Fuzzy Sets and Systems 100 Supplement (1999) 9–34
North-Holland

FUZZY SETS AS A BASIS FOR A THEORY OF POSSIBILITY*

L.A. ZADEH

Computer Science Division, Department of Electrical Engineering and Computer Sciences and the Electronics Research Laboratory, University of California, Berkeley, CA 94720, U.S.A.

Received February 1977
Revised June 1977

The theory of possibility described in this paper is related to the theory of fuzzy sets by defining the concept of a possibility distribution as a fuzzy restriction which acts as an elastic constraint on the values that may be assigned to a variable. More specifically, if F is a fuzzy subset of a universe of discourse $U=\{u\}$ which is characterized by its membership function μ_F, then a proposition of the form "X is F," where X is a variable taking values in U, induces a possibility distribution Π_X which equates the possibility of X taking the value u to $\mu_F(u)$—the compatibility of u with F. In this way, X becomes a fuzzy variable which is associated with the possibility distribution Π_X in much the same way as a random variable is associated with a probability distribution. In general, a variable may be associated both with a possibility distribution and a probability distribution, with the weak connection between the two expressed as the possibility/probability consistency principle.

A thesis advanced in this paper is that the imprecision that is intrinsic in natural languages is, in the main, possibilistic rather than probabilistic in nature. Thus, by employing the concept of a possibility distribution, a proposition, p, in a natural language may be translated into a procedure which computes the probability distribution of a set of attributes which are implied by p. Several types of conditional translation rules are discussed and, in particular, a translation rule for propositions of the form "X is F is α-possible," where α is a number in the interval [0, 1], is formulated and illustrated by examples.

1. Introduction

The pioneering work of Wiener and Shannon on the statistical theory of communication has led to a universal acceptance of the belief that information is intrinsically statistical in nature and, as such, must be dealt with by the methods provided by probability theory.

Unquestionably, the statistical point of view has contributed deep insights into the fundamental processes involved in the coding, transmission and reception of data, and played a key role in the development of modern communication, detection and telemetering systems. In recent years, however, a number of other important applications have come to the fore in which the major issues center not on the

To Professor Arnold Kaufmann.

*Research supported by Naval Electronic Systems Command Contract N00039–77–C–0022, U.S. Army Research Office Contract DAHC04–75–G0056 and National Science Foundation Grant MCS76–06693.

Reprinted from Fuzzy Sets and Systems 1 (1978) 3–28

transmission of information but on its meaning. In such applications, what matters is the ability to answer questions relating to information that is stored in a database—as in natural language processing, knowledge representation, speech recognition, robotics, medical diagnosis, analysis of rare events, decision-making under uncertainty, picture analysis, information retrieval and related areas.

A thesis advanced in this paper is that when our main concern is with the meaning of information—rather than with its measure—the proper framework for information analysis is possibilistic[1] rather than probabilistic in nature, thus implying that what is needed for such an analysis is not probability theory but an analogous—and yet different—theory which might be called the *theory of possibility*.[2]

As will be seen in the sequel, the mathematical apparatus of the theory of fuzzy sets provides a natural basis for the theory of possibility, playing a role which is similar to that of measure theory in relation to the theory of probability. Viewed in this perspective, a fuzzy restriction may be interpreted as a possibility distribution, with its membership function playing the role of a possibility distribution function, and a fuzzy variable is associated with a possibility distribution in much the same manner as a random variable is associated with a probability distribution. In general, however, a variable may be associated both with a possibility distribution and a probability distribution, with the connection between the two expressible as the *possibility/probability consistency principle*. This principle—which is an expression of a weak connection between possibility and probability—will be described in greater detail at a later point in this paper.

The importance of the theory of possibility stems from the fact that—contrary to what has become a widely accepted assumption—much of the information on which human decisions are based is possibilistic rather than probabilistic in nature. In particular, the intrinsic fuzziness of natural languages—which is a logical consequence of the necessity to express information in a summarized form—is, in the main, possibilistic in origin. Based on this premise, it is possible to construct a universal language[3] in which the translation of a proposition expressed in a natural language takes the form of a procedure for computing the possibility distribution of a set of fuzzy relations in a data base. This procedure, then, may be interpreted as the meaning of the proposition in question, with the computed possibility distribution playing the role of the information which is conveyed by the proposition.

The present paper has the limited objective of exploring some of the elementary properties of the concept of a possibility distribution, motivated principally by the application of this concept to the representation of meaning in natural languages. Since our intuition concerning the properties of possibility distributions is not as yet well developed, some of the definitions which are formulated in the sequel should be viewed as provisional in nature.

[1]The term possibilistic—in the sense used here—was coined by Gaines and Kohout in their paper on possible automata [1].

[2]The interpretation of the concept of possibility in the theory of possibility is quite different from that of modal logic [2] in which propositions of the form "It is possible that..." and "It is necessary that..." are considered.

[3]Such a language, called PRUF (Possibilistic Relational Universal Fuzzy), is described in [30].

2. The concept of a possibility distribution

What is a possibility distribution? It is convenient to answer this question in terms of another concept, namely, that of a *fuzzy restriction* [4, 5], to which the concept of a possibility distribution bears a close relation.

Let X be a variable which takes values in a universe of discourse U, with the generic element of U denoted by u and

$$X = u \tag{2.1}$$

signifying that X is assigned the value u, $u \in U$.

Let F be a fuzzy subset of U which is characterized by a membership function μ_F. Then F is a *fuzzy restriction on* X (or *associated with* X) if F acts as an elastic constraint on the values that may be assigned to X—in the sense that the assignment of a value u to X has the form

$$X = u : \mu_F(u) \tag{2.2}$$

where $\mu_F(u)$ is interpreted as the degree to which the constraint represented by F is satisfied when u is assigned to X. Equivalently, (2.2) implies that $1-\mu_F(u)$ is the degree to which the constraint in question must be stretched in order to allow the assignment of u to X.[4]

Let $R(X)$ denote a fuzzy restriction associated with X. Then, to express that F plays the role of a fuzzy restriction in relation to X, we write

$$R(X) = F. \tag{2.3}$$

An equation of this form is called a *relational assignment equation* because it represents the assignment of a fuzzy set (or a fuzzy relation) to the restriction associated with X.

To illustrate the concept of a fuzzy restriction, consider a proposition of the form $p \triangleq X$ is F,[5] where X is the name of an object, a variable or a proposition, and F is the name of a fuzzy subset of U, as in "Jessie is very intelligent," "X is a small number," "Harriet is blonde is quite true," etc. As shown in [4] and [6], the translation of such a proposition may be expressed as

$$R(A(X)) = F \tag{2.4}$$

where $A(X)$ is an implied attribute of X which takes values in U, and (2.4) signifies that the proposition $p \triangleq X$ is F has the effect of assigning F to the fuzzy restriction on the values of $A(X)$.

As a simple example of (2.4), let p be the proposition "John is young," in which young is a fuzzy subset of $U = [0, 100]$ characterized by the membership function

$$\mu_{\text{young}}(u) = 1 - S(u; 20, 30, 40) \tag{2.5}$$

[4] A point that must be stressed is that a fuzzy set *per sé* is not a fuzzy restriction. To be a fuzzy restriction, it must be acting as a constraint on the values of a variable.

[5] The symbol $\triangleq$ stands for "denotes" or "is defined to be".

where u is the numerical age and the S-function is defined by [4].

$$\begin{aligned} S(u;\alpha,\beta,\gamma) &= 0 && \text{for } u \leqq \alpha \\ &= 2\left(\frac{u-\alpha}{\gamma-\alpha}\right)^2 && \text{for } \alpha \leqq u \leqq \beta \\ &= 1-2\left(\frac{u-\gamma}{\gamma-\alpha}\right)^2 && \text{for } \beta \leqq u \leqq \gamma \\ &= 1 && \text{for } u \geqq \gamma, \end{aligned} \tag{2.6}$$

in which the parameter $\beta \triangleq (\alpha+\gamma)/2$ is the crossover point, that is, $S(\beta;\alpha,\beta,\gamma)=0.5$. In this case, the implied attribute $A(X)$ is Age(John) and the translation of "John is young" assumes the form:

$$\text{John is young} \to R(\text{Age(John)}) = \text{young}. \tag{2.7}$$

To relate the concept of a fuzzy restriction to that of a possibility distribution, we interpret the right-hand member of (2.7) in the following manner.

Consider a numerical age, say $u=28$, whose grade of membership in the fuzzy set young (as defined by (2.5)) is approximately 0.7. First, we interpret 0.7 as the degree of *compatibility* of 28 with the concept labeled young. Then, we postulate that the proposition "John is young" converts the meaning of 0.7 from the degree of compatibility of 28 with young to the degree of possibility that John is 28 given the proposition "John is young." In short, the compatibility of a value of u with young becomes converted into the possibility of that value of u given "John is young."

Stated in more general terms, the concept of a possibility distribution may be defined as follows. (For simplicity, we assume that $A(X)=X$.)

Definition 2.1. Let F be a fuzzy subset of a universe of discourse U which is characterized by its membership function μ_F, with the grade of membership, $\mu_F(u)$, interpreted as the compatibility of u with the concept labeled F.

Let X be a variable taking values in U, and let F act as a fuzzy restriction, $R(X)$, associated with X. Then the proposition "X is F," which translates into

$$R(X)=F, \tag{2.8}$$

associates a *possibility distribution*, Π_X, *with* X which is postulated to be equal to $R(X)$, i.e.,

$$\Pi_X = \mathrm{R(X)}. \tag{2.9}$$

Correspondingly, the *possibility distribution function associated with* X (or the possibility distribution function of Π_X) is denoted by π_X and is defined to be numerically equal to the membership function of F, i.e.,

$$\pi_X \triangleq \mu_F. \tag{2.10}$$

Thus, $\pi_X(u)$, the *possibility* that $X=u$, is postulated to be equal to $\mu_F(u)$.

In view of (2.9), the relational assignment equation (2.8) may be expressed equivalently in the form

$$\Pi_X = F, \tag{2.11}$$

placing in evidence that the proposition $p \triangleq X$ is F has the effect of associating X with a possibility distribution Π_X which, by (2.9), is equal to F. When expressed in the form of (2.11), a relational assignment equation will be referred to as a *possibility assignment equation*, with the understanding that Π_X is *induced by p*.

As a simple illustration, let U be the universe of positive integers and let F be the fuzzy set of small integers defined by ($+ \triangleq$ union)

$$\text{small integer} = 1/1 + 1/2 + 0.8/3 + 0.6/4 + 0.4/5 + 0.2/6.$$

Then, the proposition "X is a small integer" associates with X the possibility distribution

$$\Pi_X = 1/1 + 1/2 + 0.8/3 + 0.6/4 + 0.4/5 + 0.2/6 \tag{2.12}$$

in which a term such as 0.8/3 signifies that the possibility that X is 3, given that X is a small integer, is 0.8.

There are several important points relating to the above definition which are in need of comment.

First, (2.9) implies that the possibility distribution Π_X may be regarded as an interpretation of the concept of a fuzzy restriction and, consequently, that the mathematical apparatus of the theory of fuzzy sets—and, especially, the calculus of fuzzy restrictions [4]—provides a basis for the manipulation of possibility distributions by the rules of this calculus.

Second, the definition implies the assumption that our intuitive perception of the ways in which possibilities combine is in accord with the rules of combination of fuzzy restrictions. Although the validity of this assumption cannot be proved at this juncture, it appears that there is a fairly close agreement between such basic operations as the union and intersection of fuzzy sets, on the one hand, and the possibility distributions associated with the disjunctions and conjunctions of propositions of the form "X is F." However, since our intuition concerning the behaviour of possibilities is not very reliable, a great deal of empirical work would have to be done to provide us with a better understanding of the ways in which possibility distributions are manipulated by humans. Such an understanding would be enhanced by the development of an axiomatic approach to the definition of subjective possibilities—an approach which might be in the spirit of the axiomatic approaches to the definition of subjective probabilities [7, 8].

Third, the definition of $\pi_X(u)$ implies that the degree of possibility may be any number in the interval [0, 1] rather than just 0 or 1. In this connection, it should be noted that the existence of intermediate degrees of possibility is implicit in such commonly encountered propositions as "There is a slight possibility that Marilyn is very rich," "It is quite possible that Jean-Paul will be promoted," "It is almost impossible to find a needle in a haystack," etc.

It could be argued, of course, that a characterization of an intermediate degree of possibility by a label such as "slight possibility" is commonly meant to be interpreted as "slight probability." Unquestionably, this is frequently the case in everyday discourse. Nevertheless, there is a fundamental difference between probability and possibility which, once better understood, will lead to a more careful differentiation between the characterizations of degrees of possibility vs. degrees of probability—especially in legal discourse, medical diagnosis, synthetic languages and, more generally, those applications in which a high degree of precision of meaning is an important desideratum.

To illustrate the difference between probability and possibility by a simple example, consider the statement "Hans ate X eggs for breakfast," with X taking values in $U = \{1, 2, 3, 4, \ldots\}$. We may associate a possibility distribution with X by interpreting $\pi_X(u)$ as the degree of ease with which Hans can eat u eggs. We may also associate a probability distribution with X by interpreting $P_X(u)$ as the probability of Hans eating u eggs for breakfast. Assuming that we employ some explicit or implicit criterion for assessing the degree of ease with which Hans can eat u eggs for breakfast, the values of $\pi_X(u)$ and $p_X(u)$ might be as shown in Table 1.

Table 1

The possibility and probability distributions associated with X

u	1	2	3	4	5	6	7	8
$\pi_X(u)$	1	1	1	1	0.8	0.6	0.4	0.2
$P_X(u)$	0.1	0.8	0.1	0	0	0	0	0

We observe that, whereas the possibility that Hans may eat 3 eggs for breakfast is 1, the probability that he may do so might be quite small, e.g., 0.1. Thus, a high degree of possibility does not imply a high degree of probability, nor does a low degree of probability imply a low degree of possibility. However, if an event is impossible, it is bound to be improbable. This heuristic connection between possibilities and probabilities may be stated in the form of what might be called the *possibility/probability consistency principle*, namely:

If a variable X can take the values $u_1,\ldots,u_n$ with respective possibilities $\Pi = (\pi_1,\ldots,\pi_n)$ and probabilities $P = (p_1,\ldots,p_n)$, then the degree of consistency of the probability distribution P with the possibility distribution Π is expressed by ($+ \triangleq$ arithmetic sum)

$$\gamma = \pi_1 p_1 + \cdots + \pi_n p_n. \tag{2.13}$$

It should be understood, of course, that the possibility/probability consistency principle is not a precise law or a relationship that is intrinsic in the concepts of possibility and probability. Rather it is an approximate formalization of the heuristic observation that a lessening of the possibility of an event tends to lessen its probability—but not vice-versa. In this sense, the principle is of use in situations in which what is known about a variable X is its possibility—rather than its probability—distribution. In such cases—which occur far more frequently than those in which the reverse is true—the possibility/probability consistency principle provides a basis for the computation of the possibility distribution of the probability distribution of X. Such computations play a particularly important role in decision-making under uncertainty and in the theories of evidence and belief [9–12].

In the example discussed above, the possibility of X assuming a value u is interpreted as the degree of ease with which u may be assigned to X, e.g., the degree of ease with which Hans may eat u eggs for breakfast. It should be understood, however, that this "degree of ease" may or may not have physical reality. Thus, the proposition "John is young" induces a possibility distribution whose possibility distribution function is expressed by (2.5). In this case, the possibility that the variable Age(John) may take the value 28 is 0.7, with 0.7 representing the degree of ease with which 28 may be assigned to Age(John) given the elasticity of the fuzzy restriction labeled young. Thus, in this case "the degree of ease" has a figurative rather than physical significance.

2.1. Possibility measure

Additional insight into the distinction between probability and possibility may be gained by comparing the concept of a *possibility measure* with the familiar concept of a probability measure. More specifically, let A be a nonfuzzy subset of U and let Π_X be a possibility distribution associated with a variable X which takes values in U. Then, the *possibility measure*, $\pi(A)$, of A is defined as a number in [0, 1] given by[6]

$$\pi(A) \triangleq \mathrm{Sup}_{u \in A}\, \pi_X(u), \tag{2.14}$$

where $\pi_X(u)$ is the possibility distribution function of Π_X. This number, then, may be interpreted as the possibility *that a value of X belongs to A*, that is

$$\begin{aligned}\mathrm{Poss}\{X \in A\} &\triangleq \pi(A) \\ &\triangleq \mathrm{Sup}_{u \in A}\, \pi_X(u).\end{aligned} \tag{2.15}$$

When A is a fuzzy set, the belonging of a value of X to A is not meaningful. A more general definition of possibility measure which extends (2.15) to fuzzy sets is the following.

Definition 2.2. Let A be a fuzzy subset of U and let Π_X be a possibility distribution associated with a variable X which takes values in U. The *possibility measure*, $\pi(A)$, of A is defined by

$$\begin{aligned}\mathrm{Poss}\{X \text{ is } A\} &\triangleq \pi(A) \\ &\triangleq \mathrm{Sup}_{u \in U}\, \mu_A(u) \wedge \pi_X(u),\end{aligned} \tag{2.16}$$

[6]The measure defined by (2.14) may be viewed as a particular case of the fuzzy measure defined by Sugeno and Terano [20, 21]. Furthermore, $\pi(A)$ as defined by (2.14) provides a possibilistic interpretation for the *scale function*, $\sigma(A)$, which is defined by Nahmias [22] as the supremum of a membership function over a nonfuzzy set A.

where "X is A" replaces "$X \in A$" in (2.15), μ_A is the membership function of A, and $\wedge$ stands, as usual, for min. It should be noted that, in terms of the *height* of a fuzzy set, which is defined as the supremum of its membership function [23], (2.16) may be expressed compactly by the equation

$$\pi(A) \triangleq \text{Height}(A \cap \Pi_X). \tag{2.17}$$

As a simple illustration, consider the possibility distribution (2.12) which is induced by the proposition "X is a small integer." In this case, let A be the set $\{3, 4, 5\}$. Then

$$\pi(A) = 0.8 \vee 0.6 \vee 0.4 = 0.8,$$

where $\vee$ stands, as usual, for max.

On the other hand, if A is the fuzzy set of integers which are not small, i.e.,

$$A \triangleq 0.2/3 + 0.4/4 + 0.6/5 + 0.8/6 + 1/7 + \cdots \tag{2.18}$$

then

$$\begin{aligned} \text{Poss}\{X \text{ is not a small integer}\} &= \text{Height}(0.2/3 + 0.4/4 + 0.4/5 + 0.2/6) \\ &= 0.4. \end{aligned} \tag{2.19}$$

It should be noted that (2.19) is an immediate consequence of the assertion

$$X \text{ is } F \Rightarrow \text{Poss}\{X \text{ is } A\} = \text{Height}(F \cap A), \tag{2.20}$$

which is implied by (2.11) and (2.17). In particular, if A is a normal fuzzy set (i.e., $\text{Height}(A) = 1$), then, as should be expected

$$X \text{ is } A \Rightarrow \text{Poss}\{X \text{ is } A\} = 1. \tag{2.21}$$

Let A and B be arbitrary fuzzy subsets of U. Then, from the definition of the possibility measure of a fuzzy set (2.16), it follows that[7]

$$\pi(A \cup B) = \pi(A) \vee \pi(B). \tag{2.22}$$

By comparison, the corresponding relation for probability measures of A, B and $A \cup B$ (if they exist) is

$$P(A \cup B) \leqq P(A) + P(B) \tag{2.23}$$

and, if A and B are disjoint (i.e., $\mu_A(u)\mu_B(u) \equiv 0$),

$$P(A \cup B) = P(A) + P(B), \tag{2.24}$$

[7]It is of interest that (2.22) is analogous to the extension principle for fuzzy sets [5], with + (union) in the right-hand side of the statement of the principle replaced by $\vee$.

which expresses the basic additivity property of probability measures. Thus, in contrast to probability measure, possibility measure is not additive. Instead, it has the property expressed by (2.22), which may be viewed as an analog of (2.24) with $+$ replaced by $\vee$.

In a similar fashion, the possibility measure of the intersection of A and B is related to those of A and B by

$$\pi(A \cap B) \leqq \pi(A) \wedge \pi(B). \tag{2.25}$$

In particular, if A and B are *noninteractive*,[8] (2.25) holds with the equality sign, i.e.,

$$\pi(A \cap B) = \pi(A) \wedge \pi(B). \tag{2.26}$$

By comparison, in the case of probability measures, we have

$$P(A \cap B) \leqq P(A) \wedge P(B) \tag{2.27}$$

and

$$P(A \cap B) = P(A)P(B) \tag{2.28}$$

if A and B are independent and nonfuzzy. As in the case of (2.22), (2.26) is analogous to (2.28), with product corresponding to min.

2.2. *Possibility and information*

If p is a proposition of the form $p \triangleq X$ is F which translates into the possibility assignment equation

$$\Pi_{A(X)} = F, \tag{2.29}$$

where F is a fuzzy subset of U and $A(X)$ is an implied attribute of X taking values in U, then the *information conveyed by* p, $I(p)$, may be identified with the possibility distribution, $\Pi_{A(X)}$, of the fuzzy variable $A(X)$. Thus, the connection between $I(p)$, $\Pi_{A(X)}$, $R(A(X))$ and F is expressed by

$$I(p) \triangleq \Pi_{A(X)}, \tag{2.30}$$

where

$$\Pi_{A(X)} = R(A(X)) = F. \tag{2.31}$$

For example, if the proposition $p \triangleq$ John is young translates into the possibility assignment equation

$$\Pi_{\text{Age(John)}} = \text{young}, \tag{2.32}$$

[8]Noninteraction in the sense defined here is closely related to the concept of noninteraction of fuzzy restrictions [5, 6]. It should also be noted that (2.26) provides a possibilistic interpretation for "unrelatedness' as defined by Nahmias [22].

where young is defined by (2.5), then

$$I(\text{John is young}) = \Pi_{\text{Age(John)}} \tag{2.33}$$

in which the possibility distribution function of Age(John) is given by

$$\pi_{\text{Age(John)}}(u) = 1 - S(u;20,30,40), \qquad u \in [0, 100]. \tag{2.34}$$

From the definition of $I(p)$ it follows that if $p \triangleq X$ is F and $q \triangleq X$ is G, then p is at least as informative as q, expressed as $I(p) \geqq I(q)$, if $F \subset G$. Thus, we have a partial ordering of the $I(p)$ defined by

$$F \subset G \Rightarrow I(X \text{ is } F) \geqq I(X \text{ is } G) \tag{2.35}$$

which implies that the more restrictive a possibility distribution is, the more informative is the proposition with which it is associated. For example, since very tall $\subset$ tall, we have

$$I(\text{Lucy is very tall}) \geqq I(\text{Lucy is tall}). \tag{2.36}$$

3. *N*-ary possibility distributions

In asserting that the translation of a proposition of the form $p \triangleq X$ is F is expressed by

$$X \text{ is } F \rightarrow R(A(X)) = F \tag{3.1}$$

or, equivalently,

$$X \text{ is } F \rightarrow \Pi_{A(X)} = F, \tag{3.2}$$

we are tacitly assuming that p contains a single implied attribute $A(X)$ whose possibility distribution is given by the right-hand member of (3.2).

More generally, p may contain n implied attributes $A_1(X), \ldots, A_n(X)$, with $A_i(X)$ taking values in U_i, $i = 1, \ldots, n$. In this case, the translation of $p \triangleq X$ is F, where F is a fuzzy relation in the cartesian product $U = U_1 \times \cdots \times U_n$, assumes the form

$$X \text{ is } F \rightarrow R(A_1(X), \ldots, A_n(X)) = F \tag{3.3}$$

or, equivalently,

$$X \text{ is } F \rightarrow \Pi_{(A_1(X), \ldots, A_n(X))} = F \tag{3.4}$$

where $R(A_1(X), \ldots, A_n(X))$ is an n-ary fuzzy restriction and $\Pi_{(A_1(X), \ldots, A_n(X))}$ is an *n-ary possibility distribution* which is *induced by p*. Correspondingly, the *n-ary possibility distribution function induced by p* is given by

$$\pi_{(A_1(X), \ldots, A_n(X))}(u_1, \ldots, u_n) = \mu_F(u_1, \ldots, u_n), \qquad (u_1, \ldots, u_n) \in U, \tag{3.5}$$

where μ_F is the membership function of F. In particular, if F is a cartesian product of n

unary fuzzy relations $F_1, \ldots, F_n$, then the righthand member of (3.3) decomposes into a system of n unary relational assignment equations, i.e.,

$$X \text{ is } F \to R(A_1(X) = F_1 \qquad (3.6)$$
$$R(A_2(X)) = F_2$$
$$\vdots \qquad \vdots$$
$$R(A_n(X)) = F_n.$$

Correspondingly,[9]

$$\Pi_{(A_1(X), \ldots, A_n(X))} = \Pi_{A_1(X)} \times \cdots \times \Pi_{A_n(X)} \qquad (3.7)$$

and

$$\pi_{(A_1(X), \ldots, A_n(X))}(u_1, \ldots, u_n) = \pi_{A_1(X)}(u_1) \wedge \cdots \wedge \pi_{A_n(X)}(u_n), \qquad (3.8)$$

where

$$\pi_{A_i(X)}(u_i) = \mu_{F_i}(u_i), \qquad u_i \in U_i, \qquad i = 1, \ldots, n \qquad (3.9)$$

and $\wedge$ denotes min (in infix form).

As a simple illustration, consider the proposition $p \triangleq$ carpet is large, in which large is a fuzzy relation whose tableau is of the form shown in Table 2 (with length and width expressed in metric units).

Table 2

Tableau of large

Large	Width	Length	μ
	250	300	0.6
	250	350	0.7
	.	.	.
	300	400	0.8
	.	.	.
	400	600	1

In this case, the translation (3.3) leads to the possibility assignment equation

$$\Pi_{(\text{width(carpet)}, \text{length(carpet)})} = \text{large}, \qquad (3.10)$$

which implies that if the compatibility of a carpet whose width is, say, 250 cm and length is 350 cm with "large carpet" is 0.7, then the possibility that the width of the carpet is 250 cm and its length is 350 cm—given the proposition $p \triangleq$ carpet is large—is 0.7.

Now, if large is defined as

$$\text{large} = \text{wide} \times \text{long} \qquad (3.11)$$

[9] If F and G are fuzzy relations in U and V, respectively, then their cartesian product $F \times G$ is a fuzzy relation in $U \times V$ whose membership function is given by $\mu_F(u) \wedge \mu_G(v)$.

where long and wide are unary fuzzy relations, then (3.10) decomposes into the possibility association equations

$$\Pi_{\text{width(carpet)}} = \text{wide}$$

and

$$\Pi_{\text{length(carpet)}} = \text{long}$$

where the tableaux of long and wide are of the form shown in Table 3.

Table 3

Tableaux of wide and long

Wide	Width	μ	Long	Length	μ
	250	0.7		300	0.6
	300	0.8		350	0.7
	350	0.8		400	0.8
	.	.		.	.
	400	1		500	1

3.1. *Marginal possibility distributions*

The concept of a marginal possibility distribution bears a close relation to the concept of a marginal fuzzy restriction [4], which in turn is analogous to the concept of a marginal probability distribution.

More specifically, let $X = (X_1, \ldots, X_n)$ be an n-ary fuzzy variable taking values in $U = U_1 \times \cdots \times U_n$, and let Π_X be a possibility distribution associated with X, with $\pi_X(u_1, \ldots, u_n)$ denoting the possibility distribution function of Π_X.

Let $q \triangleq (i_1, \ldots, i_k)$ be a subsequence of the index sequence $(1, \ldots, n)$ and let $X_{(q)}$ be the q-ary fuzzy variable $X_{(q)} \triangleq (X_{i_1}, \ldots, X_{i_k})$. The *marginal possibility distribution* $\Pi_{X_{(q)}}$ is a possibility distribution associated with $X_{(q)}$ which is *induced* by Π_X as the projection of Π_X on $U_{(q)} \triangleq U_{i_1} \times \cdots \times U_{i_k}$. Thus, by definition,

$$\Pi_{X_{(q)}} \triangleq \text{Proj}_{U_{(q)}} \Pi_X, \qquad (3.12)$$

which implies that the probability distribution function of $X_{(q)}$ is related to that of X by

$$\pi_{X_{(q)}}(u_{(q)} = \bigvee_{u_{(q')}} \pi_X(u) \qquad (3.13)$$

where $u_{(q)} \triangleq (u_{i_1}, \ldots, u_{i_k})$, $q' \triangleq (j_1, \ldots, j_m)$ is a subsequence of $(1, \ldots, n)$ which is complementary to q (e.g., if $n = 5$ and $q \triangleq (i_1\, i_2) = (2, 4)$, then $q' = (j_1, j_2, j_3) = (1, 3, 5)$, $u_{(q')} \triangleq (u_{j_1}, \ldots, u_{j_m})$ and $\vee_{u_{(q')}}$ denotes the supremum over $(u_{j_1}, \ldots, u_{j_m}) \in U_{j_1} \times \ldots \times U_{j_m}$.

As a simple illustration, assume that $U_1 = U_2 = U_3 = \{a, b\}$ and the tableau of Π_X is given by

Table 4
Tableau of Π_X

Π_X	X_1	X_2	X_3	π
	a	a	a	0.8
	a	a	b	1
	b	a	a	0.6
	b	a	b	0.2
	b	b	b	0.5

Then,

$$\Pi_{(X_1,X_2)} = \mathrm{Proj}_{U_1 \times U_2} \Pi_X = 1/(a,a) + 0.6/(b,a) + 0.5/(b,b) \tag{3.14}$$

which in tabular form reads

Table 5
Tableau of $\Pi_{(X_1,X_2)}$

$\Pi_{(X_1,X_2)}$	X_1	X_2	π
	a	a	1
	b	a	0.6
	b	b	0.5

Then, from Π_X it follows that the possibility that $X_1 = b, X_2 = a$ and $X_3 = b$ is 0.2, while from $\Pi_{(X_1,X_2)}$ it follows that the possibility of $X_1 = b$ and $X_2 = a$ is 0.6.

By analogy with the concept of independence of random variables, the fuzzy variables

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

and

$$X_{(q')} \triangleq (X_{j_1, \ldots, j_m})$$

are *noninteractive* [5] if and only if the possibility distribution associated with $X = (X_1, \ldots, X_n)$ is the cartesian product of the possibility distributions associated with $X_{(q)}$ and $X_{(q')}$, i.e.,

$$\Pi_X = \Pi_{X_{(q)}} \times \Pi_{X_{(q')}} \tag{3.15}$$

or, equivalently,

$$\pi_X(u_1, \ldots, u_n) = \pi_{X_{(q)}}(u_{i_1}, \ldots, u_{i_k}) \wedge \pi_{X_{(q')}}(u_{j_1}, \ldots, u_{j_m}). \tag{3.16}$$

In particular, the variables $X_1, \ldots, X_n$ are noninteractive if and only if

$$\Pi_X = \Pi_{X_1} \times \Pi_{X_2} \times \cdots \times \Pi_{X_n}. \tag{3.17}$$

The intuitive significance of noninteraction may be clarified by a simple example. Suppose that $X \triangleq (X_1, X_2)$, and X_1 and X_2 are noninteractive, i.e.,

$$\pi_X(u_1, u_2) = \pi_{X_1}(u_1) \wedge \pi_{X_2}(u_2). \tag{3.18}$$

Furthermore, suppose that for some particular values of u_1 and u_2, $\pi_{X_1}(u_1) = \alpha_1$, $\pi_{X_2}(u_2) = \alpha_2 < \alpha_1$ and hence $\pi_X(u_1, u_2) = \alpha_2$. Now, if the value of $\pi_{X_1}(u_1)$ is increased to $\alpha_1 + \delta_1$, $\delta_1 > 0$, it is not possible to decrease the value of $\pi_{X_2}(u_2)$ by a positive amount, say δ_2, such that the value of $\pi_X(u_1, u_2)$ remains unchanged. In this sense, an increase in the possibility of u_1 cannot be compensated by a decrease in the possibility of u_2, and vice-versa. Thus, in essence, noninteraction may be viewed as a form of noncompensation in which a variation in one or more components of a possibility distribution cannot be compensated by variations in the complementary components.

In the manipulation of possibility distributions, it is convenient to employ a type of symbolic representation which is commonly used in the case of fuzzy sets. Specifically, assume, for simplicity, that $U_1, \ldots, U_n$ are finite sets, and let $r^i \triangleq (r_1^i, \ldots, r_n^i)$ denote an n-tuple of values drawn from $U_1, \ldots, U_n$, respectively. Furthermore, let π_i denote the possibility of r^i and let the n-tuple $(r_1^i, \ldots, r_n^i)$ be written as the string $r_1^i \cdots r_n^i$.

Using this notation, a possibility distribution Π_X may be expressed in the symbolic form

$$\Pi_X = \sum_{i=1}^{N} \pi_i r_1^i \, r_2^i \cdots r_n^i \tag{3.19}$$

or, in case a separator symbol is needed, as

$$\Pi_X = \sum_{i=1}^{N} \pi_i / r_1^i r_2^i \cdots r_n^i, \tag{3.20}$$

where N is the number of n-tuples in the tableau of Π_X, and the summation should be interpreted as the union of the fuzzy singletons $\pi_i/(r_1^i, \ldots, r_n^i)$. As an illustration, in the notation of (3.19), the possibility distribution defined in Table 4 reads

$$\Pi_X = 0.8aaa + 1aab + 0.6baa + 0.2bab + 0.5bbb. \tag{3.21}$$

The advantage of this notation is that it allows the possibility distributions to be manipulated in much the same manner as linear forms in n variables, with the understanding that, if r and s are two tuples and α and β are their respective possibilities, then

$$\alpha r + \beta r = (\alpha \vee \beta) r \tag{3.22}$$

$$\alpha r \cap \beta r = (\alpha \wedge \beta) r \tag{3.23}$$

and

$$\alpha r \times \beta s = (\alpha \wedge \beta) rs. \tag{3.24}$$

where rs denotes the concatenation of r and s. For example, if

$$\Pi_X = 0.8aa + 0.5ab + 1bb \tag{3.25}$$

and

$$\Pi_Y = 0.9ba + 0.6bb \tag{3.26}$$

then

$$\Pi_X + \Pi_Y = 0.8aa + 0.5ab + 0.9ba + 1bb \tag{3.27}$$

$$\Pi_X \cap \Pi_Y = 0.6bb \tag{3.28}$$

and

$$\Pi_X \times \Pi_Y = 0.8aaba + 0.5abba + 0.9bbba + 0.6aabb + 0.5abbb + 0.6bbbb. \tag{3.29}$$

To obtain the projection of a possibility distribution Π_X on $U_{(q)} \triangleq (U_{i_1},\ldots,U_{i_k})$, it is sufficient to set the values of $X_{j_1},\ldots,X_{j_m}$ in each tuple in Π_X equal to the null string Λ (i.e., multiplicative identity). As an illustration, the projection of the possibility distribution defined by Table 4 on $U_1 \times U_2$ is given by

$$\begin{aligned}\mathrm{Proj}_{U_1 \times U_2}\Pi_X &= 0.8aa + 1aa + 0.6ba + 0.2ba + 0.5bb \\ &= 1aa + 0.6ba + 0.5bb\end{aligned} \tag{3.30}$$

which agrees with Table 5.

3.2. *Conditioned possibility distributions*

In the theory of possibilities, the concept of a conditioned possibility distribution plays a role that is analogous—though not completely—to that of a conditional possibility distribution in the theory of probabilities.

More concretely, let a variable $X = (X_1,\ldots,X_n)$ be associated with a possibility distribution Π_X, with Π_X characterized by a possibility distribution function $\pi_X(u_1,\ldots,u_n)$ which assigns to each n-tuple $(u_1,\ldots,u_n)$ in $U_1 \times \cdots \times U_n$ its possibility $\pi_X(u_1,\ldots,u_n)$.

Let $q = (i_1,\ldots,i_k)$ and $s = (j_1,\ldots,j_m)$ be subsequences of the index sequence $(1,\ldots,n)$, and let $(a_{j_1},\ldots,a_{j_m})$ be an n-tuple of values assigned to $X_{(q')} = (X_{j_1},\ldots,X_{j_m})$. By definition, the *conditioned possibility distribution* of

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

given

$$X_{(q')} = (a_{j_1},\ldots,a_{j_m})$$

is a possibility distribution expressed as

$$\Pi_{X_{(q)}}[X_{j_1} = a_{j_1};\ldots;X_{j_m} = a_{j_m}]$$

whose possibility distribution function is given by[10]

$$\pi_{X_{(q)}}(u_{i_1},\ldots,u_{i_k} \mid X_{j_1}=a_{j_1};\ldots;X_{j_m}=a_{j_m}) \tag{3.31}$$

$$\triangleq \pi_X(u_1,\ldots,u_n)\Big|_{u_{j_1}=a_{j_1},\ldots,u_{j_m}=a_{j_m}}.$$

As a simple example, in the case of (3.21), we have

$$\Pi_{(X_2,X_3)}[X_1=a]=0.8aa+1ab \tag{3.32}$$

as the expression for the conditioned possibility distribution of (X_2,X_3) given $X_1=a$.

An equivalent expression for the conditioned possibility distribution which makes clearer the connection between

$$\Pi_{X_{(q)}}[X_{j_1}=a_{j_1};\ldots;X_{j_m}=a_{j_m}]$$

and Π_X may be derived as follows.

Let

$$\Pi_X[X_{j_1}=a_{j_1};\ldots;X_{j_m}=a_{j_m}]$$

denote a possibility distribution which consists of those terms in (3.19) in which the j_1th element is a_{j_1}, the j_2th element is $a_{j_2},\ldots$, and the j_mth element is a_{j_m}. For example, in the case of (3.21)

$$\Pi_X(X_1=a]=0.8aaa+1aab. \tag{3.33}$$

Expressed in the above notation, the conditioned possibility distribution of $X_{(q)}=(X_{i_1},\ldots,X_{i_k})$ given $X_{j_1}=a_{j_1},\ldots,X_{j_m}=a_{j_m}$ may be written as

$$\Pi_{X_{(q)}}[X_{j_1}=a_{j_1};\ldots;X_{j_m}=a_{j_m}]$$

$$=\text{Proj}_{U_{(q)}}\Pi_X[X_{j_1}=a_{j_1};\ldots;X_{j_m}=a_{j_m}] \tag{3.34}$$

which places in evidence that $\Pi_{X_{(q)}}$ (conditioned on $X_{(s)}=a_{(s)}$) is a marginal possibility distribution induced by Π_X (conditioned on $X_{(s)}=a_{(s)}$). Thus, by employing (3.33) and (3.34), we obtain

$$\Pi_{(X_2,X_3)}[X_1=a]=0.8aa+1ab \tag{3.35}$$

which agrees with (3.32).

In the foregoing discussion, we have assumed that the possibility distribution of $X=(X_1,\ldots,X_n)$ is conditioned on the values assigned to a specified subset, $X_{(s)}$, of the constituent variables of X. In a more general setting, what might be specified is a

[10] In some applications, it may be appropriate to normalize the expression for the conditioned possibility distribution function by dividing the right-hand member of (3.31) by its supremum over $U_{i_1}\times\cdots\times U_{i_k}$.

possibility distribution associated with $X_{(s)}$ rather than the values of $X_{j_1}, \ldots, X_{j_m}$. In such cases, we shall say that Π_X is *particularized*[11] by specifying that $\Pi_{X_{(s)}} = G$, where G is a given m-ary possibility distribution. It should be noted that in the present context $\Pi_{X_{(s)}}$ is a given possibility distribution rather than a marginal distribution that is induced by Π_X.

To analyze this case, it is convenient to assume—in order to simplify the notation—that $X_{j_1} = X_1, X_{j_2} = X_2, \ldots, X_{j_m} = X_m$, $m < n$. Let $\bar{G}$ denote the *cylindrical extension* of G, that is, the possibility distribution defined by

$$\bar{G} \triangleq G \times U_{m+1} \times \cdots \times U_n \tag{3.36}$$

which implies that

$$\mu_{\bar{G}}(u_1, \ldots, u_n) \triangleq \mu_G(u_1, \ldots, u_m), \qquad u_j \in U_j, \qquad j = 1, \ldots, n, \tag{3.37}$$

where μ_G is the membership function of the fuzzy relation G.

The assumption that we are given Π_X and G is equivalent to assuming that we are given the intersection $\Pi_X \cap \bar{G}$. From this intersection, then, we can deduce the particularized possibility distribution $\Pi_{X_{(q)}}[\Pi_{X_{(s)}} = G]$ by projection on $U_{(q)}$. Thus

$$\Pi_{X_{(q)}}[\Pi_{X_{(s)}} = G] = \mathrm{Proj}_{U_{(q)}} \Pi_X \cap \bar{G}. \tag{3.38}$$

Equivalently, the left-hand member of (3.38) may be regarded as the composition of Π_X and G [5].

As a simple illustration, consider the possibility distribution defined by (3.21) and assume that

$$G = 0.4aa + 0.8ba + 1bb. \tag{3.39}$$

Then

$$\bar{G} = 0.4aaa + 0.4aab + 0.8baa + 0.8bab + 1bba + 1bbb \tag{3.40}$$

$$\Pi_X \cap \bar{G} = 0.4aaa + 0.4aab + 0.6baa + 0.2bab + 0.5bbb \tag{3.41}$$

and

$$\Pi_{X_3}[\Pi_{(X_1, X_2)} = G] = 0.6a + 0.5b. \tag{3.42}$$

As an elementary application of (3.38), consider the proposition $p \triangleq$ John is big, where big is a relation whose tableau is of the form shown in Table 6 (with height and weight expressed in metric units).

[11] In the case of nonfuzzy relations, particularization is closely related to what is commonly referred to as *restriction*. We are not employing this more conventional term here because of our use of the term "fuzzy restriction" to denote an elastic constraint on the values that may be assigned to a variable.

Table 6

Tableau of big

Big	Height	Weight	μ
	170	70	0.7
	170	80	0.8
	180	80	0.9
	.	.	.
	190	90	1

Now, suppose that in addition to knowing that John is big, we also know that $q \triangleq$ John is tall, where the tableau of tall is given (in partially tabulated form) by Table 7.

Table 7

Tableau of tall

Tall	Height	μ
	170	0.8
	180	0.9
	190	1

The question is: What is the weight of John? By making use of (3.38), the possibility distribution of the weight of John may be expressed as

$$\Pi_{\text{weight}} = \text{Proj}_{\text{weight}}\, \Pi_{(\text{height, weight})}[\Pi_{\text{height}} = \text{tall}] \tag{3.39}$$

$$= 0.7/70 + 0.9/80 + 1/90.$$

An acceptable linguistic approximation [5], [13] to the right-hand side of (3.39) might be "somewhat heavy," where "somewhat" is a modifier which has a specified effect on the fuzzy set labeled "heavy." Correspondingly, an approximate answer to the question would be "John is somewhat heavy."

4. Possibility distributions of composite and qualified propositions

As was stated in the Introduction, the concept of a possibility distribution provides a natural way for defining the meaning as well as the information content of a proposition in a natural language. Thus, if p is a proposition in a natural language NL and M is its meaning, then M may be viewed as a procedure which acts on a set of relations in a universe of discourse associated with NL and yields the possibility distribution of a set of variables or relations which are explicit or implicit in p.

In constructing the meaning of a given proposition, it is convenient to have a collection of what might be called *conditional translation rules* [30] which relate the meaning of a proposition to the meaning of its modifications or combinations with other propositions. In what follows, we shall discuss briefly some of the basic rules of this type and, in particular, will formulate a rule governing the modification of possibility distributions by the *possibility qualification* of a proposition.

4.1. Rules of type I

Let p be a proposition of the form X is F, and let m be a modifier such as very, quite, rather, etc. The so-called *modifier rule* [6] which defines the modification in the possibility distribution induced by p may be stated as follows.

If

$$X \text{ is } F \rightarrow \Pi_{A(X)} = F \tag{4.1}$$

then

$$X \text{ is } mF \rightarrow \Pi_{A(X)} = F^{+} \tag{4.2}$$

where $A(X)$ is an implied attribute of X and F^{+} is a modification of F defined by m.[12] For example, if $m \triangleq$ very, then $F^{+} = F^2$; if $m \triangleq$ more or less then $F^{+} = \sqrt{F}$; and if $m \triangleq$ not then $F^{+} = F' \triangleq$ complement of F. As an illustration:

If

$$\text{John is young} \rightarrow \Pi_{\text{Age(John)}} = \text{young} \tag{4.3}$$

then

$$\text{John is very young} \rightarrow \Pi_{\text{Age(John)}} = \text{young}^2.$$

In particular, if

$$\text{young} = 1 - S(20, 30, 40) \tag{4.4}$$

then

$$\text{young}^2 = (1 - S(20, 30, 40))^2,$$

where the S-function (with its argument suppressed) is defined by (2.6).

4.2. Rules of type II

If p and q are propositions, then $r \triangleq p * q$ denotes a proposition which is a *composition* of p and q. The three most commonly used modes of composition are (i) conjunctive, involving the connective "and"; (ii) disjunctive, involving the connective "or"; and (iii) conditional, involving the connective "if...then." The conditional translation rules relating to these modes of composition are stated below.

[12] A more detailed discussion of the effect of modifiers (or hedges) may be found in [15, 16, 17, 8, 6, 13 and 18].

Conjunctive (noninteractive): If

$$X \text{ is } F \to \Pi_{A(X)} = F \tag{4.5}$$

and

$$Y \text{ is } G \to \Pi_{B(Y)} = G \tag{4.6}$$

then

$$X \text{ is } F \text{ and } Y \text{ is } G \to \Pi_{(A(X),B(Y))} = F \times G \tag{4.7}$$

where $A(X)$ and $B(Y)$ are the implied attributes of X and Y, respectively, $\Pi_{(A(X),B(Y))}$ is the possibility distribution of the variables $A(X)$ and $B(Y)$, and $F \times G$ is the cartesian product of F and G. It should be noted that $F \times G$ may be expressed equivalently as

$$F \times G = \bar{F} \cap \bar{G} \tag{4.8}$$

where $\bar{F}$ and $\bar{G}$ are the cylindrical extensions of F and G, respectively.

Disjunctive (noninteractive): If (4.5) and (4.6) hold, then

$$X \text{ is } F \text{ or } Y \text{ is } G \to \Pi_{(A(X),B(Y))} = \bar{F} + \bar{G} \tag{4.9}$$

where the symbols have the same meaning as in (4.5) and (4.6), and $+$ denotes the union.

Conditional (noninteractive): If (4.5) and (4.6) hold, then

$$\text{If } X \text{ is } F \text{ then } Y \text{ is } G \to \Pi_{(A(X),B(Y))} = \bar{F}' \oplus \bar{G} \tag{4.10}$$

where F' is the complement of F and $\oplus$ is the bounded sum defined by

$$\mu_{F' \oplus G} = 1 \wedge (1 - \mu_F + \mu_G), \tag{4.11}$$

in which $+$ and $-$ denote the arithmetic addition and subtraction, and μ_F and μ_G are the membership functions of F and G, respectively. Illustrations of these rules—expressed in terms of fuzzy restrictions rather than possibility distributions—may be found in [6 and 14].

4.3. *Truth qualification, probability qualification and possibility qualification*

In natural languages, an important mechanism for the modification of the meaning of a proposition is provided by the adjuction of three types of qualifiers: (i) is τ, where τ is a linguistic truth-value, e.g., true, very true, more or less true, false, etc.; (ii) is λ, where λ is a linguistic probability-value (or likelihood), e.g., likely, very likely, very unlikely, etc.; and (iii) is π, where π is a linguistic possibility-value, e.g., possible, quite possible, slightly possible, impossible, etc. These modes of qualification will be referred to, respectively, as *truth qualification, probability qualification* and *possibility qualification.* The rules governing these qualifications may be stated as follows.

Truth qualification: If

$$X \text{ is } F \to \Pi_{A(X)} = F \tag{4.12}$$

then

$$X \text{ is } F \text{ is } \tau \to \Pi_{A(X)} = F^{+},$$

where

$$\mu_{F^+}(u) = \mu_\tau(\mu_F(u)), \qquad u \in U; \tag{4.13}$$

μ_τ and μ_F are the membership functions of τ and F, respectively, and U is the universe of discourse associated with $A(X)$. As an illustration, if young is defined by (4.4); τ = very true is defined by

$$\text{very true} = S^2(0.6, 0.8, 1) \tag{4.14}$$

and

$$\text{John is young} \to \Pi_{\text{Age(John)}} = \text{young}$$

then

$$\text{John is young is very true} \to \Pi_{\text{Age(John)}} = \text{young}^{+}$$

where

$$\mu_{\text{young}^+}(u) = S^2(1 - S(u;\ 20,\ 30,\ 40);\ 0.6,\ 0.8,\ 1), \qquad u \in U.$$

It should be noted that for the *unitary* truth-value, *u*-true, defined by

$$\mu_{u\text{-true}}(v) = v, \qquad v \in [0, 1] \tag{4.15}$$

(4.13) reduces to

$$\mu_{F^+}(u) = \mu_F(u), \qquad u \in U$$

and hence

$$X \text{ is } F \text{ is } u\text{-true} \to \Pi_X = F. \tag{4.16}$$

Thus, the possibility distribution induced by any proposition is invariant under unitary truth qualification.

Probability qualification: If

$$X \text{ is } F \to \Pi_{A(X)} = F$$

then

$$X \text{ is } F \text{ is } \lambda \to \Pi_{\int_U p(u)\mu_F(u)\,du} = \lambda \tag{4.17}$$

where $p(u)\mathrm{d}u$ is the probability that the value of $A(X)$ falls in the interval $(u, u+\mathrm{d}u)$; the integral

$$\int_U p(u)\mu_F(u)\mathrm{d}u$$

is the probability of the fuzzy event F [19]; and λ is a linguistic probability-value. Thus, (4.17) defines a possibility distribution of probability distributions, with the possibility of a probability density $p(\cdot)$ given implicitly by

$$\pi[\int_U p(u)\mu_F(u)\mathrm{d}u] = \mu_\lambda[\int_U p(u)\mu_F(u)\mathrm{d}u]. \tag{4.18}$$

As an illustration, consider the proposition $p \triangleq$ John is young is very likely, in which young is defined by (4.4) and

$$\mu_{\text{very likely}} = S^2(0.6, 0.8, 1). \tag{4.19}$$

Then

$$\pi[\int_U p(u)\mu_F(u)\mathrm{d}u] = S^2[\int_0^1 p(u)(1 - S(u;20,30,40))\mathrm{d}u;0.6,0.8,1].$$

It should be noted that the probability qualification rule is a consequence of the assumption that the propositions "X is F is λ" and "Prob$\{X$ is $F\} = \lambda$" are *semantically equivalent* (i.e., induce identical possibility distributions), which is expressed in symbols as

$$X \text{ is } F \text{ is } \lambda \leftrightarrow \text{Prob}\{X \text{ is } F\} = \lambda. \tag{4.20}$$

Thus, since the probability of the fuzzy event F is given by

$$\text{Prob}\{X \text{ is } F\} = \int_U p(u)\mu_F(u)\mathrm{d}u,$$

it follows from (4.20) that we can assert the semantic equivalence

$$X \text{ is } F \text{ is } \lambda \leftrightarrow \int_U p(u)\mu_F(u)\mathrm{d}u \text{ is } \lambda,$$

which by (2.11) leads to the right-hand member of (4.17).

Possibility qualification: Our concern here is with the following question: Given that "X is F" translates into the possibility assignment equation $\Pi_{A(X)} = F$, what is the translation of "X is F is π," where π is a linguistic possibility-value such as quite possible, very possible, more or less possible, etc.? Since our intuition regarding the behavior of possibility distributions is not well-developed at this juncture, the answer suggested in the following should be viewed as tentative in nature.

For simplicity, we shall interpret the qualifier "possible" as "1-possible," that is, as the assignment of the possibility-value 1 to the proposition which it qualifies. With this understanding, the translation of "X is F is possible" will be assumed to be given by

$$X \text{ is } F \text{ is possible} \to \Pi_{A(X)} = F^+, \tag{4.21}$$

in which

$$F^{+} = F \oplus \Pi \tag{4.22}$$

where Π is a fuzzy set of Type 2[13] defined by

$$\mu_{\Pi}(u) = [0, 1], \qquad u \in U, \tag{4.23}$$

and $\oplus$ is the bounded sum defined by (4.11). Equivalently,

$$\mu_{F^+}(u) = [\mu_F(u), 1], \qquad u \in U, \tag{4.24}$$

which defines μ_{F^+} as an interval-valued membership function.

In effect, the rule in question signifies that possibility qualification has the effect of weakening the proposition which it qualifies through the addition to F of a possibility distribution Π which represents total indeterminacy[14] in the sense that the degree of possibility which it associates with each point in U may be any number in the interval $[0, 1]$. An illustration of the application of this rule to the proposition $p \triangleq X$ is small is shown in Fig. 1.

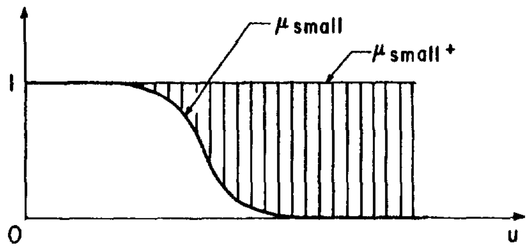

Fig. 1. The possibility distribution of "X is small is possible".

As an extension of the above rule, we have: If

$$X \text{ is } F \rightarrow \Pi_{A(X)} = F \tag{4.25}$$

then, for $0 \leq \alpha \leq 1$,

$$X \text{ is } F \text{ is } \alpha\text{-possible} \rightarrow \Pi_{A(X)} = F^{+} \tag{4.26}$$

where F^+ is a fuzzy set of Type 2 whose interval-valued membership function is given by

$$\mu_{F^+}(u) = [\alpha \wedge \mu_F(u), \alpha \oplus (1 - \mu_F(u))], \qquad u \in U. \tag{4.27}$$

[13] The membership function of a fuzzy set of Type 2 takes values in the set of fuzzy subsets of the unit interval [5, 6].

[14] Π may be interpreted as the possibilistic counterpart of white noise.

As an illustration, the result of the application of this rule to the proposition $p \triangleq X$ is small is shown in Fig. 2. Note that the rule expressed by (4.24) may be regarded as a special case of (4.27) corresponding to $\alpha = 1$.

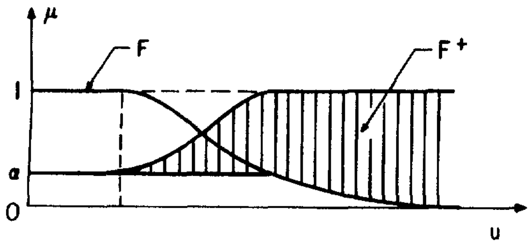

Fig. 2. The possibility distribution of "X is small is α-possible".

A further extension of the rule expressed by (4.25) to linguistic possibility-values may be obtained by an application of the extension principle, leading to the *linguistic possibility qualification rule*:

If

$$X \text{ is } F \rightarrow \Pi_{A(X)} = F$$

then

$$X \text{ is } F \text{ is } \pi \rightarrow \Pi_{A(X)} = F^+ \tag{4.28}$$

where F^+ is a fuzzy set of Type 2 whose membership function is given by

$$\mu_F(u) = \{\geqq \circ (\pi \wedge \mu_F(u)) \cap (\leqq \circ (\pi \oplus (1 - \mu_F(u))))\}, \tag{4.29}$$

where π is the linguistic possibility (e.g., quite possible, almost impossible, etc.) and $\circ$ denotes the composition of fuzzy relations. This rule should be regarded as speculative in nature since the implications of a linguistic possibility qualification are not as yet well understood

An alternative approach to the translation of "X is F is π" is to interpret this proposition as

$$X \text{ is } F \text{ is } \pi \leftrightarrow \text{Poss}\{X \text{ is } F\} = \pi, \tag{4.30}$$

which is in the spirit of (4.20), and then formulate a rule of the form (4.28) in which $\Pi_{A(X)}$ is the largest (i.e., least restrictive) possibility distribution satisfying the constraint $\text{Poss}\{X \text{ is } F\} = \pi$. A complicating factor in this case is that the proposition "X is F is π" may be associated with other implicit propositions such as "X is not F is [0, 1]-possible," or "X is not F is not impossible," which affect the translation of "X is F is π." In this connection, it would be useful to deduce the translation rules (4.21), (4.26) and (4.29) (or their variants) from a conjunction of "X is F is π" with other implicit propositions involving the negation of "X is F."

An interesting aspect of possibility qualification relates to the invariance of implication under this mode of qualification. Thus, from the definition of implication [6], it follows at once that

$$X \text{ is } F \Rightarrow X \text{ is } G \quad \text{if } F \subset G.$$

Now, it can readily be shown that

$$F \subset G \Rightarrow F^{+} \subset G^{+} \tag{4.31}$$

where $\subset$ in the right-hand member of (4.31) should be interpreted as the relation of containment for fuzzy sets of Type 2. In consequence of (4.31), then, we can assert that

$$X \text{ is } F \text{ is possible} \Rightarrow X \text{ is } G \text{ is possible} \quad \text{if } F \subset G. \tag{4.32}$$

5. Concluding remarks

The exposition of the theory of possibility in the present paper touches upon only a few of the many facets of this—as yet largely unexplored—theory. Clearly, the intuitive concepts of possibility and probability play a central role in human decision-making and underlie much of the human ability to reason in approximate terms. Consequently, it will be essential to develop a better understanding of the interplay between possibility and probability—especially in relation to the roles which these concepts play in natural languages—in order to enhance our ability to develop machines which can simulate the remarkable human ability to attain imprecisely defined goals in a fuzzy environment.

Acknowledgment

The idea of employing the theory of fuzzy sets as a basis for the theory of possibility was inspired by the paper of B.R. Gaines and L. Kohout on possible automata [1]. In addition, our work was stimulated by discussions with H.J. Zimmermann regarding the interpretation of the operations of conjunction and disjunction; the results of a psychological study conducted by Eleanor Rosch and Louis Gomez; and discussions with Barbara Cerny.

References

[1] B.R. Gaines and L.J. Kohout, Possible automata, Proc. Int. Symp. on Multiple-Valued Logic, University of Indiana, Bloomington, IN (1975) 183–196.
[2] G.E. Hughes and M.J. Cresswell, An Introduction to Modal Logic (Methuen, London, 1968).
[3] A. Kaufmann, Valuation and probabilization, in: A. Kaufmann and E. Sanchez, eds., Theory of Fuzzy Subsets, Vol. 5 (Masson and Co., Paris, 1977).
[4] L.A. Zadeh, Calculus of fuzzy restrictions, in: L.A. Zadeh, K.S. Fu, K. Tanaka and M. Shimura, eds., Fuzzy Sets and Their Applications to Cognitive and Decision Processes (Academic Press, New York, 1975) 1–39.

[5] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning, Part I, Information Sci. 8 (1975) 199–249; Part II, Information Sci. 8 (1975) 301–357; Part III, Information Sci. 9 (1975) 43–80.
[6] R.E. Bellman and L.A. Zadeh, Local and fuzzy logics, ERL Memo. M-584, University of California, Berkeley, CA (1976); Modern Uses of Multiple-Valued Logics, D. Epstein, ed. (D. Reidel, Dordrecht, 1977).
[7] B. DeFinetti, Probability Theory (Wiley, New York, 1974).
[8] T. Fine, Theories of Probability (Academic Press, New York, 1973).
[9] A. Dempster, Upper and lower probabilities induced by multi-valued mapping, Ann. Math. Statist. 38 (1967) 325–339.
[10] G. Shafer, A Mathematical Theory of Evidence (Princeton University Press, Princeton, NJ, 1976).
[11] E.H. Shortliffe, A model of inexact reasoning in medicine, Math. Biosciences 23 (1975) 351–379.
[12] R.O. Duda, P.F. Hart and N.J. Nilsson, Subjective Bayesian methods for rule-based inference systems, Stanford Research Institute Tech. Note 124, Stanford, CA (1976).
[13] F. Wenstop, Deductive verbal models of organization, Int. J. Man-Machine Studies 8 (1976) 293–311.
[14] L.A. Zadeh, Theory of fuzzy sets, Memo. No. UCB/ERL M77/1, University of California, Berkeley, CA (1977).
[15] L.A. Zadeh, A fuzzy-set-theoretic interpretation of linguistic hedges, J. Cybernet. 2 (1972) 4–34.
[16] G. Lakoff, Hedges: a study in meaning criteria and the logic of fuzzy concepts, J. Phil. Logic 2 (1973) 458–508; also paper presented at 8th Regional Meeting of Chicago Linguistic Soc. (1972) and in: D. Hockney, W. Harper and B. Freed, eds., Contemporary Research in Philosophical Logic and Linguistic Semantics (D. Reidel, Dordrecht, 1975) 221–271.
[17] H.M. Hersch and A. Caramazza, A fuzzy set approach to modifiers and vagueness in natural languages, Dept. of Psych., The Johns Hopkins University, Baltimore, MD (1975).
[18] P.J. MacVicar-Whelan, Fuzzy sets, the concept of height and the hedge very, Tech. Memo 1, Physics Dept., Grand Valley State College, Allendale, MI (1974).
[19] L.A. Zadeh, Probability measures of fuzzy events, J. Math. Anal. Appl. 23 (1968) 421–427.
[20] M. Sugeno, Theory of fuzzy integrals and its applications, Ph.D. Thesis, Tokyo Institute of Technology, Tokyo (1974).
[21] T. Terano and M. Sugeno, Conditional fuzzy measures and their applications, in: L.A. Zadeh, K.S. Fu, K. Tanaka and M. Shimura, eds., Fuzzy Sets and Their Applications to Cognitive and Decision Processes (Academic Press, New York, 1975) 151–170.
[22] S. Nahmias, Fuzzy variables, Tech. Rep. 33, Dept. of Ind. Eng., Systems Manag. Eng. and Operations Research, University of Pittsburgh, PA (1976); Presented at the ORSA/TIMS meeting, Miami, FL (November 1976).
[23] L.A. Zadeh, Similarity relations and fuzzy orderings, Information Sci. 3 (1971) 177–200.
[24] W. Rödder, On 'and' and 'or' connectives in fuzzy set theory, Inst. for Oper. Res., Technical University of Aachen, Aachen (1975).
[25] L.A. Zadeh, K.S. Fu, K. Tanaka and M. Shimura, eds., Fuzzy Sets and Their Applications to Cognitive and Decision Processes (Academic Press, New York, 1975).
[26] E. Mamdani, Application of fuzzy logic to approximate reasoning using linguistic synthesis. Proc. 6th Int. Symp. on Multiple-Valued Logic, Utah State University, Logan, UT (1976) 196–202.
[27] V.V. Nalimov, Probabilistic model of language, Moscow State University, Moscow (1974).
[28] J.A. Goguen, Concept representation in natural and artificial languages: axioms, extensions and applications for fuzzy sets, Int. J. Man-Machine Studies 6 (1974) 513–561.
[29] C.V. Negoita and D.A. Ralescu, Applications of Fuzzy Sets to Systems Analysis (Birkhauser, Stuttgart, 1975).
[30] L.A. Zadeh, PRUF–A meaning representation language for natural languages, Memo No. UCB/ERL M77/61, University of California, Berkeley, CA (1977).