than the nonfuzzy theorem which served as our point of departure.

At this point, what we can say about fuzzy theorems is highly preliminary and incomplete in nature. Nonetheless, it appears to be an intriguing area for further study and eventually may prove to be of use in various types of ill-defined decision processes.

Graphical Representation by Fuzzy Flowcharts

As pointed out in [7], in the representation and execution of fuzzy algorithms it is frequently very convenient to employ flowcharts for the purpose of defining relations between variables and assigning values to them.

In what follows, we shall not concern ourselves with the many complex issues arising in the representation and execution of fuzzy algorithms. Thus, our limited objective is merely to clarify the role played by the decision boxes which are associated with fuzzy rather than nonfuzzy predicates by relating their function to the assignment of restrictions on base variables.

In the conventional flowchart, a decision box such as A in Fig. 8.9, represents a unary[4] predicate, A(x). Thus, transfer from point 1 to point 2 signifies that A(x) is <u>true</u>, while transfer from 1 to 3 signifies that A(x) is <u>false</u>.

The concepts introduced in the preceding sections provide us with a basis for extending the notion of a decision box to fuzzy sets (or predicates). Specifically, with reference to Fig. 8.9, suppose that A is a fuzzy subset of U, and the question associated with the decision box is: "Is x A," as in "Is x <u>small</u>," where x is a generic name for the input variable. Flowcharts containing decision boxes of this type will be referred to as <u>fuzzy flowcharts</u>.

If the answer is simply YES, we assign A to the restriction on x. That is, we set

$$R(x) = A \tag{8.63}$$

and transfer x from 1 to 2.

On the other hand, if the answer is NO, we set

$$R(x) = \neg A \tag{8.64}$$

and transfer x from 1 to 3.

As an illustration, if $A \triangleq \underline{small}$, then (8.63) would read

$$R(x) = \underline{small}. \tag{8.65}$$

If the answer is Yes/μ, where $0 \leq \mu \leq 1$, then we transfer x to 2 with the conclusion that the grade of membership of x in A is μ. We

[4]For simplicity, we shall not consider decision boxes having more than one input and two outputs.

also transfer x to 3 with the conclusion that the grade of membership of x in $\neg A$ is $1 - \mu$.

If the grade of membership μ is linguistic rather than numerical, we represent it as a linguistic truth-value. Typically, then, the answer would have the form Yes/true or YES/very true or YES/more or less true, etc. As before, we conclude that the grade of membership of x in A is μ, where μ is a linguistic truth-value, and transfer x to 3 with the conclusion that the grade of membership of x in $\neg A$ is $1 - \mu$.

If we have a chain of decision boxes as in Fig. 8.10, a succession of YES answers would transfer x from 1 to n + 1 and would result in the assignment to R(x) of the intersection of $A_1,\ldots,A_n$. Thus,

$$R(x) = A_1 \cap \ldots \cap A_n. \tag{8.66}$$

where $\cap$ denotes the intersection of fuzzy sets. (See also Fig. 8.11.)

As a simple illustration, suppose that x = John, A_1 = tall and A_2 = fat. Then, if the response to the question "Is John tall," is YES, and the response to "Is John fat," is YES, the restriction imposed by John is expressed by

$$R(\text{John}) = \underline{\text{tall}} \cap \underline{\text{fat}} \tag{8.67}$$

It should be noted that "John" is actually the name of a binary linguistic variable with two components named Height and Weight. Thus, (8.67) is equivalent to the assignment equations

$$\underline{\text{Height}} = \underline{\text{tall}} \tag{8.68}$$

and

$$\underline{\text{Weight}} = \underline{\text{fat}} \tag{8.69}$$

As implied by (8.66), a tandem connection of decision boxes represents the intersection of the fuzzy sets (or, equivalently, the conjunction of the fuzzy predicates) associated with them. In the case of nonfuzzy sets, their union may be realized by the scheme shown in Fig. 8.12. In this arrangement of decision boxes, it is clear that transfer from 1 to 2 implies that

$$R(u) = A + \neg A \cap B \tag{8.70}$$

and since

$$A \cap B \subset A \tag{8.71}$$

it follows that (8.70) may be rewritten as

$$\begin{aligned} R(u) &= A + A \cap B + \neg A \cap B \\ &= A + (A + \neg A) \cap B \\ &= A + B \end{aligned} \tag{8.72}$$

since

$$A + \neg A = U \tag{8.73}$$

and

$$U \cap B = B \tag{8.74}$$

The same scheme would not yield the union of fuzzy sets since the identity

$$A + \neg A = U \tag{8.75}$$

does not hold exactly if A is fuzzy. Nevertheless, we can agree to interpret the arrangement of decision boxes in Fig. 8.12 as one that represents the union of A and B. In this way, we can remain on the familiar ground of flowcharts involving nonfuzzy decision boxes. The flowchart shown in Fig. 8.14 illustrates the use of this convention in the definition of <u>Hippie</u>.

The conventions described above may be used to represent in a graphical form the assignment of a linguistic value to a linguistic variable. Of particular use in this connection is a tandem connection of decision boxes which represent a series of <u>bracketing</u> questions which are intended to narrow down the range of possible values of a variable. As an illustration, suppose that x = John and (see Fig. 8.13)

$$\begin{aligned} A_1 &= \underline{\text{tall}} \\ A_2 &= \underline{\text{very}}\ \underline{\text{tall}} \\ A_3 &= \underline{\text{very}}\ \underline{\text{very}}\ \underline{\text{tall}} \\ A_4 &= \underline{\text{extremely}}\ \underline{\text{tall}} \end{aligned} \tag{8.76}$$

If the answer to the first question is YES, we have

$$R(x) = \underline{\text{tall}} \tag{8.77}$$

If the answer to the second question is YES and to the third question is NO, then

$$R(\text{John}) = \underline{\text{very}}\ \underline{\text{tall}}\ \underline{\text{and}}\ \underline{\text{not}}\ \underline{\text{very}}\ \underline{\text{very}}\ \underline{\text{tall}} \tag{8.78}$$

which brackets the height of John between <u>very</u> <u>tall</u> and <u>not</u> <u>very</u> <u>very</u> <u>tall</u>.

By providing a mechanism - as in bracketing - for assigning linguistic values in stages rather than in one step, fuzzy flowcharts can be very helpful in the representation of algorithmic definitions of fuzzy concepts. The basic idea in this instance is to define a complex or a new fuzzy concept in terms of simpler or more familiar ones. Since a fuzzy concept may be viewed as a name for a fuzzy set, what is involved in this approach is, in effect, the decomposition of a fuzzy set into a combination of simpler fuzzy sets.

As an illustration, suppose that we wish to define the term Hippie, which may be viewed as a name of a fuzzy subset of the universe of humans. To this end, we employ the fuzzy flowchart[5] shown in Fig. 8.14. In essence, this flowchart defines the fuzzy set Hippie in terms of the fuzzy sets labeled Long Hair, Bald, Shaved, Job and Drugs. More specifically, it defines the fuzzy set Hippie as ($+ \triangleq$ union)

$$\underline{\text{Hippie}} = (\underline{\text{Long Hair}} + \underline{\text{Bald}} + \underline{\text{Shaved}}) \cap \underline{\text{Drugs}} \cap \neg\, \underline{\text{Job}} \quad (8.79)$$

Suppose that we pose the following questions and receive the indicated answers.

Does x have Long Hair?	YES
Does x have a Job?	NO
Does x take Drugs?	YES

Then, we assign to x the restriction

$$R(x) = \underline{\text{Long Hair}} \cap \neg\, \underline{\text{Job}} \cap \underline{\text{Drugs}}$$

and since it is contained in the right-hand member of (8.79), we conclude that x is a Hippie.

By modifying the fuzzy sets entering into the definition of Hippie through the use of hedges such as very, more or less, extremely, etc., and by allowing the answers to be of the form YES/μ or NO/μ, where μ is

---

[5] It should be understood, of course, that this highly oversimplified definition is used merely as an illustration and has no pretense at being accurate, complete or realistic.

a numerical or linguistic truth-value, the definition of Hippie can be adjusted to fit more closely our conception of what we want to define. Furthermore, we may use a soft and (see Comment 3.7) to allow some trade-offs between the characteristics which define a hippie. And, finally, we may allow our decision boxes to have multiple inputs and multiple outputs. In this way, a concept such as Hippie can be defined as completely as one may desire in terms of a set of constituent concepts each of which, in turn, may be defined algorithmically. In essence, then, in employing a fuzzy flowchart to define a fuzzy concept such as Hippie, we are decomposing a statement of the general form

$$v(u \text{ is: linguistic value of a Boolean linguistic variable } \mathcal{X}) = \text{linguistic value of a Boolean linguistic truth-variable } \mathcal{T} \tag{8.80}$$

into truth-value assignments of the same form, but involving simpler or more familiar variables in the left-hand member of (8.80).

Concluding Remarks

In this as well as in the preceding sections, our main concern centered on the development of a conceptual framework for what may be called a linguistic approach to the analysis of complex or ill-defined systems and decision processes. The substantive differences between this approach and the conventional quantitative techniques of system analysis raise many issues and problems which are novel in nature and hence require a great deal of additional study and experimentation. This is true, in particular, of some of the basic aspects of the concept of a linguistic variable on which we have dwelt only briefly in our exposition, namely: linguistic approximation, representation of linguistic hedges, nonnumerical base variables, λ- and β-interaction, fuzzy theorems, linguistic probability distributions, fuzzy flowcharts and others.

Although the linguistic approach is orthogonal to what have become the prevailing attitudes in scientific research, it may well prove to be a step in the right direction, that is, in the direction of lesser preoccupation with exact quantitative analyses and greater acceptance of the pervasiveness of imprecision in much of human thinking and perception. It is our belief that, by accepting this reality rather than assuming that the opposite is the case, we are likely to make more real progress in the understanding of the behavior of humanistic systems than is possible within the confines of traditional methods.

REFERENCES

1. Sir William Thomson, Popular Lectures and Addresses, McMillan and Co., London, 1891.

2. E. Feigenbaum, Computers and Thought, McGraw-Hill Book Co., New York,1963.

3. M. Minsky and S. Papert, Perceptrons: An Introduction to Computational Geometry, The M.I.T. Press, Cambridge, Mass., 1969.

4. M. Arbib, The Metaphorical Brain, Wiley-Interscience, New York, 1972.

5. A. Newell and H. Simon, Human Problem Solving, Prentice-Hall, Englewood Hills, N.J., 1972.

6. L. A. Zadeh, "Fuzzy Languages and Their Relation to Human and Machine Intelligence," Proc. of Inter. Conference on Man and Computer, Bordeaux, France, pp.130-165, S. Karger, Basel, 1972.

7. L. A. Zadeh, "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," IEEE Trans. on Systems, Man and Cybernetics, vol. SMC-3, pp.28-44, January 1973.

8. R. E. Bellman, R.E.Kalaba and L.A.Zadeh,"Abstraction and Pattern Classification," Jour. Math. Analysis and Appl., vol.13, pp.1-7, 1966.

9. M. Black, "Reasoning With Loose Concepts," Dialogue, vol.2, pp.1-12,1963.

10. L. Wittgenstein, "Logical Form," Proc. of the Aristotelian Society, vol.9, pp.162-171,1929.

11. M. Scriven, "The Logic of Criteria," Jour. of Philosophy, vol.56, pp.857-868, 1959.

12. H. Khatchadourian, "Vagueness, Meaning and Absurdity," Amer. Phil. Quart., vol.2, pp.119-129, 1965.

13. R.R.Verma, "Vagueness and the Principle of Excluded Middle," Mind, vol.79, pp.67-77, 1970.

14. J. A. Goguen, "The Logic of Inexact Concepts," Synthese, vol.19, pp.325-373, 1969.

15. E. Adams, "The Logic of 'Almost All'," to appear in Jour. of Philosophical Logic.

16. K. Fine, "Vagueness, Truth and Logic," Department of Philosophy, University of Edingburgh, 1973.

17. B. S. van Frassen, "Presuppositions, Supervaluations and Free Logic," in The Logical Way of Doing Things, K. Lambert, ed., Yale Univ. Press, New Haven, 1969.

18. G. Lakoff, "Linguistics and Natural Logic," in Semantics of Natural Languages, D. Davidson and G. Harman (eds.), Doudrecht, The Netherlands, D. Reidel, 1971.

19. L. A. Zadeh, "Shadows of Fuzzy Sets," Prob. in Trans. of Information, vol.2, pp.37-44, 1966.

20. A. Kaufmann, Theory of Fuzzy Sets, Masson, Paris, 1972.

21. J. Goguen, "L-Fuzzy Sets," Jour. Math. Analysis and Appl., vol.18, pp.145-174,1967.

22. J. G. Brown, "A Note on Fuzzy Sets," Inf. Control, vol.18, pp.32-39,1971

23. M. Mizumoto, J. Toyoda and K. Tanaka, "General Formulation of Formal Grammars," Inf. Sci., vol.4, pp.87-100, 1972.

24. L. A. Zadeh, "Similarity Relations and Fuzzy Orderings," Inf. Sci., vol.3, pp.177-200, 1971.

25. R. E. Bellman and L. A. Zadeh, "Decision-Making in a Fuzzy Environment," Management Science, vol. 17, pp. B-141-B-164, 1970.

26. R. E. Bellman and M. Giertz, "On the Analytic Formalism of the Theory of Fuzzy Sets," Inf. Sci., vol. 5, pp. 149-156, 1973.

27. L. A. Zadeh, "A Fuzzy-Set-Theoretic Interpretation of Linguistic Hedges," Jour. of Cybernetics, vol.2, pp.4-34, 1972.

28. E. F. Codd, "Relational Completeness of Data Base Sublanguages," in Courant Computer Science Symposia, vol.6, Prentice-Hall, Englewood Hills, N.J., 1971.

29. L. A. Zadeh, "Fuzzy Sets," Inf. Control, vol.8, pp.338-353, 1965.

30. A. Thomasian, The Structure of Probability Theory With Applications, McGraw-Hill Book Co., New York, 1969.

31. R. E. Moore, Interval Analysis, Prentice-Hall, Englewood Hills, N.J.,1966.

32. J. A. Brzozowski, "Regular Expressions for Linear Sequential Circuits," IEEE Trans. on Elec. Comp., vol. EC-14, pp.148-156, 1965.

33. E. Shamir, "Algebraic, Rational and Context-Free Power Series in Noncommuting Variables," in M. Arbib's Algebraic Theory of Machines, Languages and Semigroups, pp.329-341, Academic Press, New York, 1968.

34. A. Blikle, "Equational Languages," Inf. Control, vol.21, pp.134-147,1972.

35. D. J. Rosenkrantz, "Matrix Equation and Normal Forms for Context-Free Grammars," Jour. Assoc. Comp. Mach., vol.14, pp.501-507, 1967.

36. J. E. Hopcroft and J.D.Ullman, Formal Languages and Their Relations to Automata, Addison-Wesley, Reading, Mass., 1969.

37. A. V. Aho and J. D. Ullman, The Theory of Parsing, Translation and Compiling, Prentice-Hall, Englewood Hills, N.J., 1973.

38. G. Lakoff, "Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts," Proc. 8 th Regional Meeting of Chicago Linguistic Society, Univ. of Chicago Linguistics Dept., April 1972.

39. L. A. Zadeh, "Quantitative Fuzzy Semantics," Inf. Sci., vol.3, pp.159-176, 1971.

40. D. Knuth, "Semantics of Context-Free Languages," Math. Systems Theory, vol.2, pp.127-145, 1968.

41. P. Lucas et al, "Method and Notation for the Formal Definition of Programming Languages," Report TR 25.087, IBM Laboratory, Vienna, 1968.

42. P. Wegner, "The Vienna Definition Language," ACM Computing Surveys, vol.4, pp.5-63, 1972.

43. J. A. Lee, Computer Semantics, Van Nostrand-Reinhold, New York, 1972.

44. Z. Kohavi, Switching and Finite Automata Theory, McGraw-Hill Book Co., New York, 1970.

45. G. E. Hughes and M. J. Cresswell, An Introduction to Modal Logic, Methuen, London, 1968.

46. N. Rescher, Many-Valued Logics, McGraw-Hill Book Co., New York, 1969.

47. R. Barkan, "A Functional Calculus of First Order Based on Strict Implication," Jour. Symbolic Logic, vol.11, pp.1-16, 1946.

48. L. A. Zadeh, "Probability Measures of Fuzzy Events," Jour. Math. Analysis and Appl., vol,23, pp.421-427, 1968.

49. A. DeLuca and S. Termini, "A Definition of Non-Probabilistic Entropy in the Setting of Fuzzy Set Theory," Inf. Control, vol.20, pp.201-312, 1972.

50. J. Hintikka and P. Suppes, eds., Aspects of Inductive Logic, North-Holland Publ. Co., Amsterdam, 1966.

51. T. Winograd, Understanding Natural Language, Academic Press, New York, 1972.

52. A. DeLuca and S. Termini, "Algebraic Properties of Fuzzy Sets," Jour. Math. Analysis and Appl., vol.40, pp.373-386, 1972.

53. A. Rosenfeld, "Fuzzy Groups," Jour. Math. Analysis and Appl., vol.35, pp.512-517, 1971.

54. L. A. Zadeh, "Fuzzy Algorithms," Inf. Control, vol.12, pp.94-102,1968.

55. E. Santos, "Fuzzy Algorithms," Inf. Control, vol.17, pp.326-339, 1970.

56. S. K. Chang, " On the Execution of Fuzzy Programs Using Finite State Machines," IEEE Trans. Elec. Comp., vol.C-21, pp.241-253, 1972.

57. S. S. L. Chang and L. A. Zadeh, "Fuzzy Mapping and Control," IEEE Trans. Systems, Man and Cybernetics, vol. SMC-2, pp.30-34, 1972.

58. E. T. Lee, "Fuzzy Languages and Their Relation to Automata," Dissertation, Dept. of Elec. Eng. and Computer Sci., Univ. of Calif., Berkeley, 1972.

59. R. C. T. Lee, "Fuzzy Logic and the Resolution Principle," Jour. Assoc. Comp. Mach., vol. 19, pp. 109-119, 1972.

60. K. M. Colby, S. Weber and F. D. Hilf, "Artificial Paranoia," Jour. Art. Intelligence, vol. 2, pp. 1-25, 1971.

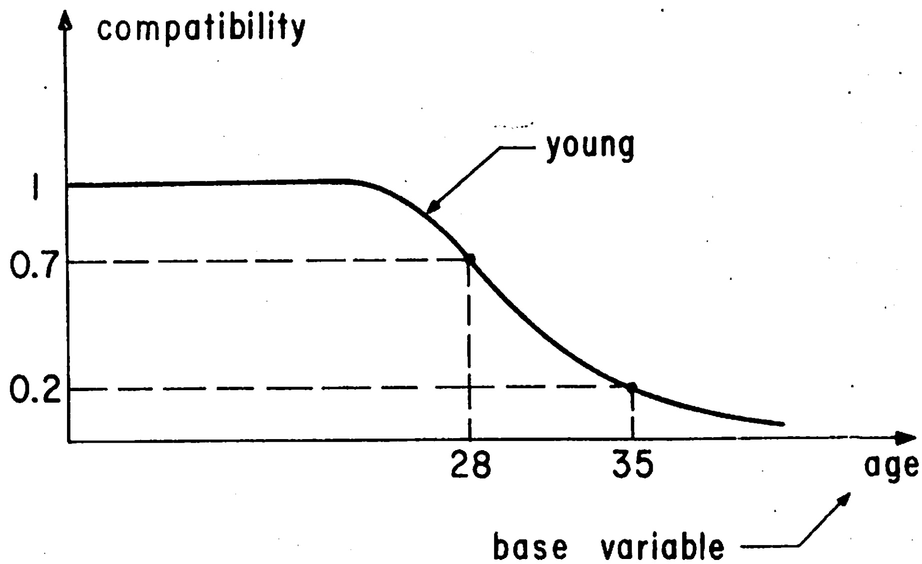


Fig.1.1. Compatibility function for young.

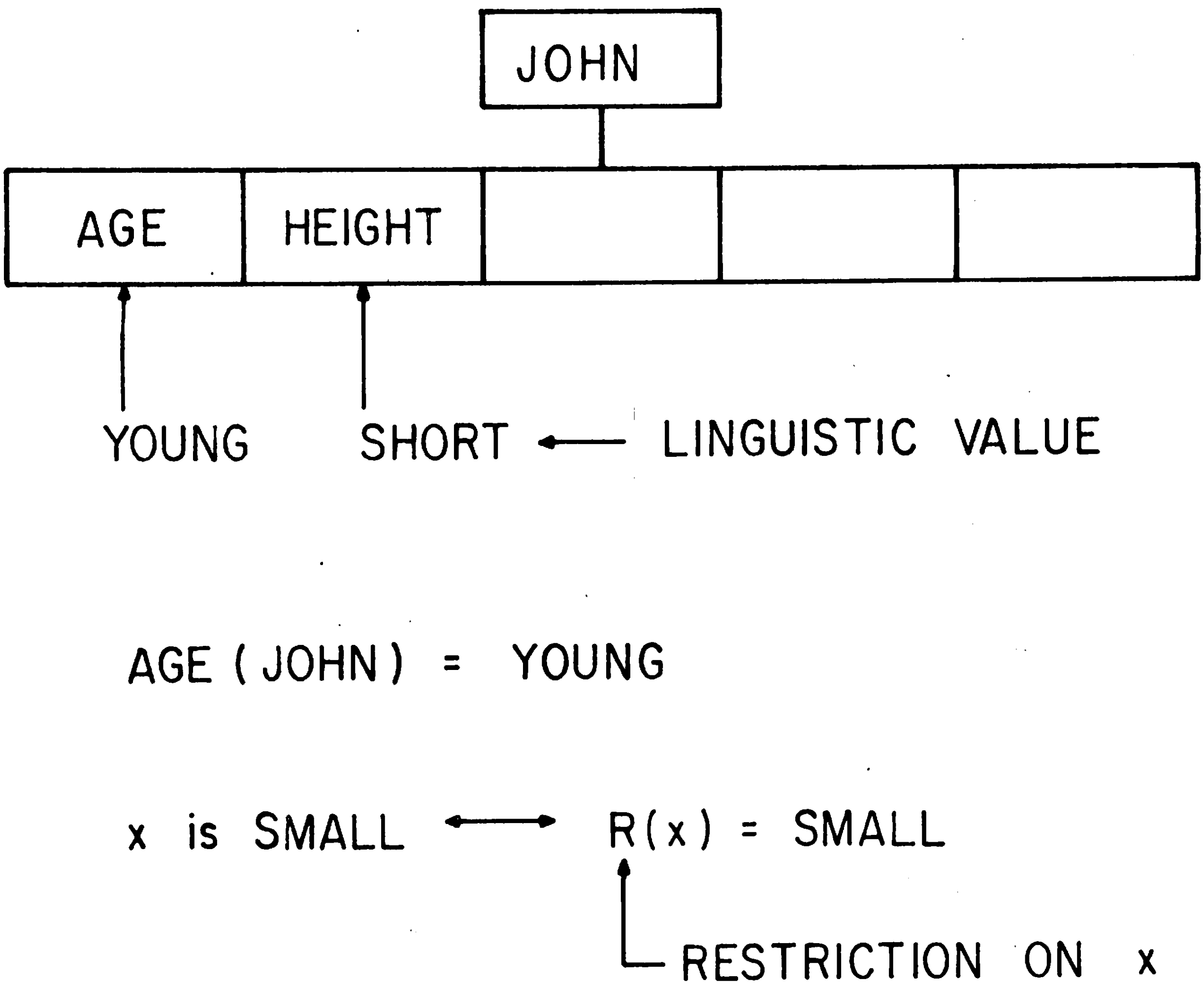


Fig. 1.2 Assignment of linguistic values to attributes of John and x.

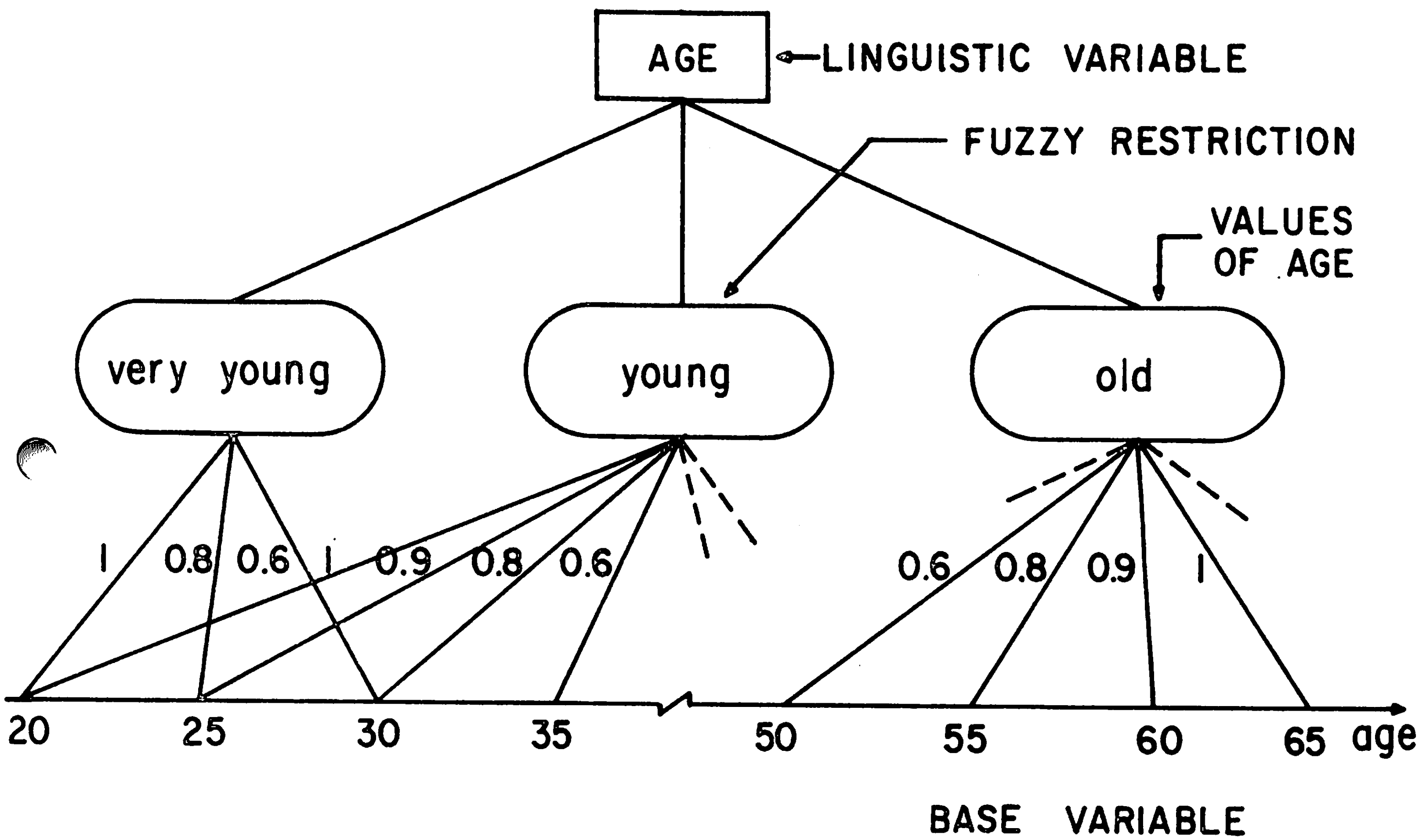


Fig.1.3. Hierarchical structure of a linguistic variable.

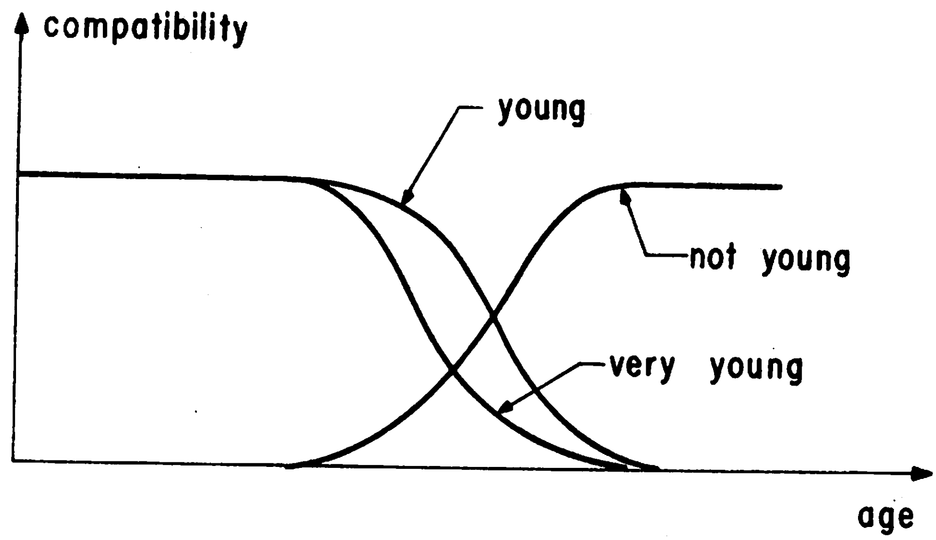


Fig.1.4. Compatibilities of young, not young, and very young.

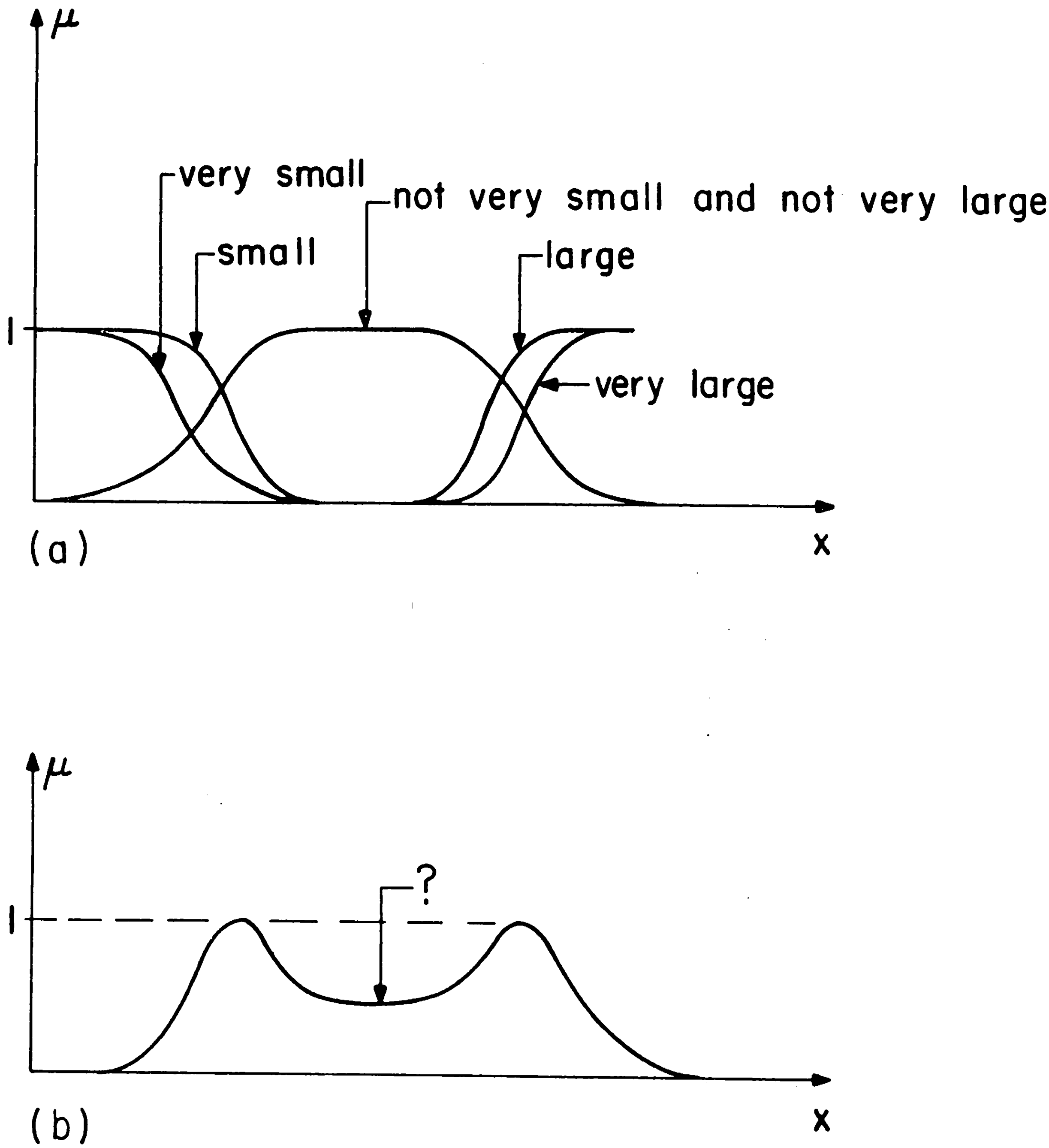


Fig. 1.5 (a) Compatibilities of <u>small</u>, <u>very small</u>, <u>large</u>, <u>very large</u> and <u>not very small and not very large</u>.

(b) The problem of linguistic approximation is that of finding an approximate linguistic characterization of a given compatibility function.

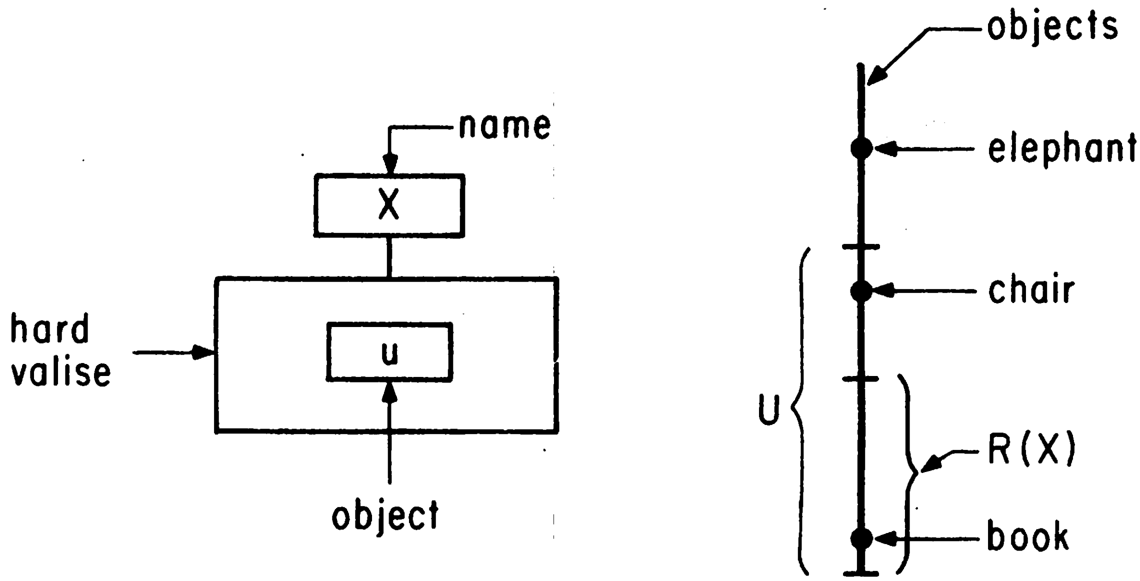


Fig. 2.1. Illustration of the valise analogy for a unary nonfuzzy variable.

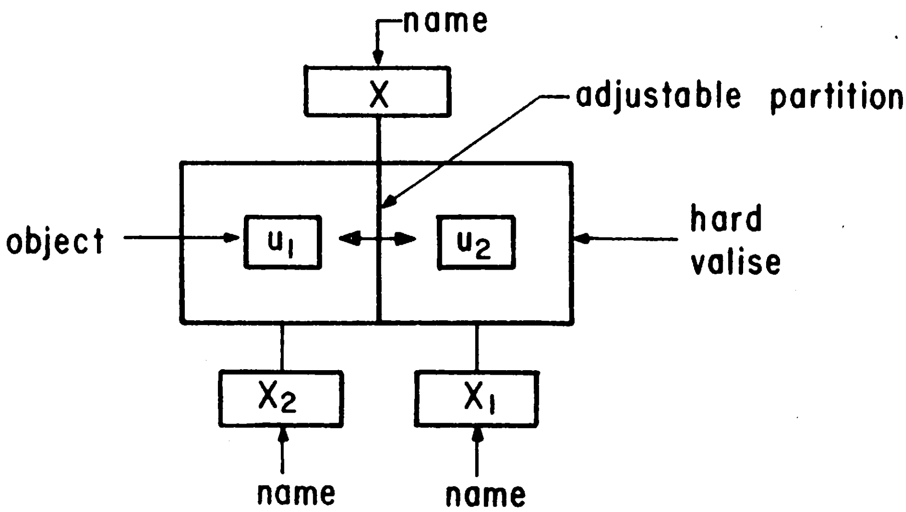


Fig.2.2. Valise analogy for a binary nonfuzzy variable.

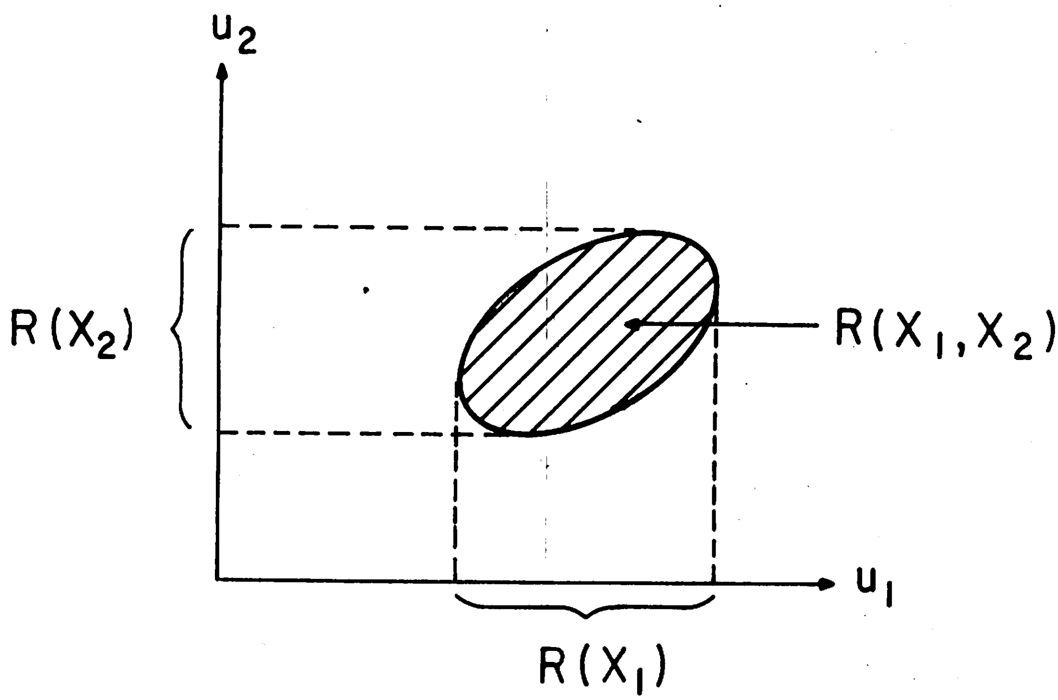


Fig. 2.3. Marginal restrictions induced by $R(X_1,X_2)$

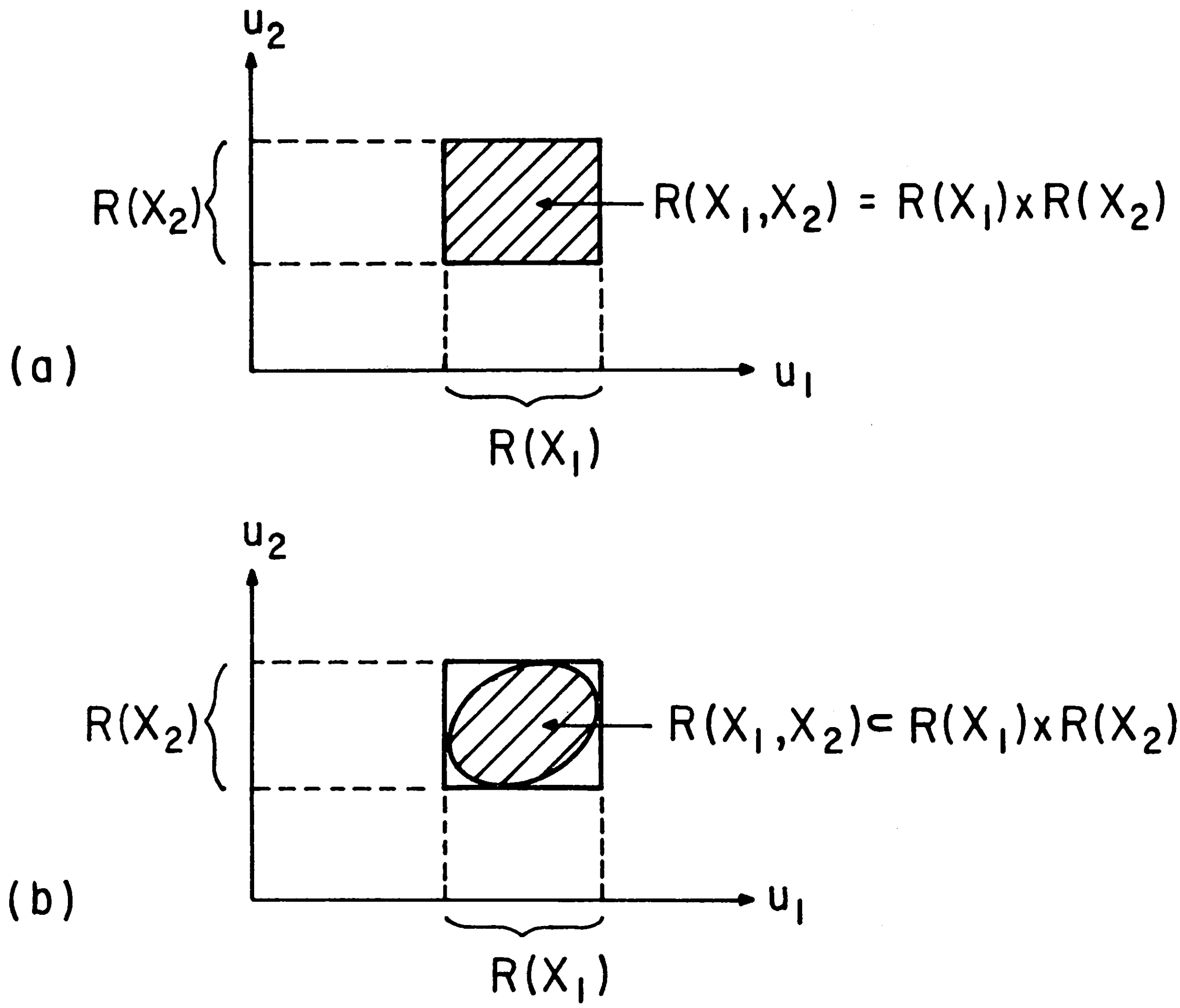


Fig.2.4. (a) X_1 and X_2 are noninteractive.

(b) X_1 and X_2 are interactive.

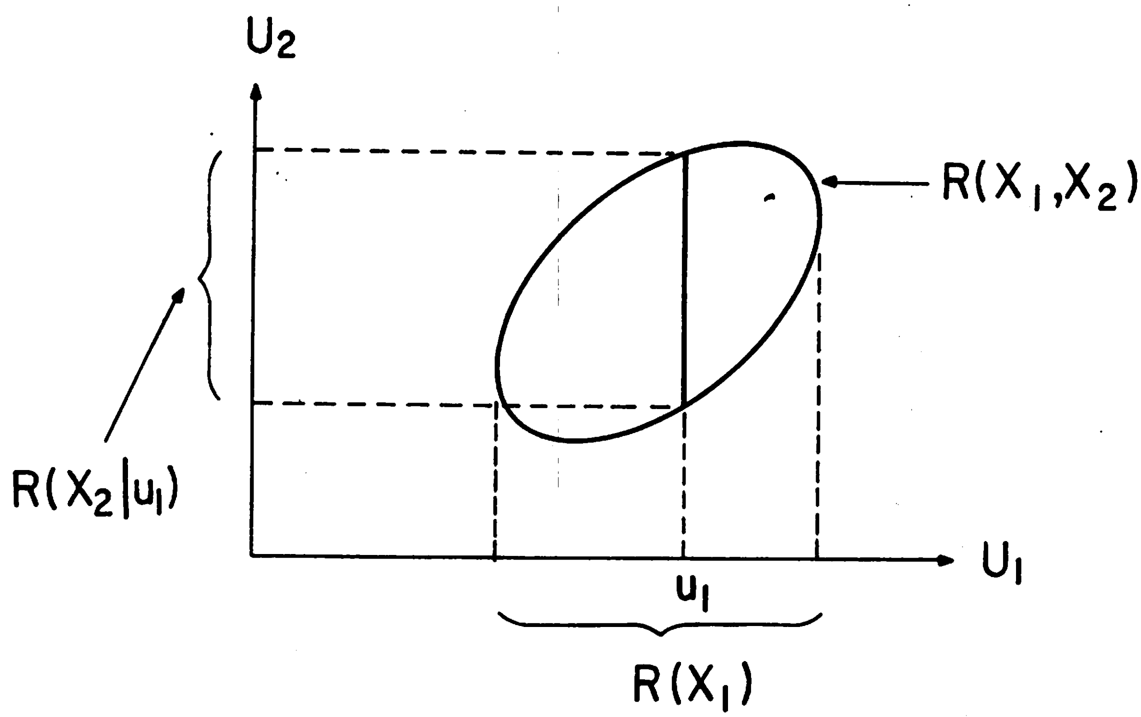


Fig. 2.5. $R(X_2|u_1)$ is the restriction on u_2 conditioned on u_1.

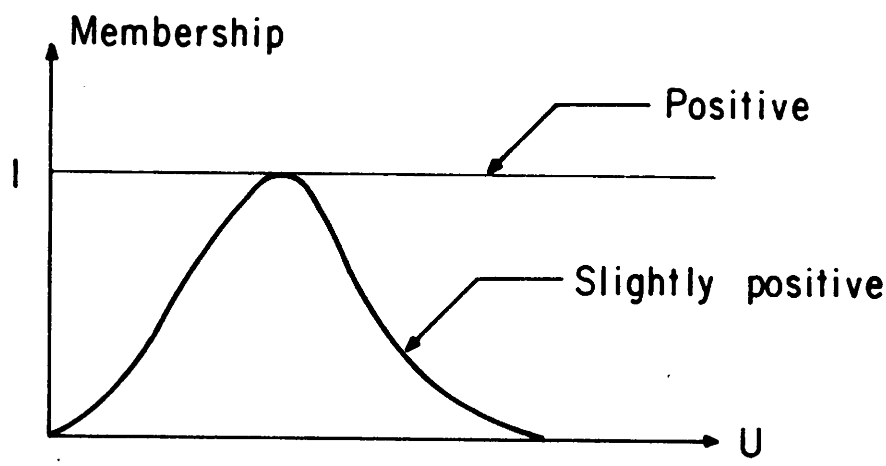


Fig.3.1. Membership functions of positive and slightly positive.

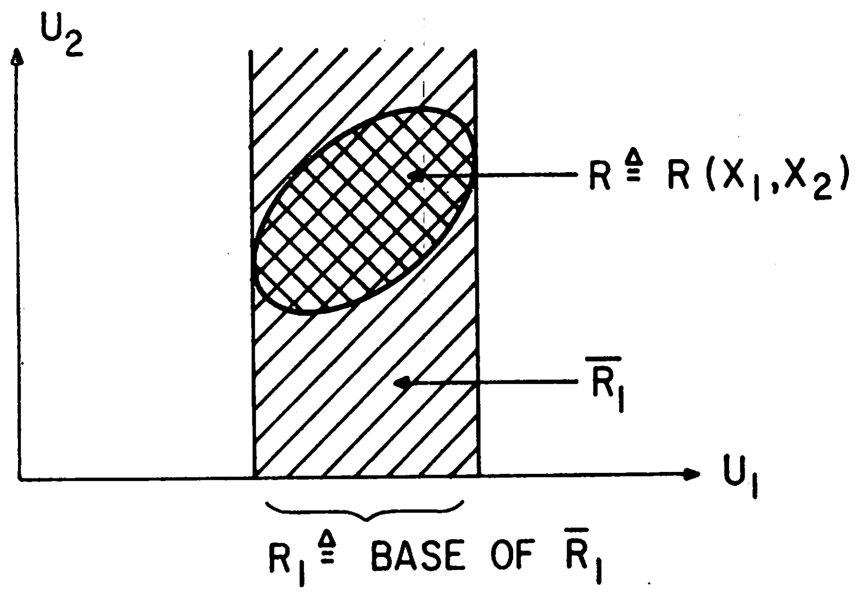


Fig.3.2. R_1 is the base of the cylindrical set $\overline{R}_1$.

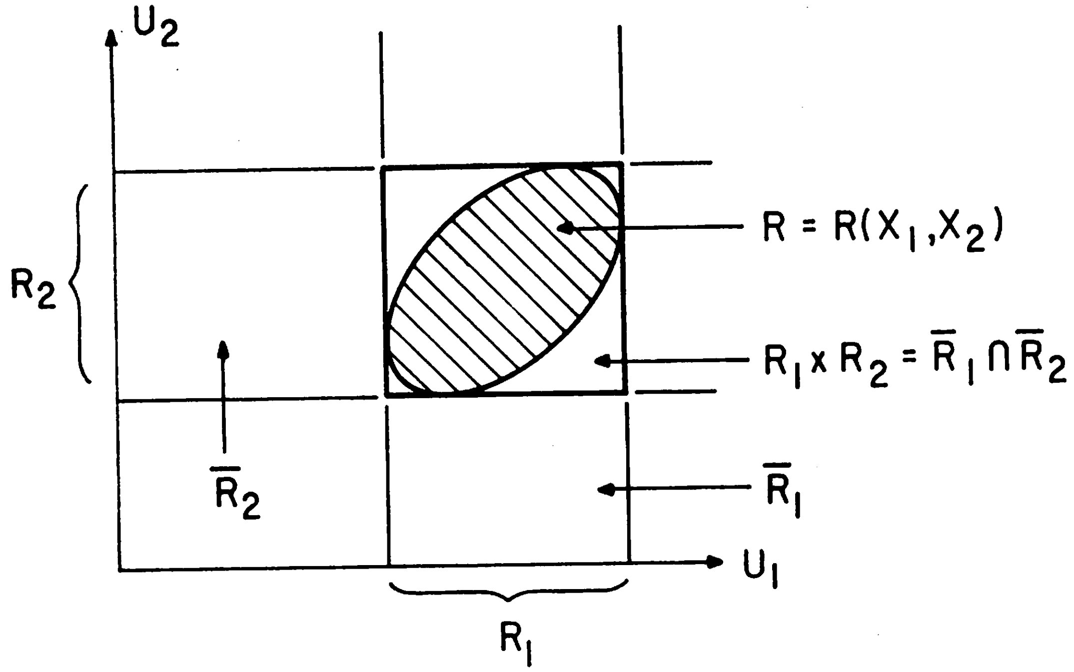


Fig.3.3. Relation between the cartesian product and intersection of cylindrical sets.

X	1	2	3	4	1v2	2v4
1	1	2	3	4	1v2	2v4
2	2	4	6	8	1v4	4v8
3	3	6	9	12	3v6	6v12
4	1	8	12	16	4v8	8v16
1v2	1v2	2v4	3v6	4v8	1v2v4	2v4v8

$$
\begin{array}{rl}
 & 3 \vee 5 \vee 6 \\
\times & 2 \vee 4 \vee 6 \\
\hline
 & 6 \vee 10 \vee 12 \\
 & 12 \vee 20 \vee 24 \\
 & 18 \vee 30 \vee 36 \\
\hline
\end{array}
$$

$$6 \vee 10 \vee 12 \vee 18 \vee 20 \vee 24 \vee 30 \vee 36$$

Table 3.4. Extension of the multiplication table to subsets of integers.

1 v 2 means 1 or 2.

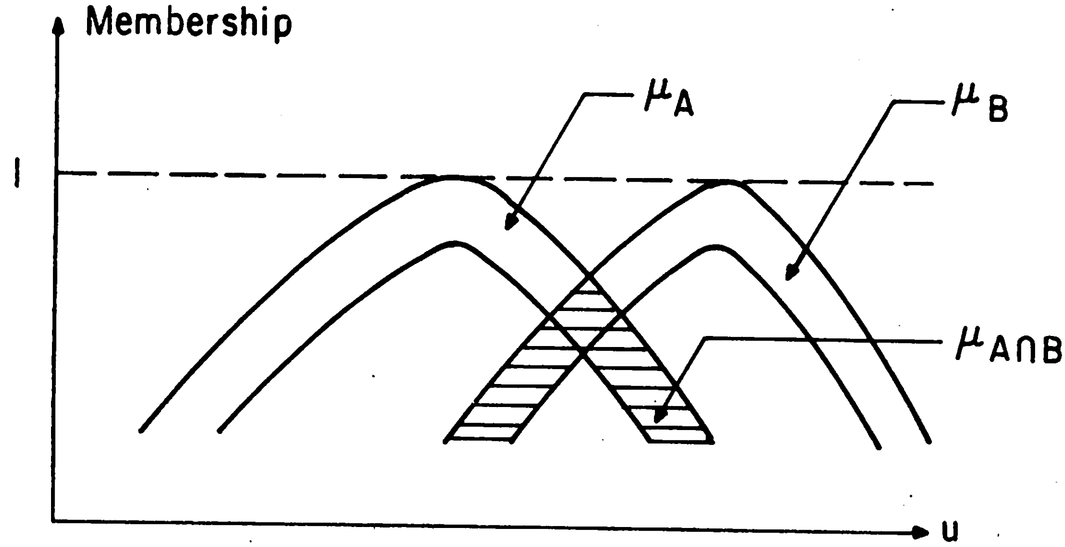


Fig. 3.5. Intersection of fuzzy sets with interval-valued membership functions.

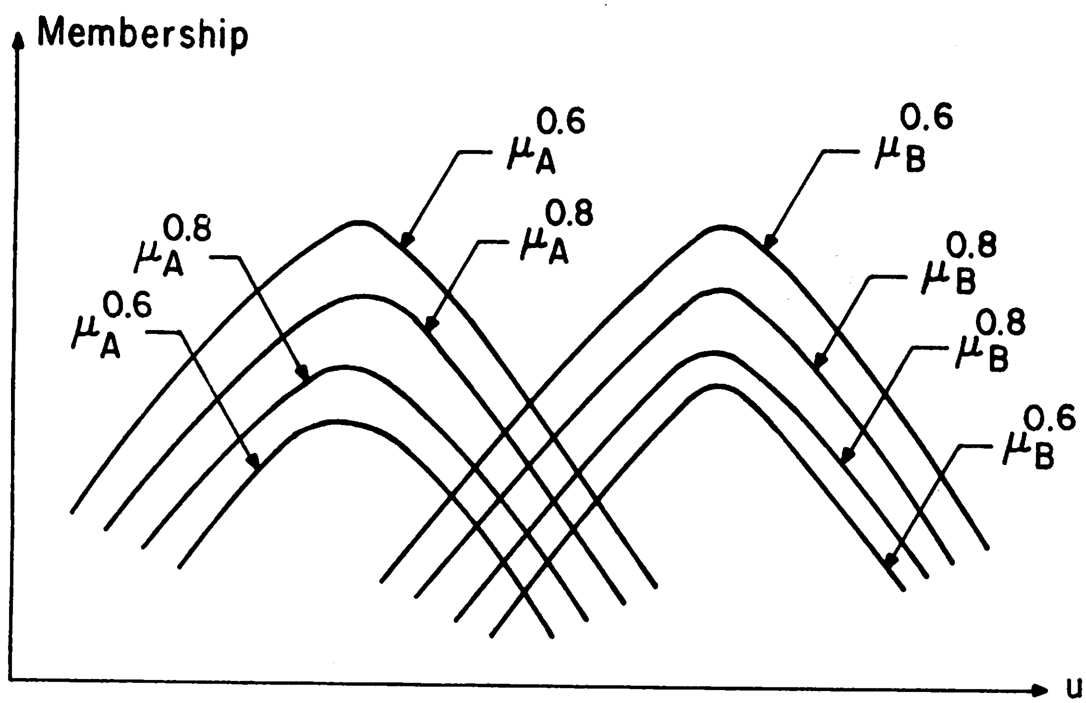
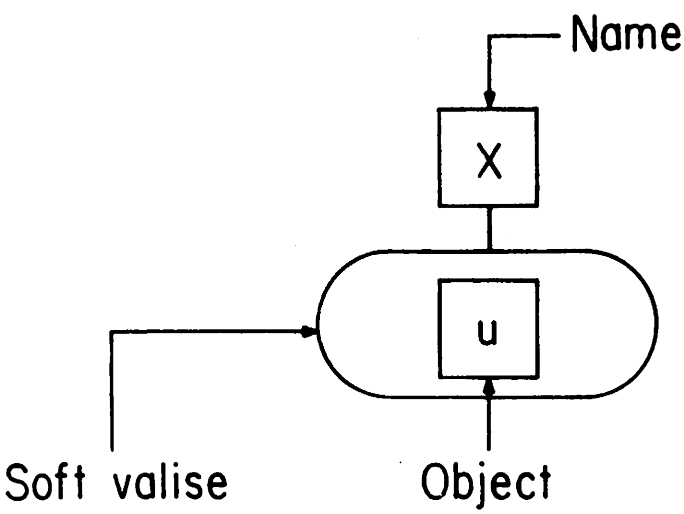


Fig.3.6. Level-sets of fuzzy membership functions μ_A and μ_B.



R(X)	
u	C(u)
Chair	0
Coat	0.6
Shoes	0.9
Shirt	1

Fig.4.1. Valise analogy for a unary fuzzy variable.

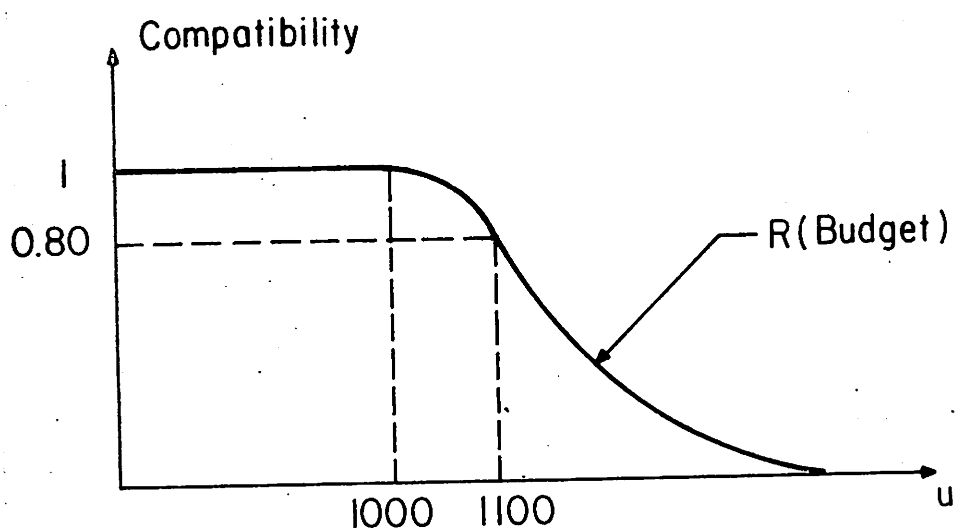
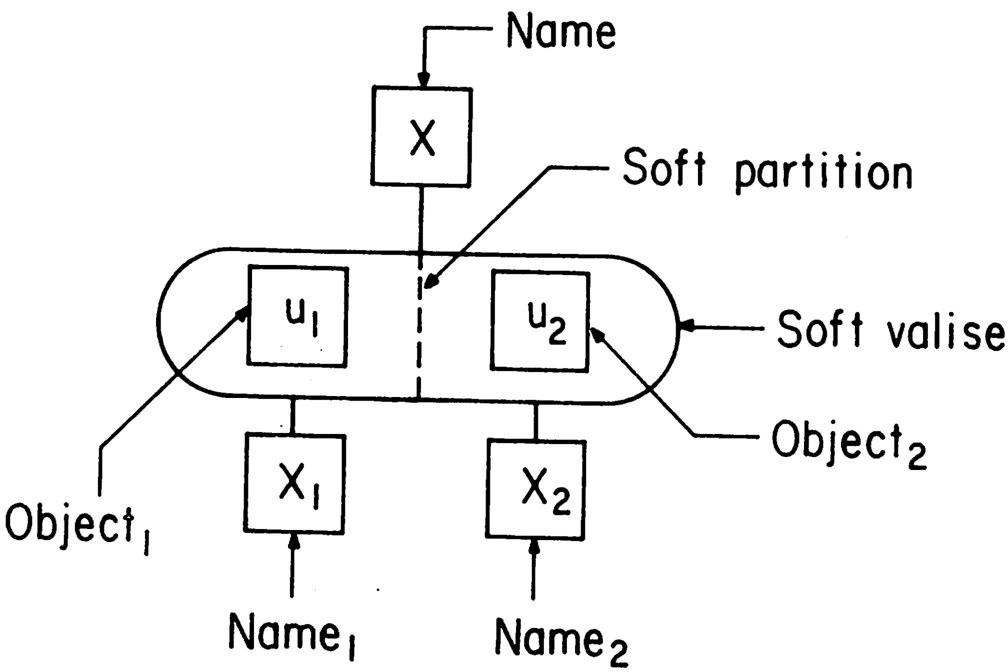


Fig. 4.2. Compatibility function of budget.



$R(X_1,X_2)$		
u_1	u_2	$c(u_1,u_2)$
Coat	Shoes	0.8
Coat	Shirt	1
Coat	Coat	0.6

Fig. 4.3. Valise analogy for a binary fuzzy variable.

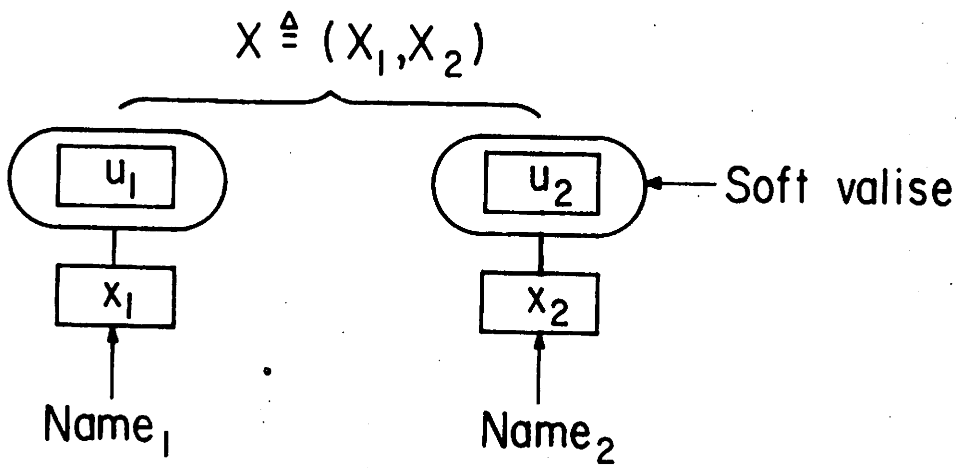


Fig.4.4. Valise analogy for noninteractive fuzzy variables.

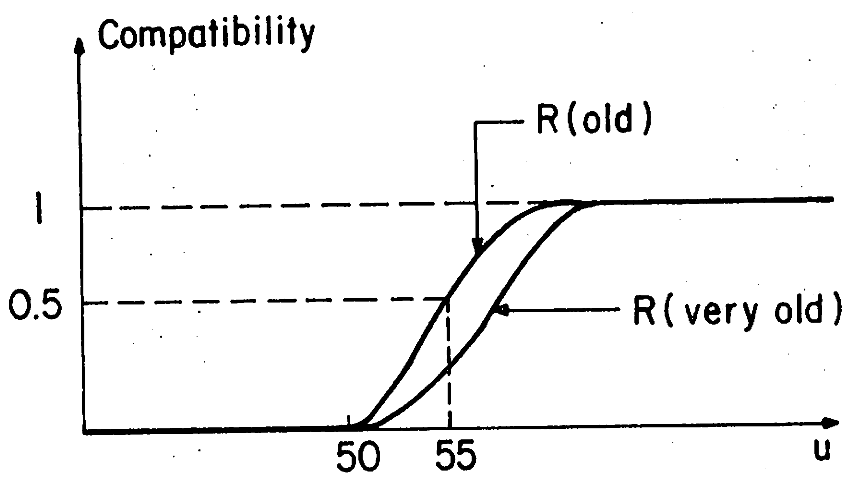


Fig. 5.1. Compatibility functions of old and very old.

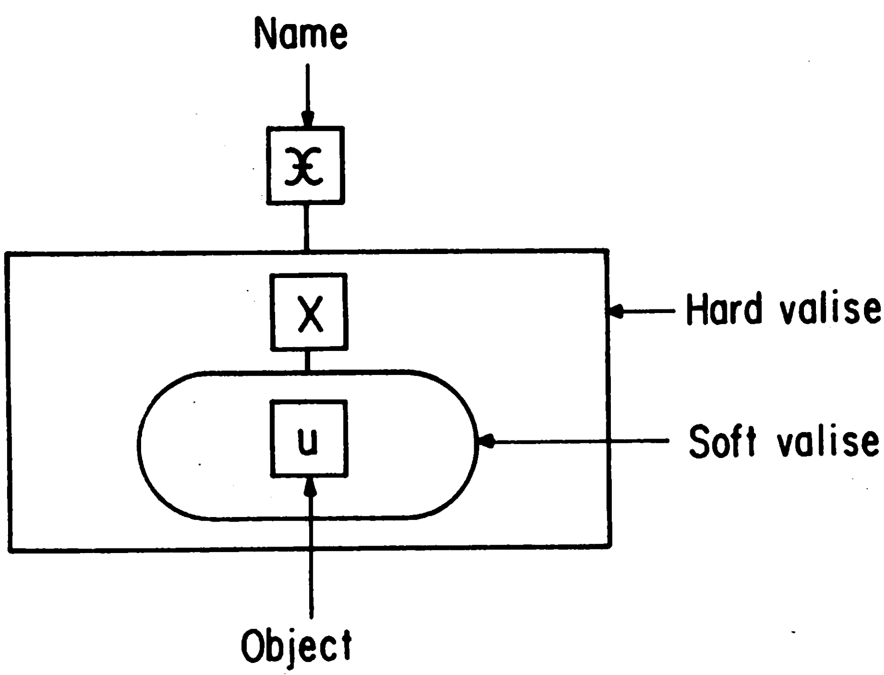


Fig.5.2. Valise analogy for a linguistic variable.

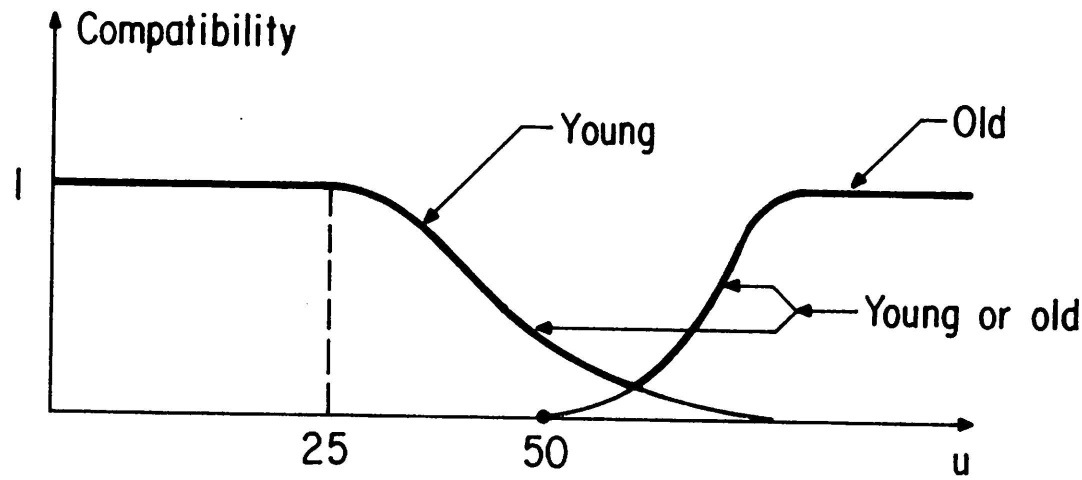


Fig.5.3. Compatibility function for young or old.

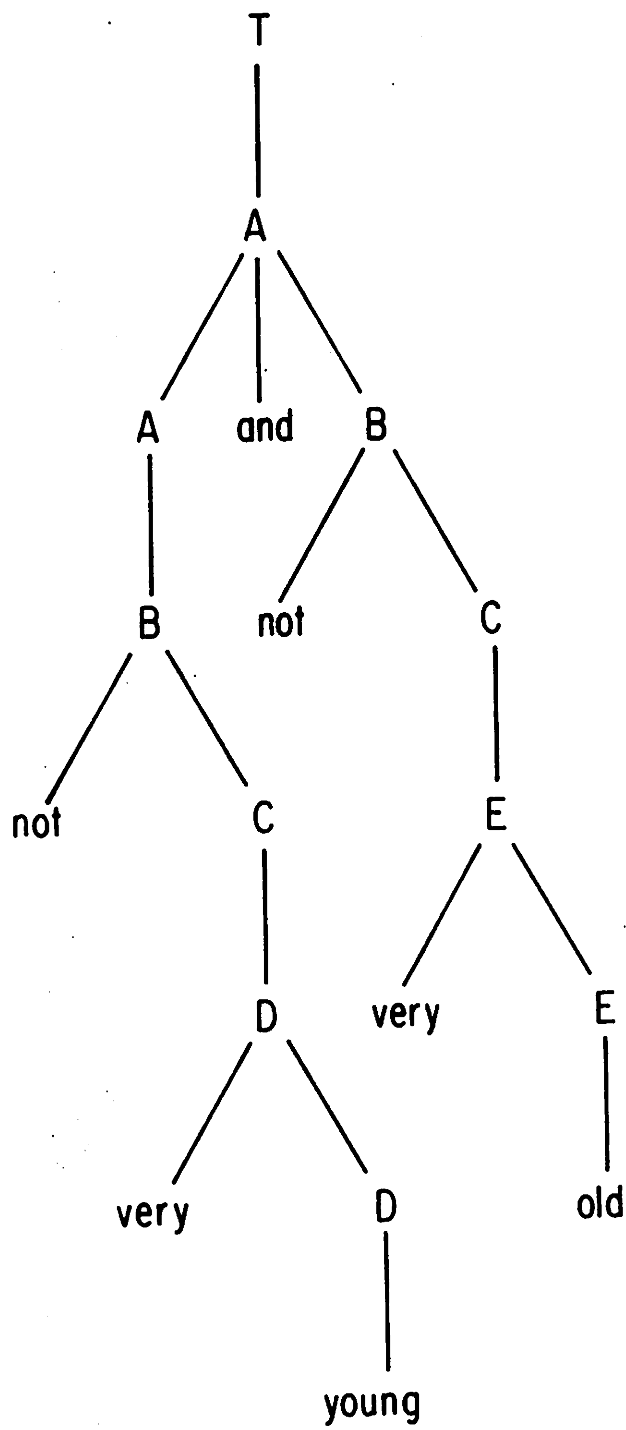


Fig.5.4. Syntax tree for not very young and not very old.

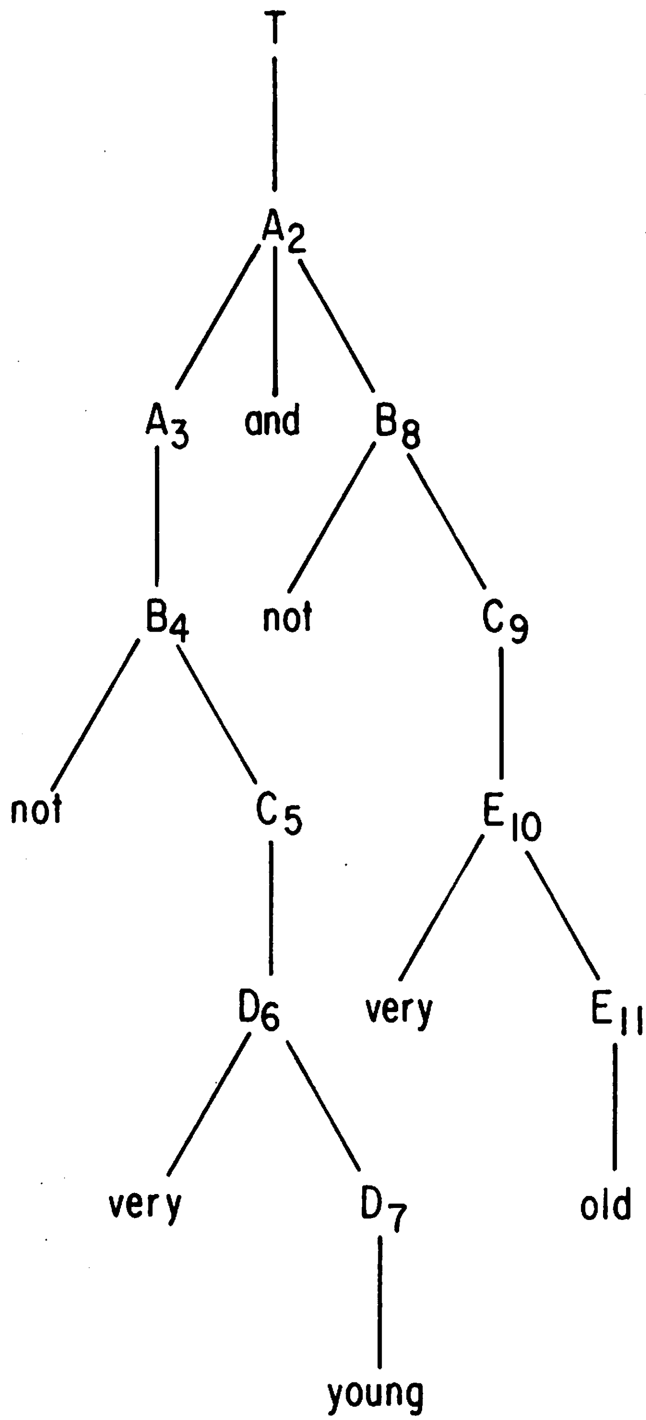


Fig.5.5. Computation of the meaning of <u>not very young and not very old</u>.

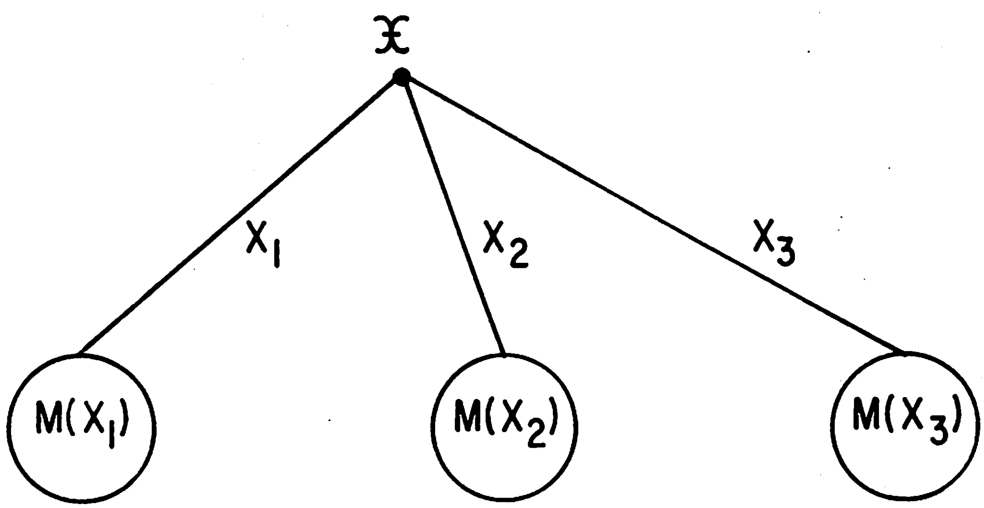


Fig.5.6. Representation of a linguistic variable as a Vienna definition language object.

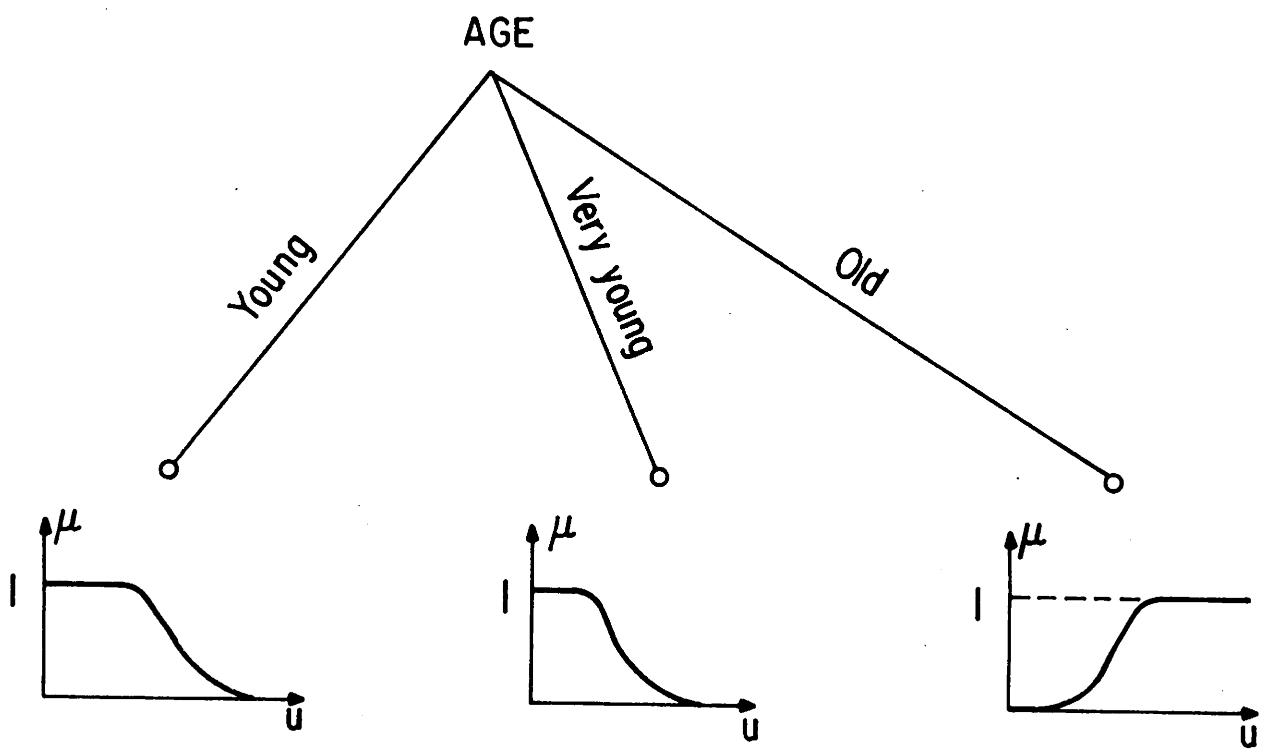


Fig.5.7. Representation of the linguistic variable Age as a Vienna definition language object.

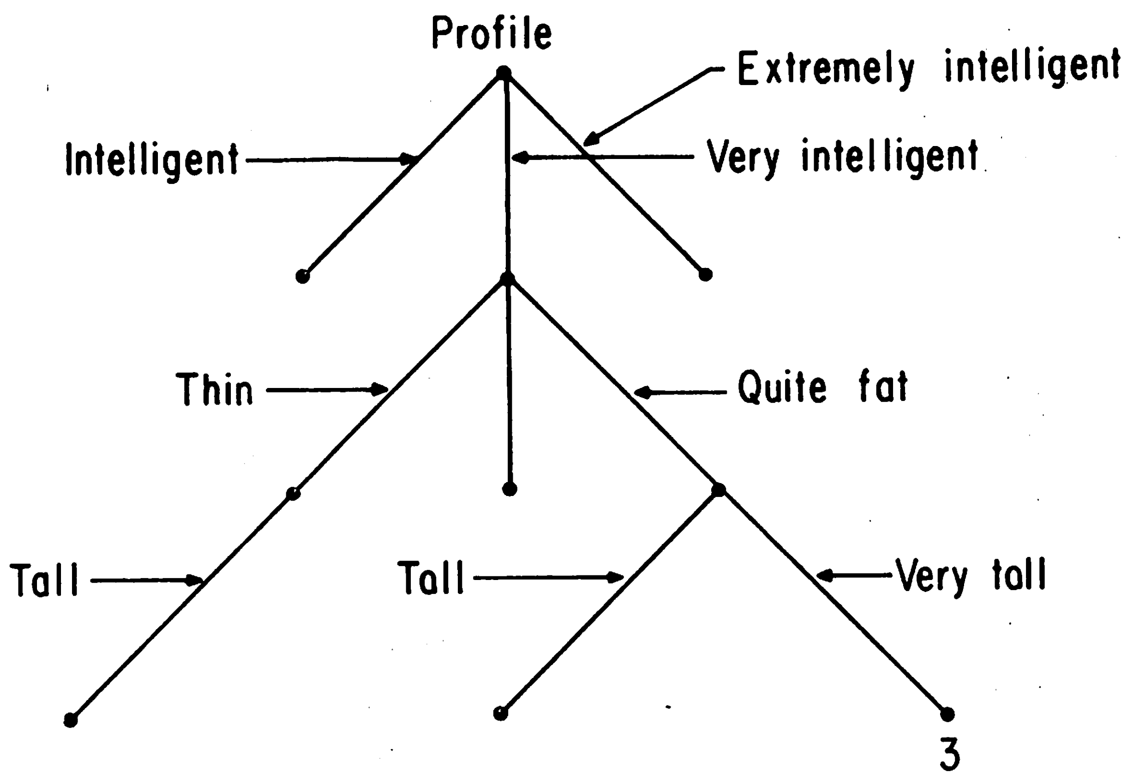


Fig.5.8. Tree representation of the linguistic variable Profile.

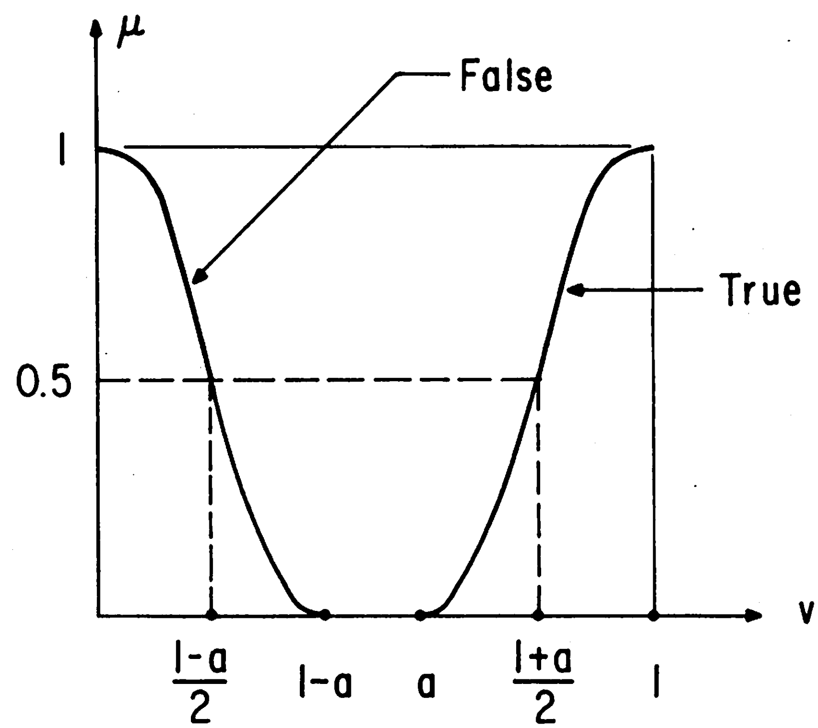


Fig.6.1. Compatibility functions of linguistic truth-values <u>true</u> and <u>false</u>.

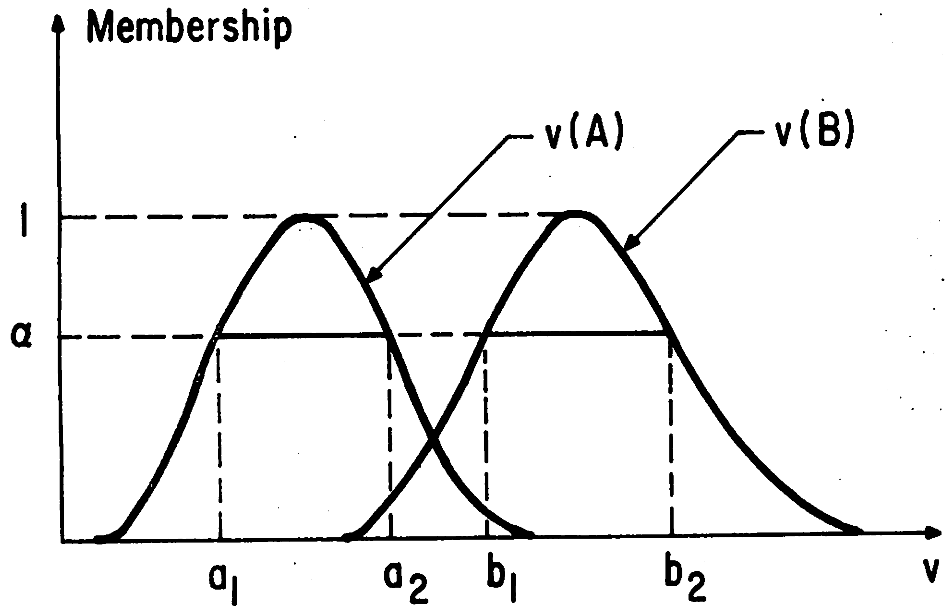


Fig.6.2. Level-sets of truth-values of A and B.

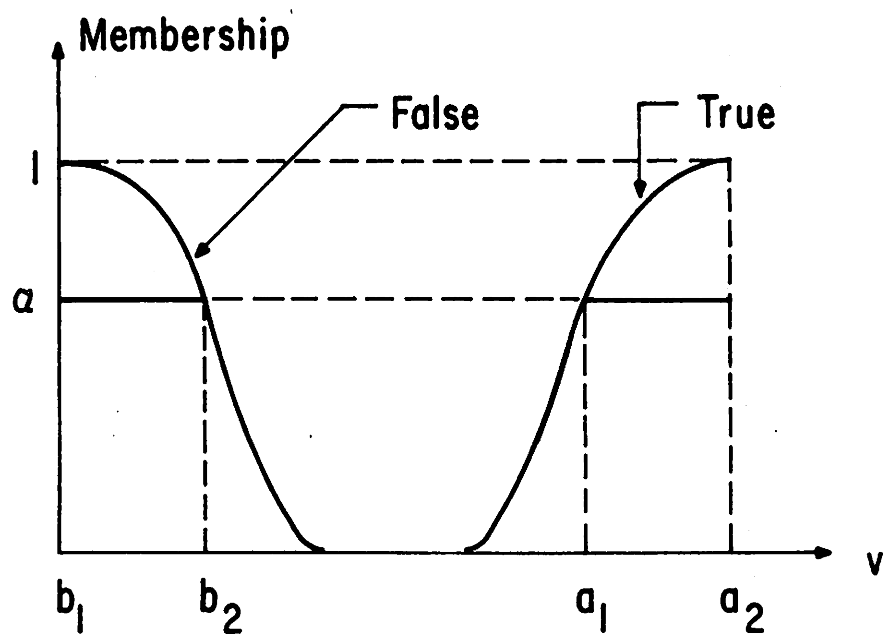


Fig.6.3. Computation of the truth-value of the conjunction of <u>true</u> and <u>false</u>.

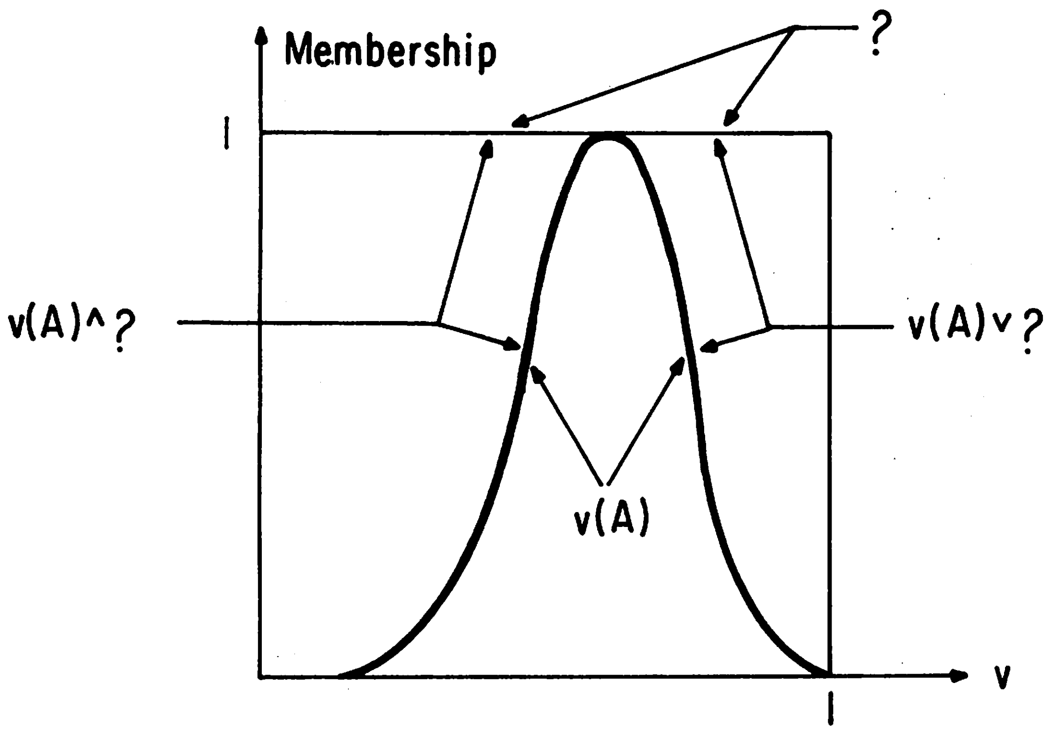


Fig.6.4. Conjunction and disjunction of the truth-value of A with the truth-value <u>unknown</u> ($\triangleq$?).

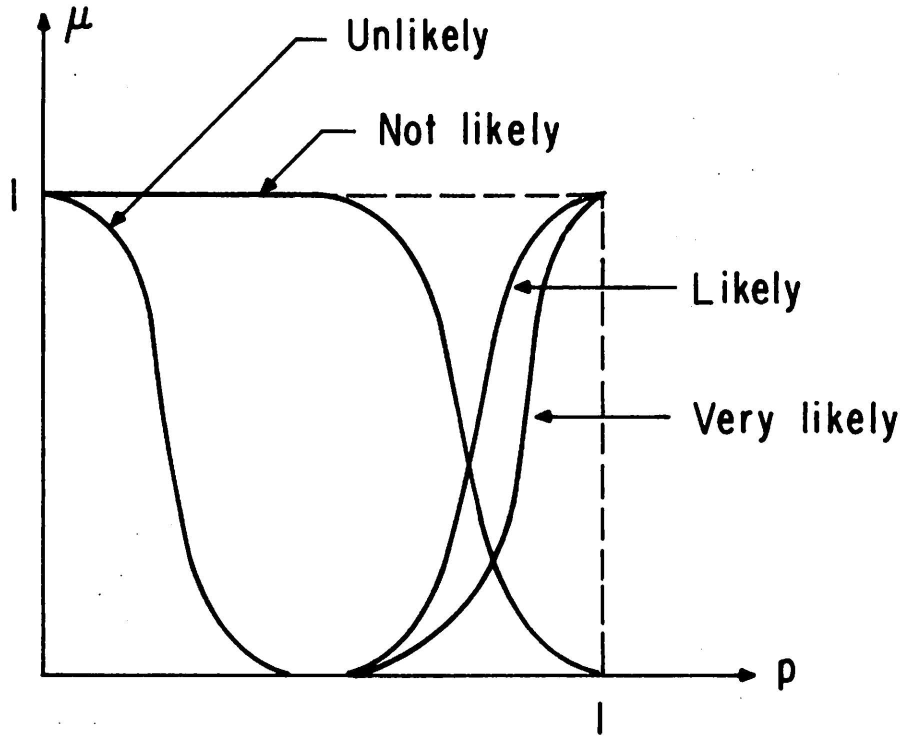


Fig.7.1. Compatibility functions of likely, not likely, unlikely and very likely.

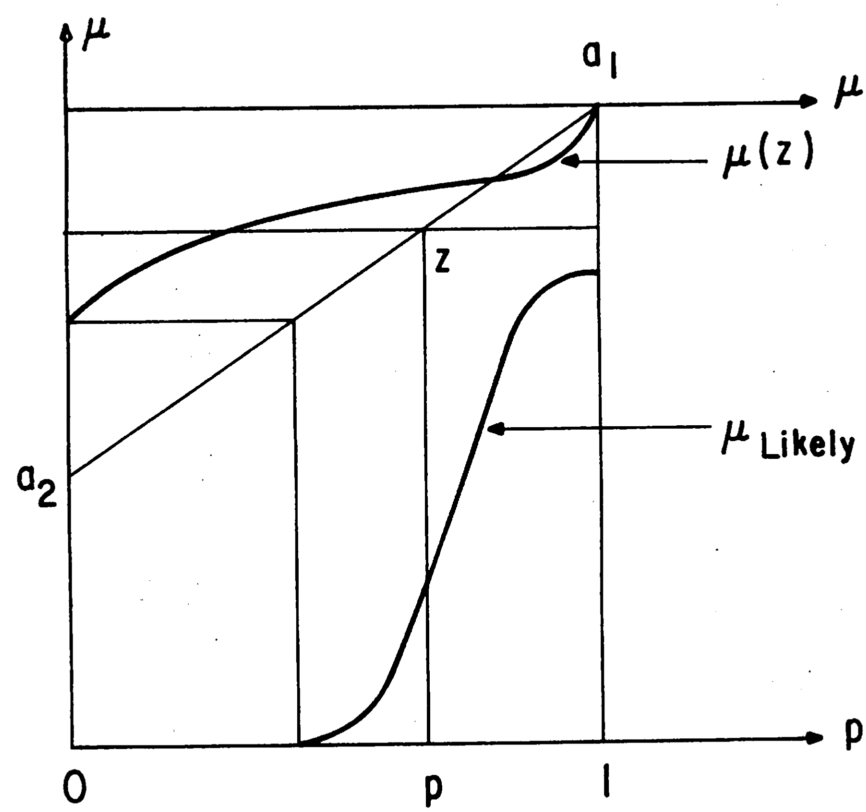


Fig.7.2. Computation of the linguistic value of $a_1p_1 + a_2p_2$.

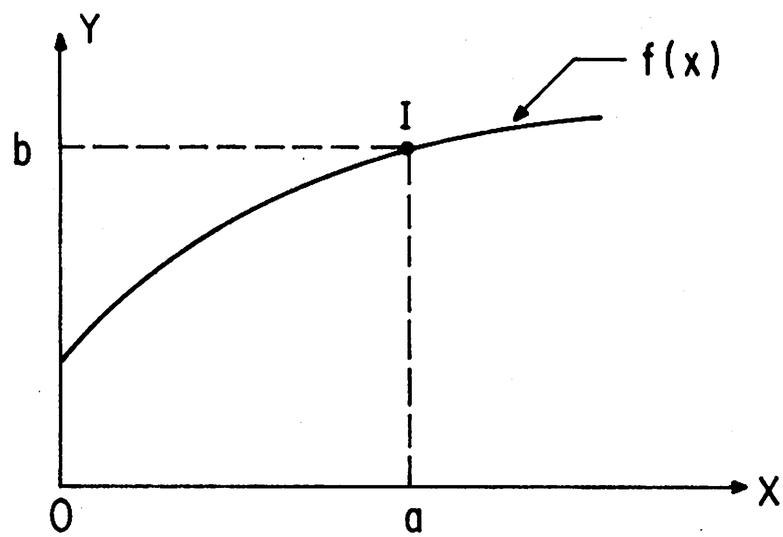


Fig. 8.1 Infering y = b from x = a and y = f(x).

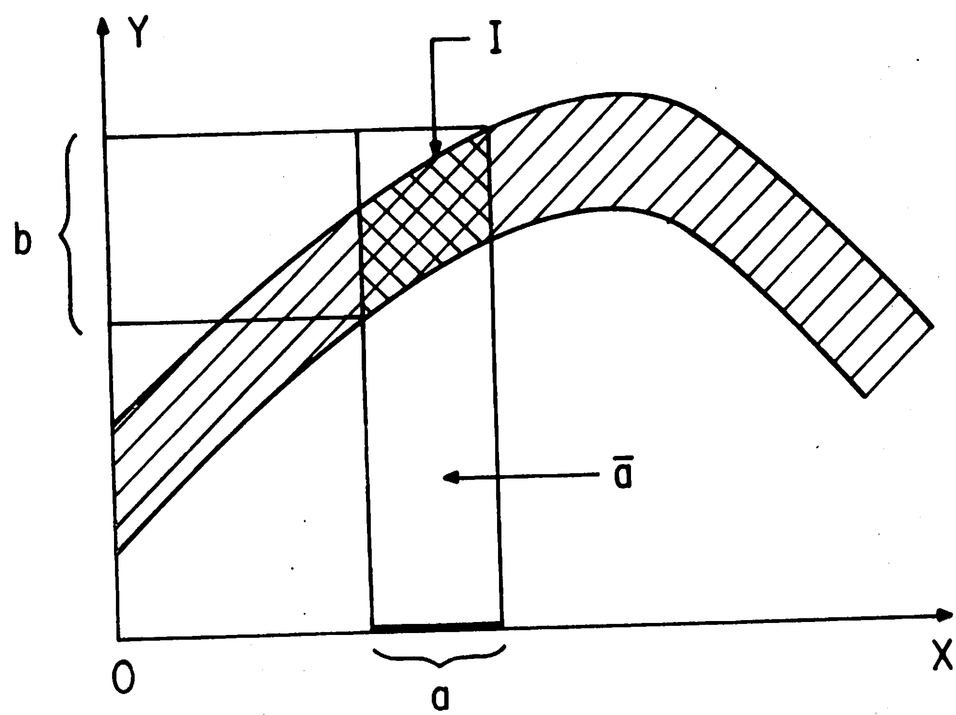


Fig.8.2. Illustration of the compositional rule of inference in the case of interval-valued variables.

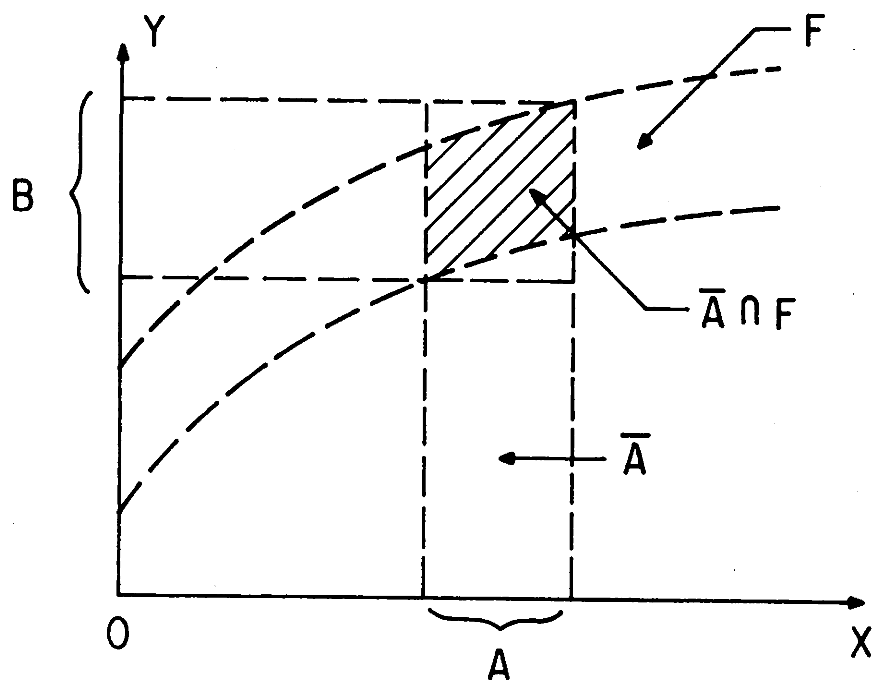


Fig. 8.3. Illustration of the compositional rule of inference for fuzzy variables.

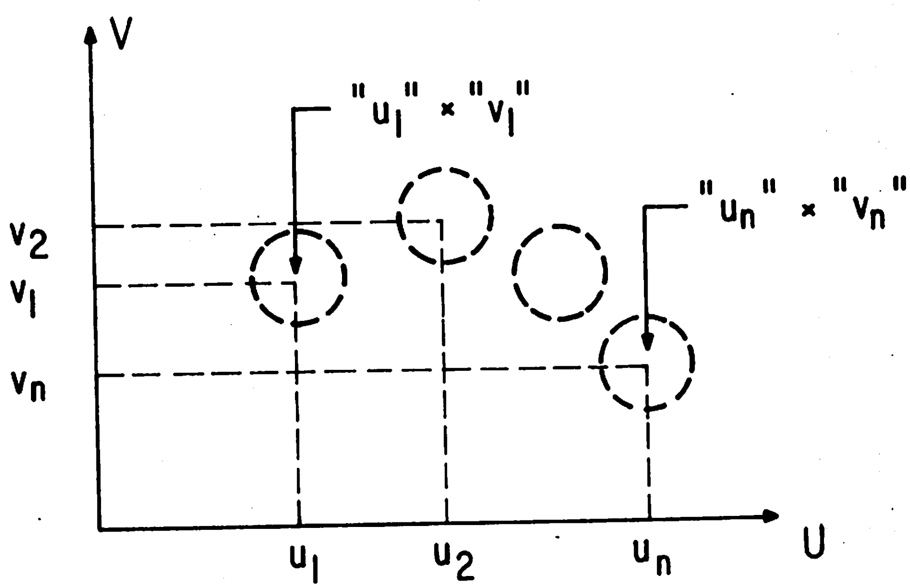


Fig.8.4. Representation of a fuzzy graph as a union of fuzzy points.

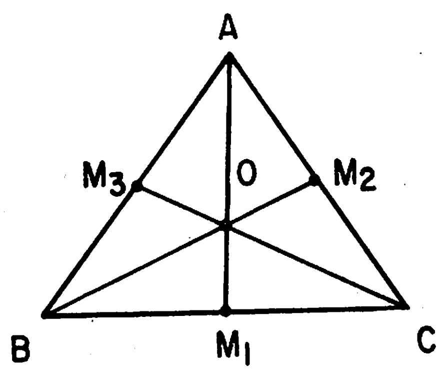


Fig.8.5. An elementary theorem in geometry.

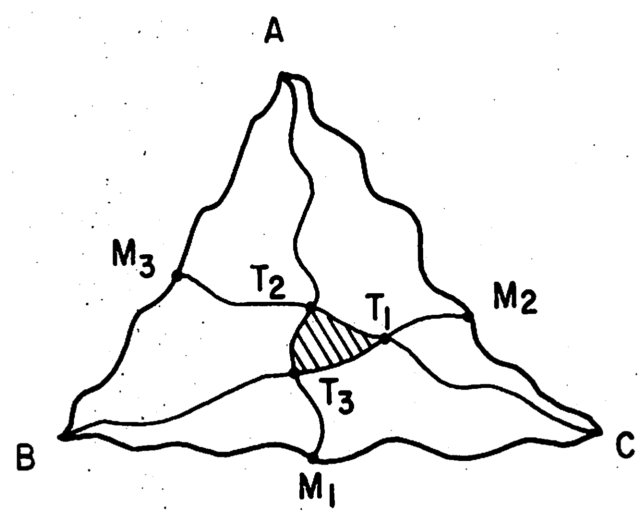


Fig.8.6. A fuzzy theorem in geometry.

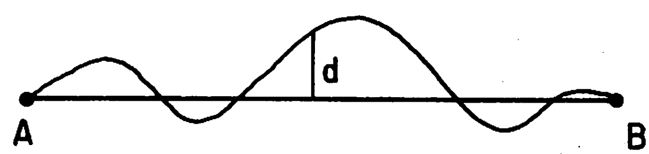


Fig.8.7. Definition of approximately straight line.

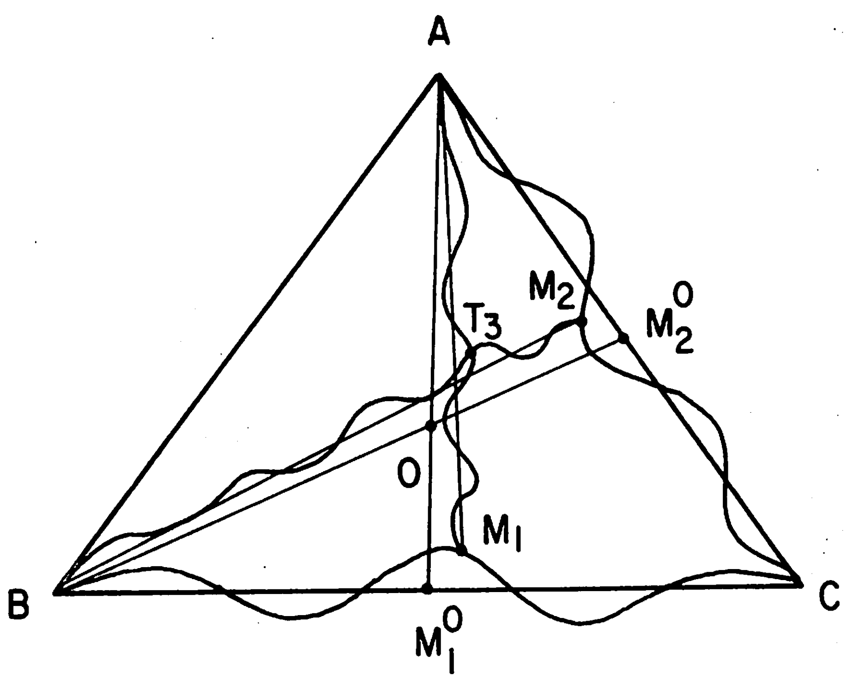


Fig.8.8. Illustration of an approximate proof of the fuzzy theorem.

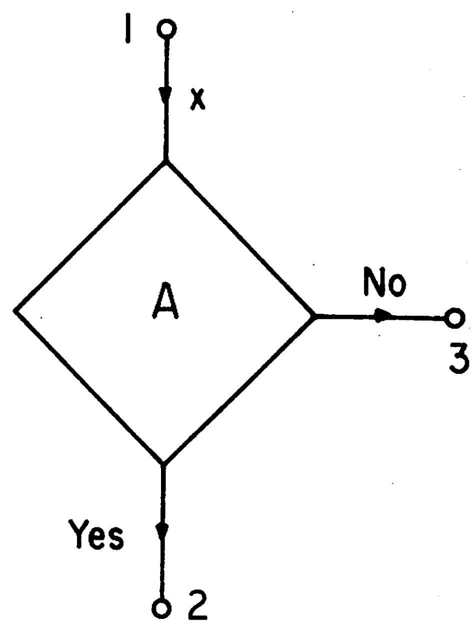


Fig.8.9. A fuzzy decision box.

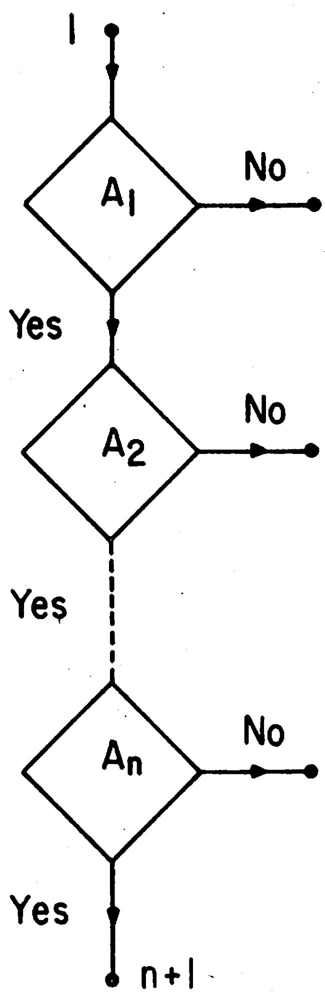


Fig. 8.10. A tandem combination of decision boxes.

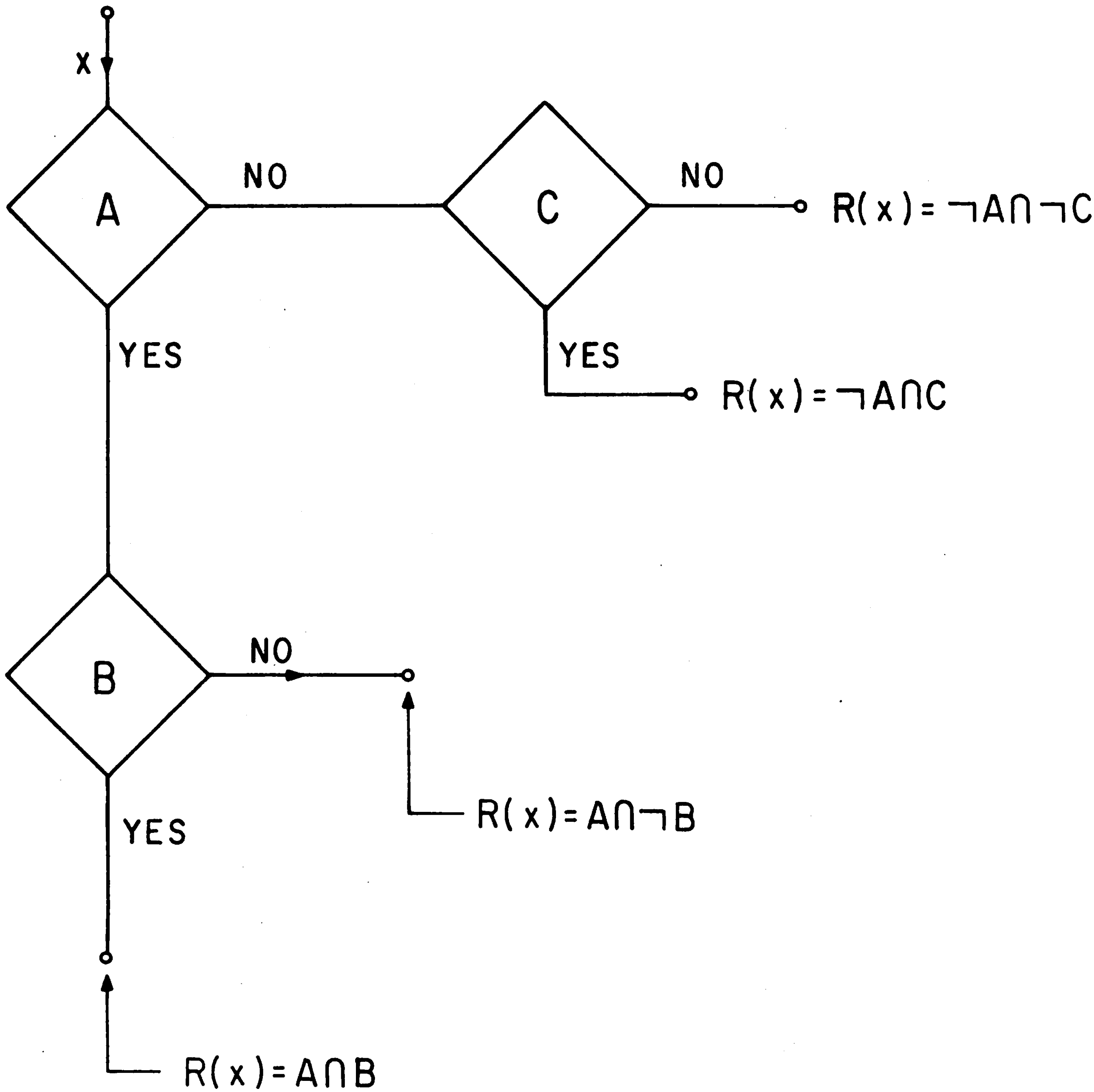


Fig. 8.11. Restrictions associated with various exits from a fuzzy flowchart.

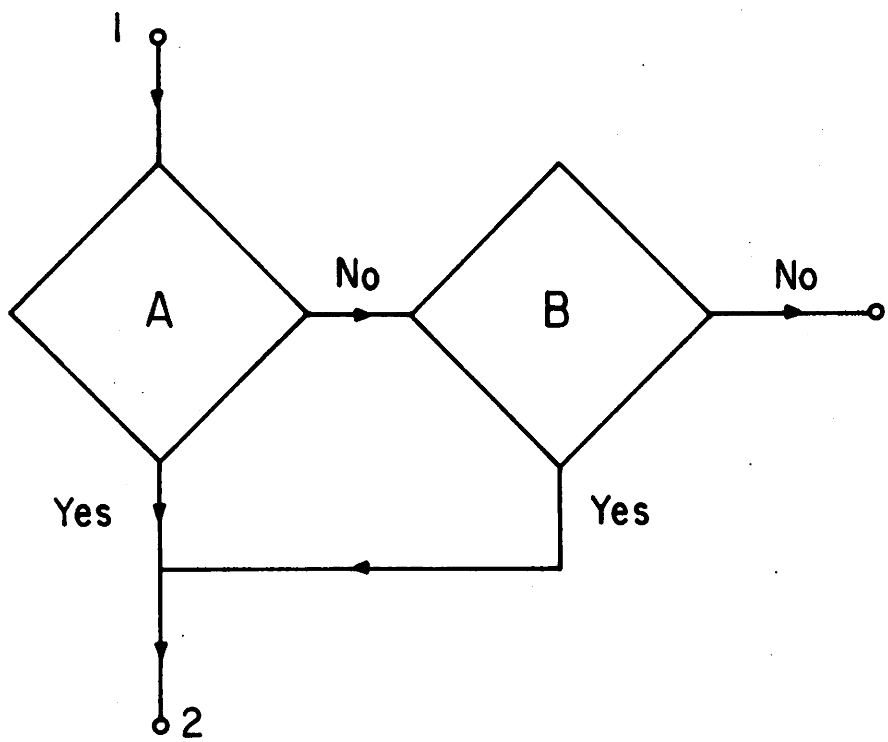


Fig.8.12. A graphical representation of the disjunction of fuzzy predicates.

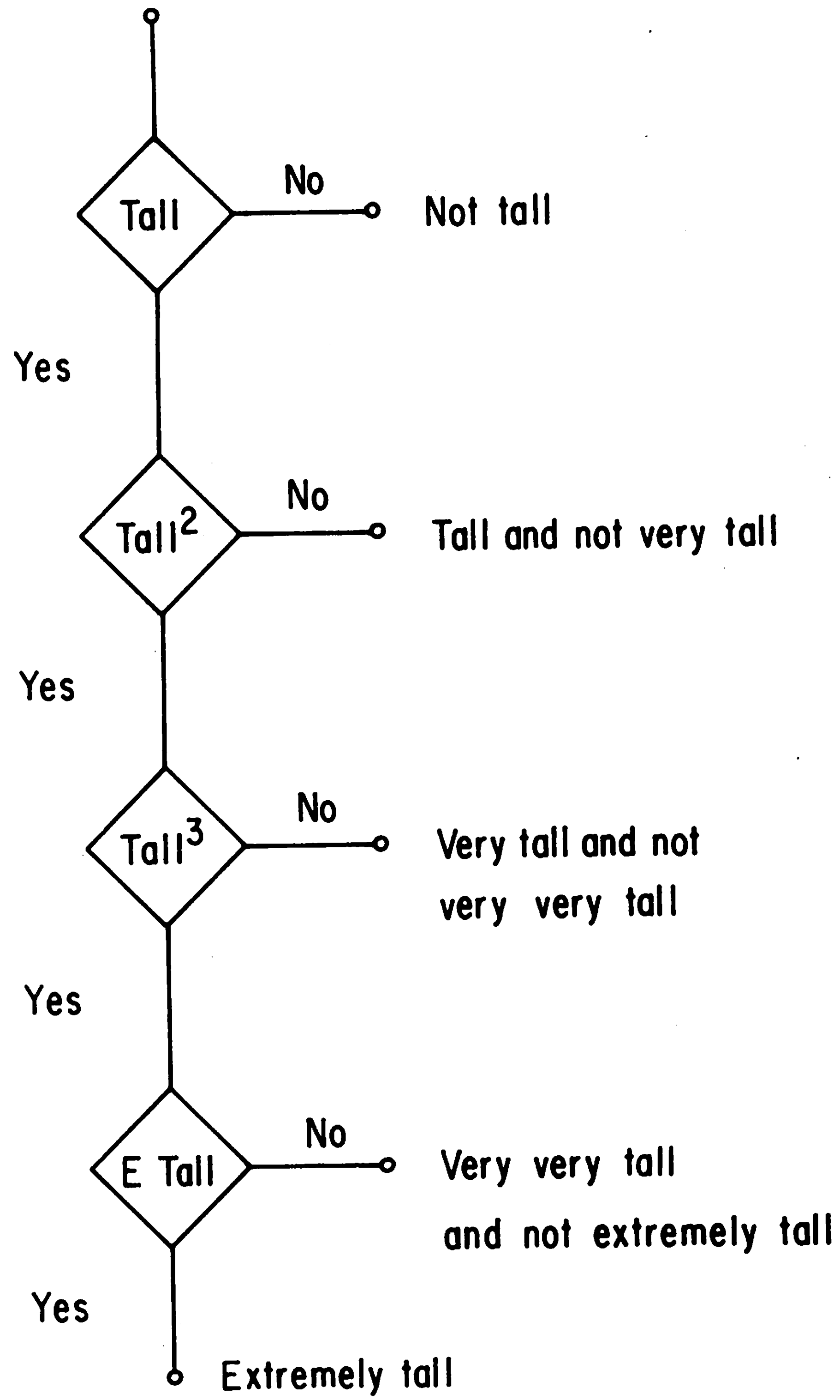


Fig. 8.13. Use of a tandem combination of decision boxes for purposes of bracketing.

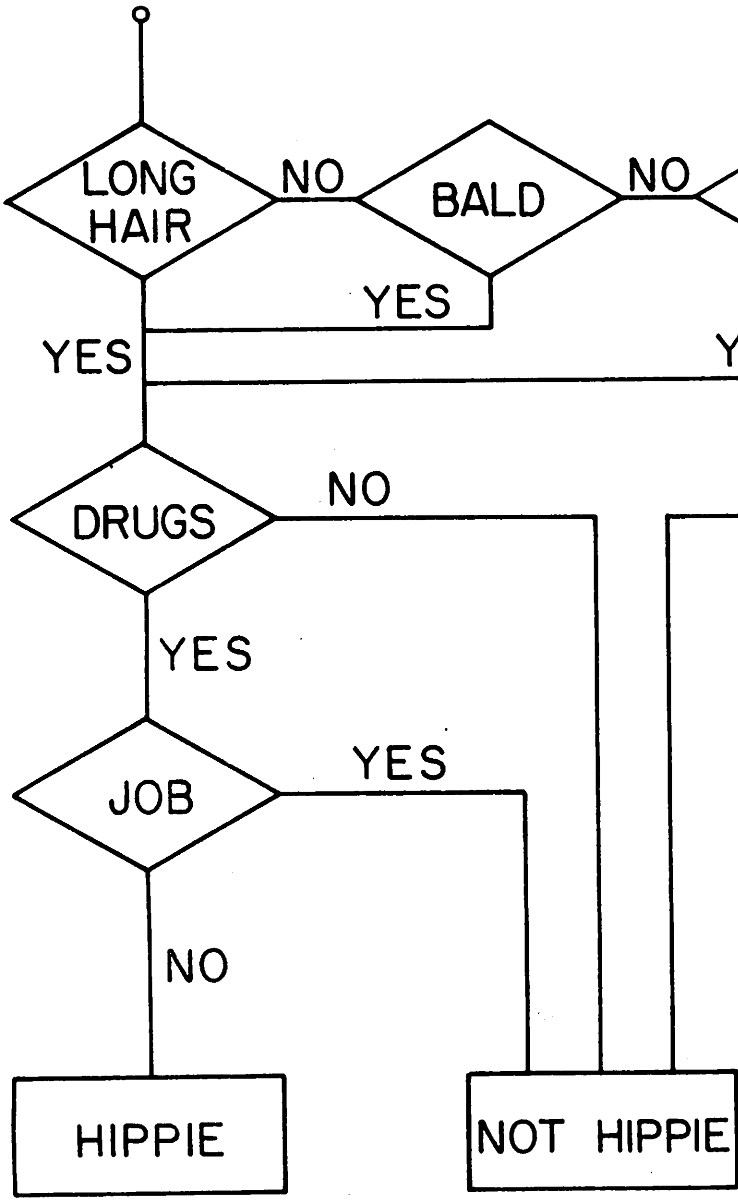


Fig. 8.14. Algorithmic definition of <u>Hippie</u> presented in the form of a fuzzy flowchart.

Captions

Fig.1.1. Compatibility function for *young*.

Fig.1.2. Assignment of linguistic values to attributes of John and x.

Fig.1.3. Hierarchical structure of a linguistic variable.

Fig.1.4. Compatibilities of *young*, *not young*, and *very young*.

Fig.1.5. (a) Compatibilities of *small*, *very small*, *large*, *very large* and *not very small and not very large*. (b) The problem of linguistic approximation is that of finding an approximate linguistic characterization of a given compatibility function.

Fig.2.1. Illustration of the valise analogy for a unary nonfuzzy variable.

Fig.2.2. Valise analogy for a binary nonfuzzy variable.

Fig.2.3. Marginal restrictions induced by $R(X_1,X_2)$.

Fig.2.4. (a) X_1 and X_2 are noninteractive.
(b) X_1 and X_2 are interactive.

Fig.2.5. $R(X_2 \mid u_1)$ is the restriction on u_2 conditioned on u_1.

Fig.3.1. Membership functions of *positive* and *slightly positive*.

Fig.3.2. R_1 is the base of the cylindrical set $\overline{R}_1$.

Fig.3.3. Relation between the cartesian product and intersection of cylindrical sets.

Table 3.4. Extension of the multiplication table to subsets of integers. $1 \vee 2$ means 1 or 2.

Fig.3.5. Intersection of fuzzy sets with interval-valued membership functions.

Fig.3.6. Level-sets of fuzzy membership functions μ_A and μ_B.