A THEORY OF COMMONSENSE KNOWLEDGE

By

L. A. Zadeh

Cognitive Science Program

Institute of Cognitive Studies

University of California at Berkeley

October 1983

An abridged version of this paper entitled "Commonsense Knowledge Representation Based on Fuzzy Logic" is to appear in *Computer Magazine*, IEEE, New York.

The production of the Berkeley Cognitive Science Report Series is supported by a grant in cognitive science to the University of California at Berkeley from the Alfred P. Sloan Foundation.

A Theory of Commonsense Knowledge

*L. A. Zadeh**

The theory outlined in this paper is based on the idea that what is commonly called *commonsense knowledge* may be viewed as a collection of *dispositions*, that is, propositions with implied fuzzy quantifers. Typical examples of dispositions are: *Icy roads are slippery, Tall men are not very agile, Overeating causes obesity, Bob loves women, What is rare is expensive,* etc. It is understood that, upon restoration of fuzzy quantifiers, a disposition is converted into a proposition with explicit fuzzy quantifiers, e.g., *Tall men are not very agile* $\rightarrow$ *Most tall men are not very agile.*

Since traditional logical systems provide no methods for representing the meaning of propositions containing fuzzy quantifiers, such systems are unsuitable for dealing with commonsense knowledge. It is suggested in this paper that an appropriate computational framework for dealing with commonsense knowledge is provided by *fuzzy logic*, which, as its name implies, is the logic underlying fuzzy (or approximate) reasoning. Such a framework, with an emphasis on the representation of dispositions, is outlined and illustrated with examples.

1. Introduction

It is widely agreed at this juncture that one of the important problem-areas in AI relates to the representation of commonsense knowledge. In general, such knowledge may be regarded as a collection of propositions exemplified by: *Snow is white, Icy roads are slippery, Most Frenchmen are not very tall, Virginia is very intelligent, If a car which is offered for sale is cheap and old then it probably is not in good shape, Heavy smoking causes lung cancer,* etc. Representation of propositions of this type plays a particularly important role in the design of expert systems [3].

The conventional knowledge representation techniques based on the use of predicate calculus and related methods are not well-suited for the representation of commonsense knowledge because the predicates in propositions which represent commonsense knowledge do not, in general, have crisp denotations. For example, the proposition *Most Frenchmen are not very tall* cannot be represented as a well-formed formula in predicate calculus because the sets which constitute the denotations of the predicate *tall* and the quantifier *most* in their respective universes of discourse are fuzzy rather than crisp.

To Professor Hans J. Zimmermann.

* Computer Science Division, Department of EECS and ERL, University of California, Berkeley, CA 94720.
Research supported in part by DARPA Contract N00039-C-0235. An abridged version of this paper entitled "Commonsense Knowledge Representation Based on Fuzzy Logic" is to appear in *Computer* Magazine, IEEE, New York.

More generally, the inapplicability of predicate calculus and related logical systems to the representation of commonsense knowledge reflects the fact that such systems make no provision for dealing with uncertainty. Thus, in predicate logic, for example, a proposition is either true or false and no gradations of truth or membership are allowed. By contrast, in the case of commonsense knowledge, a typical proposition contains a multiplicity of sources of uncertainty. For example, in the case of the proposition *If a car which is offered for sale is cheap and old then it probably is not in good shape,* there are five sources of uncertainty: (i) the temporal uncertainty associated with the fuzzy predicate *old*; (ii) the uncertainty associated with the fuzzy predicate *cheap*; (iii) the uncertainty associated with the fuzzy predicate *not in good shape*; (iv) the probabilistic uncertainty associated with the event *The car is not in good shape*; and (v) the uncertainty associated with the fuzzy characterization of the probability of the event in question as *probable*.

The approach to the representation of commonsense knowledge which is described in this paper is based on the idea that propositions which characterize commonsense knowledge are, for the most part, *dispositions*, that is, propositions with implied fuzzy quantifiers. In this sense, the proposition *Tall men are not very agile* is a disposition which upon restoration is converted into the proposition *Most tall men are not very agile*. In this proposition, *most* is an explicit fuzzy quantifier which provides an approximate characterization of the proportion of *men who are not very agile* among *men who are tall* [63].

To deal with dispositions in a systematic fashion, we shall employ *fuzzy logic* -- which is the logic underlying *approximate* or *fuzzy reasoning* [5, 19, 26, 55, 58]. Basically, fuzzy logic has two principal components. The first component is, in effect, a translation system for representing the meaning of propositions and other types of semantic entities. We shall employ the suggestive term *test-score semantics* to refer to this translation system because it involves an aggregation of the test scores of elastic constraints which are induced by the semantic entity whose meaning is represented.

The second component is an inferential system for arriving at an answer to a question which relates to the information which is resident in a knowledge base. In the present paper, the focus of our attention will be the problem of meaning representation in the context of commonsense knowledge, and our discussion of the inferential component will be limited in scope.*

---

* A more detailed discussion of the inferential component of fuzzy logic may be found in [5, 58, 59]. Recent literature, [26], contains a number of papers dealing with fuzzy logic and its applications. Descriptions of implemented fuzzy-logic-based inferential systems may be found in [2], [33] and [36].

2. Meaning Representation in Test-Score Semantics

Test score semantics is concerned with representation of the meaning of various types of semantic entities, e.g., propositions, predicates, commands, questions, modifiers, etc. However, knowledge, whether commonsense or not, may be viewed as a collection of propositions. For this reason, we shall restrict our discussion of test-score semantics to the representation of the meaning of propositions.**

In test-score semantics, as in PRUF [57], a proposition is regarded as a collection of elastic, or, equivalently, fuzzy constraints. For example, the proposition *Pat is tall* represents an elastic constraint on the height of Pat. Similarly, the proposition *Charlotte is blonde* represents an elastic constraint on the color of Charlotte's hair. And, the proposition *Most tall men are not very agile* represents an elastic constraint on the proportion of men who are not very agile among tall men.

In more concrete terms, representing the meaning of a proposition, p, through the use of test-score semantics involves the following steps.

1. Identification of the variables $X_1, \ldots, X_n$ whose values are constrained by the proposition. Usually, these variables are implicit rather than explicit in p.
2. Identification of the constraints $C_1, \ldots, C_m$ which are induced by p.
3. Characterization of each constraint, C_i , by describing a testing procedure which associates with C_i a test score τ_i representing the degree to which C_i is satisfied. Usually τ_i is expressed as a number in the interval [0,1]. More generally, however, a test score may be a probability/possibility distribution over the unit interval.
4. Aggregation of the partial test scores $\tau_1, \ldots, \tau_m$ into a smaller number of test scores $\bar{\tau}_1, \ldots, \bar{\tau}_k$, which are represented as an *overall vector test score* $\tau = (\bar{\tau}_1, \ldots, \bar{\tau}_k)$. In most cases $k = 1$, so that the overall test scores is a scalar. We shall assume that this is the case unless an explicit statement to the contrary is made.

It is important to note that, in test-score semantics, the meaning of p is represented not by the overall test score τ but by the procedure which leads to it. Viewed in this perspective, test-score semantics may be regarded as a generalization of truth-conditional, possible-world and model-theoretic semantics [8, 9, 28, 32]. However, by providing a computational framework for dealing with uncertainty -- which the conventional semantical systems disregard -- test-score semantics achieves a much higher level of expressive power and thus provides a basis for representing the meaning of a much wider variety of propositions in a natural language.

In test-score semantics, the testing of the constraints induced by p is performed on a collection of fuzzy relations which constitute an *explanatory database*, or *ED* for short. A basic assumption which

** A more detailed exposition of test-score semantics may be found in [61].

is made about the explanatory database is that it is comprised of relations whose meaning is known to the addressee of the meaning-representation process. In an indirect way, then, the testing and aggregation procedures in test-score semantics may be viewed as a description of a process by which the meaning of *p* is composed from the meanings of the constituent relations in the explanatory database. It is this explanatory role of the relations in *ED* that motivates its description as an *explanatory database*.

As will be seen in the sequel, in describing the testing procedures we need not concern ourselves with the actual entries in the constituent relations. Thus, in general, the description of a test involves only the frames* of the constituent relations, that is, their names, their variables (or attributes) and the domain of each variable. When this is the case, the explanatory database will be referred to as the *explanatory database frame*, or *EDF* for short.

As a simple illustration, consider the proposition

$$p = Debbie\ is\ a\ few\ years\ older\ than\ Dana\ . \tag{2.1}$$

In this case, a suitable explanatory database frame may be represented as

$$EDF \triangleq POPULATION\ [Name\,;\ Age\,] + FEW\ [Number\,;\mu],$$

which signifies that the explanatory database frame consists of two relations: (a) a nonfuzzy relation *POPULATION [Name; Age]*, which lists names of individuals and their age; and (b) a fuzzy relation *FEW [Number;* μ *]*, which associates with each value of Number the degree, μ, to which *Number* is compatible with the intended meaning of *few*. In general, the domain of each variable in the *EDF* is implicitly determined by *p*, and is not spelled-out explicitly unless it is necessary to do so to define the testing procedure.

As another example, consider the disposition**

$$p \triangleq Snow\ is\ white,$$

which is frequently used in the literature to explain the basic ideas underlying truth-conditional semantics.***

To construct an *EDF* for this disposition, we first note that what is generally meant by *Snow is white* is *Usually snow is white*, in which *usually* may be interpreted as a fuzzy quantifier.

* In the literature of database management systems, some authors employ the term *schema* in the same sense as we employ *frame*. More commonly, however, the term schema is used in a narrower sense [11], to describe the frame of a relation together with the dependencies between the variables.

** As was stated earlier, a *disposition* is a proposition with implied fuzzy quantifiers. (Note that this definition, too, is a disposition.) For example, the proposition *Small cars are unsafe* is a disposition, since it may be viewed as an abbreviation of the proposition *Most small cars are unsafe*, in which *most* is a fuzzy quantifier. In general, a disposition may be restored in more than one way.

*** In truth-conditional semantics, the truth condition for the proposition *Snow is white* is described as snow is white, which means, in plain terms, that the proposition *Snow is white* is true if and only if snow is white.

Consequently, on the assumption that the proposition

$$p^* = Usually\ snow\ is\ white \tag{2.2}$$

is a restoration of *p*, a natural choice for the *EDF* would be

$$EDF \triangleq WHITE[Sample\,;\,\mu] + USUALLY[Proportion\,;\,\mu]. \tag{2.3}$$

In this *EDF*, the relation *WHITE* is a listing of samples of snow together with the degree, μ, to which each sample is white, while the relation *USUALLY* defines the degree to which a numerical value of *Proportion* is compatible with the intended meaning of *usually*.

In proposition (2.1), the constrained variable is the difference in the ages of Debbie and Dana. Thus,

$$X \triangleq Age(Debbie) - Age(Dana). \tag{2.4}$$

The elastic constraint which is induced by *p* is determined by the fuzzy relation *FEW*. More specifically, let Π_X denote the possibility distribution of *X*, ie., the fuzzy set of possible values of *X*. Then, the constraint on *X* may be expressed as the *possibility assignment equation*

$$\Pi_X = FEW, \tag{2.5}$$

which assigns the fuzzy relation *FEW* to the possibility distribution of *X*. This equation implies that

$$\pi_X(u) \triangleq Poss\{X = u\} = \mu_{FEW}(u)\ \ , \tag{2.6}$$

where *Poss {X = u}* is the possibility that *X* may take *u* as its value; $\mu_{FEW}(u)$ is the grade of membership of *u* in the fuzzy relation *FEW*; and the function π_X (from the domain of *X* to the unit interval) is the *possibility distribution function* associated with *X*.

For the example under consideration, what we have done so far may be summarized by stating that the proposition

$$p \triangleq Debbie\ is\ a\ few\ years\ older\ than\ Dana$$

may be expressed in a canonical form, namely,

$$X\ is\ FEW, \tag{2.7}$$

which places in evidence the implicit variable *X* which is constrained by *p*. The canonical proposition implies and is implied by the possibility assignment equation (2.5), which defines via (2.6) the possibility distribution of *X* and thus characterizes the elastic constraint on *X* which is induced by *p*.

The foregoing analysis may be viewed as an instantiation of the basic idea underlying PRUF, namely, that any proposition, *p*, in a natural language may be expressed in the canonical form

$$cf(p) \triangleq X\ is\ F, \tag{2.8}$$

where $X = (X_1, \ldots, X_n)$ is an n-ary *focal variable* whose constituent variables $X_1, \ldots, X_n$ range over the universes $U_1, \ldots, U_n$, respectively; F is an n-ary fuzzy relation in the product space $U = U_1 \times \cdots \times U_n$ and *cf(p)* is an abbreviation for *canonical form of p*. The canonical form, in turn, may be expressed more concretely as the possibility assignment equation

$$\Pi_X = F, \tag{2.9}$$

which signifies that the possibility distribution of X is given by F. Thus, we may say that *p translates* into the possibility assignment equation (2.9), i.e.

$$p \rightarrow \Pi_X = F \tag{2.10}$$

in the sense that (2.9) exhibits the implicit variable which is constrained by p and defines the elastic constraint which is induced by p.

When we employ test-score semantics, the meaning of a proposition, p, is represented as a test which associates with each *ED* (i.e., an instantiation of *EDF*) an overall test score τ which may be viewed as the *compatibility* of p with *ED*. This compatibility may be interpreted in two equivalent ways: (a) as the truth of p given *ED*; and (b) as the possibility of *ED* given p. The latter interpretation shows that the representation of the meaning of p as a test is equivalent to representing the meaning of p by a possibility assignment equation.*

The connection between the two points of view will become clearer in Section 4, where we shall discuss several examples of propositions representing commonsense knowledge. As a preliminary, we shall present in the following section a brief exposition of some of the basic techniques which will be needed in Section 4 to test the constituent relations in the explanatory database and aggregate the partial test scores.

3. Testing and Translation Rules

A typical relation in *EDF* may be expressed as $R[X_1;...;X_n;\mu]$, where R is the name of the relation; the $X_i, i = 1, \ldots, n$, are the names of the variables (or, equivalently, the attributes of R), with U_i and u_i representing, respectively, the domain of X_i and its generic value; and μ is the grade of membership of a generic n-tuple $u = (u_1, \ldots, u_n)$ in R.

In the case of nonfuzzy relations, a basic operation on R which is a generalization of the familiar operation of looking up the value of a function for a given value of its argument, is the so-called *mapping operation* [11]. The counterpart of this operation for fuzzy relations is the operation of *transduction* [61].

---

* As is pointed out in [61], the translation of p into a possibility assignment equation is an instance of a *focused* translation. By contrast, representation of the meaning of p by a test on *EDF* is an instance of an *unfocused* translation. The two are equivalent in principle but differ in detail.

Transduction may be viewed as a combination of two operations: (a) *particularization**, which constrains the values of a subset of variables of *R*; and *projection*, which reads the induced constraints on another subset of variables of *R*. The subsets in question may be viewed as the *input* and *output* variables, respectively.

To define particularization, it is helpful to view a fuzzy relation as an elastic constraint on n-tuples in $U_1 \times \cdots \times U_n$, with the μ-value for each row in *R* representing the degree (or the test score) with which the constraint is satisfied.

For concreteness, assume that the input variables are X_1, X_2 and X_3, and that the constraints on these variables are expressed as canonical propositions. For example

$$X_1 \ is \ F$$

and

$$(X_2, X_3) \ is \ G,$$

where *F* and *G* are fuzzy subsets of U_1, and $U_2 \times U_3$, respectively. Equivalently, the constraints in question may be expressed as

$$\Pi_{X_1} = F$$

and

$$\Pi_{(X_2, X_3)} = G \ ,$$

where Π_{X_1} and $\Pi_{(X_2,X_3)}$ are the respective possibility distributions of X_1 and X_2, X_3. To place in evidence the input constraints, the particularized relation is written as

$$R^* \triangleq R[X_1 \ is \ F; \ (X_2, X_{3)} \ is \ G] \tag{3.1}$$

or, equivalently, as

$$R^* \triangleq R[\Pi_{X_1} = F; \ \Pi_{(X_2}, X_{3)} = G] \ . \tag{3.2}$$

As a concrete illustration, assume that *R* is a relation whose frame is expressed as

$$RICH \ [Name \,; Age \,; Height \,; Weight \,; Sex \,; \mu] \ , \tag{3.3}$$

in which Age, Height, Weight and Sex are attributes of Name, and μ is the degree to which Name is *rich*. In this case, the input constraints might be:

$$Age \ is \ YOUNG$$

* In the case of nonfuzzy relations, particularization is usually referred to as *selection* [11], or *restriction*.

(Height, Weight) is BIG
Sex is MALE

and, correspondingly, the particularized relation reads

$$R^* \triangleq RICH[Age\ is\ YOUNG;\ (Height,\ Weight)\ is\ BIG;\ Sex\ is\ MALE]\ . \tag{3.4}$$

To concretize the meaning of a particularized relation it is necessary to perform a *row test* on each row of *R*. Specifically, with reference to (3.1), let $r_t = (u_{1t}, \ldots, u_{nt}, \mu_t)$ be the t^{th} row of *R*, where $u_{1t}, \ldots, u_{nt}, \mu_t$ are the values of $X_1, \ldots, X_n, \mu$, respectively. Furthermore, let μ_F and μ_G be the respective membership functions of *F* and *G*. Then, for r_t , the test scores for the constraints on X_1 and (X_2, X_3) may be expressed as

$$\tau_{1t} = \mu_F(u_{1t})$$

$$\tau_{2t} = \mu_G(u_{2t}, u_{3t})\ .$$

To aggregate the test scores with μ_t , we employ the min operator *andsign* * , which leads to the overall test score for r_t :

$$\tau_t = \tau_{1t}\ andsign\ \tau_{2t}\ andsign\ \mu_t\ . \tag{3.5}$$

Then, the particularized relation (3.1) is obtained by replacing each μ_t in r_t , $t = 1,2,...,$ by τ_t . An example illustrating these steps in the computation of a particularized relation may be found in [61].

As was stated earlier, when a fuzzy relation *R* is particularized by constraining a set of input variables, we may focus our attention on a subset of variables of *R* which are designated as *output* variables and ask the question: What are the induced constraints on the output variables? As in the case of nonfuzzy relations, the answer is yielded by projecting the particularized relation on the cartesian product of the domains of output variables. Thus, for example, if the input variables are X_2, X_3 and X_5 , and the output variables are X_1 and X_4 , then the induced constraints on X_1 and X_4 are determined by the projection, *G*, of the particularized relation $R^*\ [(X_2, X_3, X_5)\ is\ F]$ on $U_1 \times U_2$. The relation which represents the projection in question is expressed as in [27]*.

$$G \triangleq {}_{X_1 \times X_2} R\ [(X_2, X_3, X_5)\ is\ F]\ , \tag{3.6}$$

with the understanding that $X_1 \times X_2$ in (3.6) should be interpreted as $U_1 \times U_2$. In more transparent terms, (3.6) may be restated as the *transduction*:

$$If (X_2, X_3, X_5)\ is\ F,\ then\ (X_1, X_2)\ is\ G\ , \tag{3.7}$$

* Here and elsewhere in the paper the aggregation operation min (*andsign*) is used as a default choice when no alternative (e.g., arithmetic mean, geometric mean, etc.) is specified.
* If *R* is a fuzzy relation, its projection on $U_1 \times U_2$ is obtained by deleting from *R* all columns other than X_1 and X_2 , and forming the union of the resulting tuples.

where G is given by (3.6). Equivalently, (3.7) may be interpreted as the instruction:

$$Read\ (X_{1,}\ X_2)\ given\ that\ (X_2, X_3, X_5)\ is\ F\ . \tag{3.8}$$

For example, the transduction represented by the expression

$$\underset{Name \times \mu}{RICH\ [Age\ is\ YOUNG;\ (Height,\ Weight)\ is\ BIG;\ Sex\ is\ MALE]}$$

may be interpreted as the fuzzy set of names of rich men who are young and big. It may also be interpreted in an imperative sense as the instruction: Read the name and grade of membership in the fuzzy set of rich men of all those who are young and big.

Remark. When the constraint set which is associated with an input variable, say X_1 , is a singleton, say {a}, we write simply

$$X = a$$

instead of X is a. For example,

$$\underset{Name \times \mu}{RICH\ [Age = 25;\ Weight = 136;\ Sex = Male]}$$

represents the fuzzy set of rich men whose age and weight are equal to 25 and 136, respectively.

Composition of Elastic Constraints

In testing the constituent relations in *EDF*, it is helpful to have a collection of standardized translation rules for computing the test score of a combination of elastic constraints $C_1, \ldots, C_k$ from the knowledge of the test scores of each constraint considered in isolation. For the most part, such rules are *default* rules in the sense that they are intended to be used in the absence of alternative rules supplied by the user.

For purposes of commonsense knowledge representation, the principal rules of this type are the following.*

* A more detailed discussion of such rules in the context of PRUF may be found in [57].

3.1. Rules pertaining to modification

If the test score for an elastic constraint C in a specified context is τ, then in the same context the test score for

$$(a)\ \ not\ C\ is\ 1-\tau \quad (negation) \tag{3.9}$$

$$(b)\ \ very\ C\ is\ \tau^2 \quad (concentration) \tag{3.10}$$

$$(c)\ \ more\ or\ less\ C\ is\ \tau^{\frac{1}{2}} \quad (diffusion)\ . \tag{3.11}$$

3.2. Rules pertaining to composition

If the test scores for elastic constraints C_1 and C_2 in a specified context are τ_1 and τ_2, respectively, then in the same context the test score for

$$(a)\ \ C_1\ and\ C_2\ is\ \tau_1\ andsign\ \tau_2 \quad (conjunction), where\ andsign \triangleq min. \tag{3.12}$$

$$(b)\ \ C_1\ or\ C_2\ is\ \tau_1\ orsign\ \tau_2 \quad (disjunction), where\ orsign \triangleq max. \tag{3.13}$$

$$(c)\ \ If C_1\ then\ C_2\ is\ 1\ andsign\ (1-\tau_1+\tau_2) \quad (implication)\ . \tag{3.14}$$

3.3. Rules pertaining to quantification

The rules in question apply to propositions of the general form $Q\ A's\ are\ B's$, where Q is a fuzzy quantifier, e.g., most, many, several, few, etc, and A and B are fuzzy sets, e.g., tall men, intelligent men, etc. As was stated earlier, when the fuzzy quantifiers in a proposition are implied rather than explicit, their suppression may be placed in evidence by referring to the proposition as a *disposition*. In this sense, the proposition *Overeating causes obesity* is a disposition which results from the suppression of the fuzzy quantifer *Most* in the proposition *Most of those who overeat are obese.*

To make the concept of a fuzzy quantifer meaningful, it is necessary to define a way of counting the number of elements in a fuzzy set or, equivalently, to determine its cardinality.

There are several ways in which this can be done [61]. For our purposes, it will suffice to employ the concept of a *sigma-count*, which is defined as follows.

Let F be a fuzzy subset of $U = \{u_1, \ldots, u_n\}$

expressed symbolically as

$$F = \mu_1/u_1 + \ldots + \mu_n/u_n = \Sigma_i \mu_i/u_i \tag{3.15}$$

or, more simply, as

$$F = \mu_1 u_1 + \ldots + \mu_n u_n\ , \tag{3.16}$$

in which the term μ_i / u_i, $i = 1, \ldots, n$, signifies that μ_i is the grade of membership of u_i in F, and the plus sign represents the union.*

The sigma-count of F is defined as the arithmetic sum of the μ_i , i.e.,

$$\Sigma Count(F) \underline{\Delta} \Sigma_i \mu_i \ , \ i = 1, \ldots, n \ , \tag{3.17}$$

with the understanding that the sum may be rounded, if need be, to the nearest integer. Furthermore, one may stipulate that the terms whose grade of membership falls below a specified threshold be excluded from the summation. The purpose of such an exclusion is to avoid a situation in which a large number of terms with low grades of membership become count-equivalent to a small number of terms with high membership.

The *relative sigma-count*, denoted by $\Sigma Count(F/G)$, may be interpreted as the proportion of elements of F which are in G. More explicitly,

$$\Sigma Count(F/G) = \frac{\Sigma Count(F \cap G)}{\Sigma Count(G)} \ , \tag{3.18}$$

where $F \cap G$, the intersection of F and G, is defined by

$$\mu_{F \cap G}(u) = \mu_F(u) \ andsign \ \mu_G(u) \ , \quad u \ member \ U \ . \tag{3.19}$$

Thus, in terms of the membership functions of F and G, the Thus, in terms of the membership functions of F and G, the relative sigma-count of F in G is given by

$$\Sigma Count(F/G) = \frac{\Sigma_i \mu_F(u_i) \ andsign \ \mu_G(u_i)}{\Sigma_i \mu_G(u_i)} \ , \tag{3.20}$$

The concept of a relative sigma-count provides a basis for interpreting the meaning of propositions of the form $Q \ A's \ are \ B's$, e.g., *Most young men are healthy*. More specifically, if the focal variable (ie., the constrained variable) in the proposition in question is taken to be the proportion of *B's* in *A's*, then the corresponding translation rule may be expressed as

$$Q \ A's \ are \ B's \rightarrow \Sigma Count(B/A) \ is \ Q \tag{3.21}$$

or, equivalently, as

$$Q \ A's \ are \ B's \rightarrow \Pi_X = Q \tag{3.22}$$

where

$$X = \frac{\Sigma_i \mu_A(u_i) \ andsign \ \mu_B(u_i)}{\Sigma_i \mu_A(u_i)} \ , \tag{3.23}$$

---

* In most cases, the context is sufficient to resolve the question of whether a plus sign should be interpreted as the union or the arithmetic sum.

As will be seen in the following section, the quantification rule (3.21) together with the other rules described in this section provide a basic conceptual framework for the representation of commonsense knowledge. We shall illustrate the representation process through the medium of several examples in which the meaning of a disposition is represented as a test on a collection of fuzzy relations in an explanatory database.

4. Representation of Dispositions

To clarify the difference between the conventional approaches to meaning representation and that described in the present paper, shall consider as our first example the disposition

$$d \triangleq \textit{Snow is white} \quad , \tag{4.1}$$

which, as was stated earlier, is frequently employed as an illustration in introductory expositions of truth-conditional semantics (see footnote on p. 5).

The first step in the representation process involves a restoration of the suppressed quantifiers in *d*. We shall assume that the intended meaning of *d* is conveyed by the proposition

$$p \triangleq \textit{Usually snow is white} \quad , \tag{4.2}$$

and, as an *EDF* for (4.2), we shall use (2.3), i.e.

$$EDF \triangleq WHITE\,[Sample\,;\mu] + USUALLY\,[Proportion\,;\mu] \quad . \tag{4.3}$$

Let $S_1, \ldots, S_m$ denote samples of snow and let τ_i, $i = 1, \ldots, m$, denote the degree to which the color of S_i matches white. Thus, τ_i may be interpreted as the test score for the constraint on the color of S_i which is induced by *WHITE*.

Using this notation, the steps in the testing procedure may be described as follows:

1. Find the proportion of samples whose color is white:

$$\rho = \frac{\Sigma Count\,(WHITE)}{m} \tag{4.4}$$

$$= \frac{\tau_1 + \ldots + \tau_m}{m}$$

2. Compute the degree to which ρ satisfies the constraint induced by *USUALLY*:

$$\tau = {}_{\mu}\, USUALLLY\,[Proportion = \rho] \tag{4.5}$$

In (4.5), τ represents the overall test score and the right-hand member signifies that the relation *USUALLY* is particularized by setting Proportion equal to ρ and projecting the resulting relation on μ. The meaning of *d*, then, is represented by the test procedure which leads to the value of τ.

Equivalently, the meaning of d may be represented as a possibility assignment equation. Specifically, let X denote the focal variable (i.e., the constrained variable) in p. Then we can write

$$d \rightarrow \Pi_X = USUALLY \tag{4.6}$$

where

$$X \underline{\underline{\Delta}} \frac{1}{m} \Sigma Count(WHITE) \ .$$

Example 2.

To illustrate the use of translation rules relating to modification, we shall consider the disposition

$$d \underline{\underline{\Delta}} Frenchmen\ are\ not\ very\ tall \ . \tag{4.7}$$

After restoration, the intended meaning of d is assumed to be represented by the proposition

$$p \underline{\underline{\Delta}} Most\ Frenchmen\ are\ not\ very\ tall \ . \tag{4.8}$$

To represent the meaning of p, we shall employ an *EDF* whose constituent relations are:

(4.9)

$$EDF \underline{\underline{\Delta}} \quad POPULATION\ [Name;\ Height] + TALL\ [Height;\ \mu\ ] + MOST\ [Proportion;\ \mu\ ] \ .$$

The relation *POPULATION* is a tabulation of *Height* as a function of *Name* for a representative group of Frenchmen. In *TALL*, μ is the degree to which a value of *Height* fits the description *tall*; and in *MOST*, μ is the degree to which a numerical value of *Proportion* fits the intended meaning of *most*.

The test procedure which represents the meaning of p involves the following steps:

1. Let $Name_i$ be the name of i^{th} individual in *POPULATION*. For each $Name_i$, $i = 1, \ldots, m$, find the height of $Name_i$:

   $$Height\ (Name_i) \underline{\underline{\Delta}}_{Height} POPULATION[Name = Name_i] \ .$$

2. For each $Name_i$, compute the test score for the constraint induced by *TALL:*

   $$\tau_i = {}_{\mu}TALL[Height = Height(Name_i)] \ .$$

3. Using the translation rules (3.9) and (3.10), compute the test score for the constraint induced by *NOT.VERY.TALL*:

   $$\tau'_i = 1 - \tau_i^2 \ .$$

4. Find the relative sigma-count of Frenchmen who are not very tall:

$$\rho \triangleq \Sigma Count(NOT.VERY.TALL\,/\,POPULATION)$$

$$= \frac{\Sigma_i \tau'_i}{m}\ .$$

5. Compute the test score for the constraint induced by *MOST*:

$$\tau = {}_{\mu}MOST[Proportion = \rho]\ . \tag{4.10}$$

The test score given by (4.10) represents the overall test score for *d*, and the test procedure which yields τ represents the meaning of *d*.

Example 3.

Consider the disposition

$$d \triangleq Overeating\ causes\ obesity \tag{4.11}$$

which after restoration is assumed to read*

$$p \triangleq Most\ of\ those\ who\ overeat\ are\ obese\ . \tag{4.12}$$

To represent the meaning of *p*, we shall employ an *EDF* whose constituent relations are:

(4.13)

$$EDF \triangleq \quad POPULATION\ [Name;\ Overeat;\ Obese] +$$
$$MOST\ [Proportion;\ \mu\ ]\ .$$

The relation *POPULATION* is a list of names of individuals, with the variables *Overeat* and *Obese* representing, respectively, the degrees to which *Name* overeats and is obese. In *MOST*, μ is the degree to which a numerical value of *Proportion* fits the intended meaning of *MOST*.

The test procedure which represents the meaning of *d* involves the following steps.

1. Let $Name_i$ be the name of i^{th} individual in *POPULATION*. For each $Name_i$, $i = 1, \ldots, m$, find the degrees to which $Name_i$ overeats and is obese:

$$\alpha_i \triangleq \mu_{OVEREAT}(Name_i) \triangleq {}_{Overeat}POPULATION[Name = Name_i] \tag{4.14}$$

and

$$\beta_i \triangleq \mu_{OBESE}(Name_i) \triangleq {}_{Obese}POPULATION[Name = Name_i]\ . \tag{4.15}$$

---

* It should be understood that (4.12) is just one of many possible interpretations of (4.11), with no implication that it constitutes a prescriptive interpretation of causality. See [48] for a thorough discussion of relevant issues.

2. Compute the relative sigma-count of *OBESE* in *OVEREAT*:

$$\rho \triangleq \Sigma Count(OBESE/OVEREAT) = \frac{\Sigma_i \alpha_i \, andsign \, \beta_i}{\Sigma_i \alpha_i} \quad . \tag{4.16}$$

3. Compute the test score for the constraint induced by *MOST*:

$$\tau = {}_{\mu}MOST[Proportion = \rho] \tag{4.17}$$

This test score represents the compatibility of *d* with the explanatory database.

Example 4.

Consider the disposition

$$d \triangleq Heavy\ smoking\ causes\ lung\ cancer \quad . \tag{4.18}$$

Although it has the same form as (4.11), we shall interpret it differently. Specifically, the restored proposition will be assumed to be expressed as

$$p \triangleq The\ incidence\ of\ cases\ of\ lung\ cancer\ among\ heavy\ smokers \tag{4.19}$$

$$is\ much\ higher\ than\ among\ those\ who\ are\ not\ heavy\ smokers \quad .$$

The *EDF* for this proposition is assumed to have the following constituents:

$$EDF \triangleq \begin{array}{l} POPULATION\ [Name;\ Heavy.Smoker;\ Lung.Cancer]\ + \\ MUCH.HIGHER\ [Proportion\ 1; Proportion\ 2;\ \mu\ ]\ . \end{array} \tag{4.20}$$

In *POPULATION*, *Heavy.Smoker* represents the degree to which *Name* is a heavy smoker, and the variable *Lung.Cancer* is 1 or 0 depending on whether or not *Name* has lung cancer. In *MUCH.HIGHER*, μ is the degree to which *Proportion 1* is much higher than *Proportion 2*.

The steps in the test procedure may be summarized as follows:

1. For each $Name_i$, $i = 1, \ldots, m$, determine the degree to which $Name_i$ is a heavy smoker:

$$\alpha_i \triangleq {}_{Heavy.Smoker}POPULATION[Name = Name_i] \quad . \tag{4.21}$$

Then, the degree to which $Name_i$ is not a heavy smoker is

$$\beta_i = 1 - \alpha_i \quad . \tag{4.22}$$

2. For each $Name_i$, determine if $Name_i$ has lung cancer:

$$\lambda_i \triangleq {}_{Lung.Cancer}POPULATION[Name = Name_i] \quad . \tag{4.23}$$

3. Compute the relative sigma-counts of those who have lung cancer among (a) heavy smokers; and (b) not heavy smokers:

$$\rho_1 = \Sigma Count(LUNG.CANCER / HEAVY.SMOKER)$$

$$= \frac{\Sigma_i \lambda_i \ andsign \ \alpha_i}{\Sigma_i \alpha_i}$$

$$\rho_2 = \Sigma Count(LUNG.CANCER / NOT.HEAVY.SMOKER)$$

$$= \frac{\Sigma_i \lambda_i \ andsign \ (1 - \alpha_i)}{\Sigma_i 1 - \alpha_i} \ .$$

4. Test the constraint induced by MUCH.HIGHER:

$$\tau = {}_{\mu}MUCH.HIGHER[Proportion1 = \rho_1 ; Proportion2 = \rho_2] \tag{4.24}$$

Example 5.

Consider the disposition

$$d \triangleq Small \ families \ are \ friendly \tag{4.25}$$

which we shall interpret as the proposition

$$p \triangleq In \ most \ small \ families \ almost \ all \ of \ the \ members \tag{4.26}$$

$$are \ friendly \ with \ one \ another.$$

It should be noted that the quantifier *most* in p is a second order fuzzy quantifier in the sense that it represents a fuzzy count of fuzzy sets (i.e., *small families*).

The *EDF* for p is assumed to be expressed by

(4.27)

EDF $\triangleq$ *POPULATION [Name; Family.Identifier] +*
SMALL [Number; μ *] +*
FRIENDLY [Name 1; Name 2; μ *] +*
MOST [Proportion; μ *] +*
ALMOST.ALL. [Proportion; μ *] .*

The relation *POPULATION* is assumed to be partitioned (by rows) into disjoint families $F_1, \ldots, F_k$. In *FRIENDLY*, μ is the degree to which *Name 1* is friendly toward *Name 2*, with *Name 1* $\neq$ *Name 2*.

The test procedure may be described as follows:

1. For each family, F_i , find the count of its members:

$$C_i \triangleq Count(POPULATION[Family.Identifier = F_i] \ . \tag{4.28}$$

2. For each family, test the constraint on C_i induced by *SMALL*:

$$\alpha_i \triangleq {}_{\mu}SMALL[Number = C_i] \ . \tag{4.29}$$

3. For each family, compute the relative sigma-count of its members who are friendly with one another:

$$\beta_i = \frac{1}{(C_i^2 - C_i)} \Sigma_{j,k} (_\mu FRIENDLY[Name1 = Name_j ; Name2 = Name_k]) \tag{4.30}$$

where $Name_j$ and $Name_k$ range over the members of F_i and $Name_i \neq Name_j$. The normalizing factor $C_i^2 - C_i$ represents the total number of links between pairs of distinct individuals in F_i.

4. For each family, test the constraint on β_i which is induced by *ALMOST.ALL*:

$$\gamma_i = {}_\mu ALMOST.ALL[Proportion = \beta_i] \ . \tag{4.31}$$

5. For each family, aggregate the test scores α_i and γ_i by using the min operator (*andsign*):

$$\delta_i \stackrel{\Delta}{=} \alpha_i \ andsign \ \gamma_i \ . \tag{4.32}$$

6. Compute the relative sigma-count of small families in which almost all members are friendly with one another:

$$\rho = \frac{1}{k}(\delta_1 + ... + \delta_k) \ . \tag{4.33}$$

7. Test the constraint on ρ induced by *MOST*:

$$\tau = {}_\mu MOST[Proportion = \rho] \ . \tag{4.35}$$

The value of τ given by (4.35) represents the compatibility of d with the explanatory database.

The foregoing examples are intended to illustrate the basic idea underlying our approach to the representation of commonsense knowledge, namely, the conversion of a disposition into a proposition, and the construction of a test procedure which acts on the constituent relations in an explanatory database and yields its compatibility with the restored proposition.

5. Inference from Dispositions *

A basic issue which will be addressed only briefly in the present paper is the following. Assuming that we have represented a collection of dispositions in the manner described above, how can an answer to a query be determined from the representations in question? In what follows, we shall consider a few problems of this type which are of relevance to the computation of certainty factors in expert systems [3, 15, 45, 49, 51].

* Inference from dispositions may be viewed as an alternative approach to default reasoning and non-monotonic logic [27, 29, 30, 40].

Conjunction of consequents

Consider two dispositions d_1 and d_2 which upon restoration become propositions of the general form

$$d_1 \rightarrow p_1 \triangleq Q_1\ A's\ are\ B's \tag{5.1}$$

$$d_2 \rightarrow p_2 \triangleq Q_2\ A's\ are\ C's\ . \tag{5.2}$$

Now assume that p_1 and p_2 appear as premises in the inference schema *

$$\begin{array}{l} Q_1\ A's\ are\ B's \\ \underline{Q_2\ A's\ are\ C's} \\ ?Q\ \ A's\ are\ (B\ and\ C)'s\ , \end{array} \tag{5.3}$$

in which *?Q* is a fuzzy quantifier which is to be determined. We shall refer to this schema as the *conjunction of consequents*.

As stated in the following assertion, the fuzzy quantifier Q is bounded by two fuzzy numbers. More specifically, on interpreting Q_1 and Q_2 as fuzzy numbers, we can assert that

$$0\ circle\ (Q_1\ circle\ Q_2\ circle\ 1) \leq Q \leq Q_1\ circle\ Q_2\ , \tag{5.4}$$

in which the operators $circle$,$circle$,$circle$, and $circle$, and the inequality $\leq$ are the extensions of $andsign$,$orsign$,$+$,$-$ and $\leq$, respectively, to fuzzy numbers [14].

Proof. We shall consider the upper bound first. To this end, it will suffice to show that

$$\Sigma Count\ (B \cap C\ /A) \leq \Sigma Count\ (B/A)\ andsign\ \Sigma Count\ (C/A)\ , \tag{5.5}$$

since, in view of (3.21), the fuzzy quantifiers Q, Q_1 and Q_2 may be regarded as fuzzy characterizations of the corresponding sigma-counts.

For convenience, let α_i , β_i and δ_i, $i = 1, \ldots, n$, denote, respectively, the grades of membership of μ_i in A, B and C. Then, on using (3.18) - (3.20), we may write

$$\Sigma Count(B \cap C\ /A) = \frac{\Sigma Count((A \cap B) \cap (A \cap C))}{\Sigma Count(A)} \tag{5.6}$$

$$= \frac{1}{\Sigma Count(A)} \left[\Sigma_i (\alpha_i\ andsign\ \beta_i)\ andsign\ (\alpha_i\ andsign\ \delta_i) \right]\ .$$

Now

$$\Sigma_i (\alpha_i\ andsign\ \beta_i)\ andsign\ (\alpha_i\ andsign\ \delta_i) \leq \Sigma_i (\alpha_i\ andsign\ \beta_i)$$

* This schema has a bearing on the rule of combination of evidence for conjunctive hypotheses in MYCIN [46].

$$\Sigma_i(\alpha_i \textit{ andsign } \beta_i) \textit{ andsign } (\alpha_i \textit{ andsign } \delta_i) \le \Sigma_i(\alpha_i \textit{ andsign } \delta_i)$$

and hence

$$\Sigma_i(\alpha_i \textit{ andsign } \beta_i) \textit{ andsign } (\alpha_i \textit{ andsign } \delta_i) \le \Sigma_i(\alpha_i \textit{ andsign } \beta_i) \textit{ andsign } \Sigma_i(\alpha_i \textit{ andsign } \delta_i) \ . \quad (5.7)$$

Consequently,

$$\Sigma Count(B \cap C\,/A) \le \frac{1}{\Sigma Count(A)}\left[\Sigma Count(A \cap B) \textit{ andsign } \Sigma Count(A \cap C)\right]$$

$$\le \Sigma Count(B/A) \textit{ andsign } \Sigma Count(C/A) \ ,$$

which is what we set out to establish.

To deduce the lower bound, we note that for any real numbers a, b, we have

$$a \textit{ andsign } b = a + b \ - \ a \textit{ orsign } b \ . \quad (5.8)$$

Consequently,

$$\frac{1}{\Sigma Count(A)}\left[\Sigma_i(\alpha_i \textit{ andsign } \beta_i) \textit{ andsign } (\alpha_i \textit{ andsign } \delta_i)\right] =$$

$$\frac{1}{\Sigma Count(A)}\left[\Sigma_i(\alpha_i \textit{ andsign } \beta_i) + \Sigma_i(\alpha_i \textit{ andsign } \delta_i) - \Sigma_i((\alpha_i \textit{ andsign } \beta_i) \textit{ orsign } (\alpha_i \textit{ andsign } \delta_i))\right] \ ,$$

and since

$$\alpha_i \ge (\alpha_i \textit{ andsign } \beta_i) \textit{ orsign } (\alpha_i \textit{ andsign } \delta_{i)}$$

it follows that

$$\Sigma Count(B \cap C\,/A) \ge \frac{1}{\Sigma Count(A)}\left[\Sigma_i\alpha_i \textit{ andsign } \beta_i + \Sigma_i\alpha_i \textit{ andsign } \delta_i - \Sigma_i\alpha_i\right]$$

or, equivalently,

$$\Sigma Count(B \cap C\,/A) \ge \Sigma Count(B/A) + (C/A) - 1 \ , \quad (5.9)$$

from which (5.4) follows by an application of the extension principle and the observation that the left-hand member of (5.5) must be non-negative [63].

In conclusion, the simple proof given above establishes the validity of the following inference schema, which, for convenience, will be referred to as the *consequent conjunction syllogism*:

$$\begin{array}{l} Q_1\ A's \textit{ are } B's \\ \underline{Q_2\ A's \textit{ are } C's} \\ Q\ A's \textit{ are } (B \text{ and } C)'s \ , \end{array} \quad (5.10)$$

where

$$0 \textit{ circle } (Q_1 \textit{ circle } Q_2 \textit{ circle } 1) \le Q \le Q_1 \textit{ circle } Q_2 \ .$$

As an illustration, from

$$p_1 \triangleq \textit{Most Frenchmen are not tall} \tag{5.11}$$

$$p_2 \triangleq \textit{Most Frenchmen are not short}$$

we can infer that

$$Q \textit{ Frenchmen are not tall and not short}$$

where

$$0 \textit{ circle } (2 \textit{ most circle } 1) \leq Q \leq \textit{most} \quad . \tag{5.12}$$

In the above example, the variable of interest is the proportion of Frenchmen who are *not tall and not short*. In a more general setting, the variable of interest may be a specified function of the variables constrained by the knowledge base. The following variation on (5.11) is intended to give an idea of how the value of the variable of interest may be inferred by an application of the extension principle [56].

Example 6.

Infer from the propositions

$$p_1 \triangleq \textit{Most Frenchmen are not tall} \tag{5.13}$$

$$p_2 \triangleq \textit{Most Frenchmen are not short}$$

the answer to the question

$$q \triangleq \textit{What is the average height of a Frenchman}? \tag{5.14}$$

Because of the simplicity of p_1 and p_2, the constraints induced by the premises may be found directly. Specifically, let $h_1, \ldots, h_n$ denote the heights of $Frenchman_1, \ldots, Frenchman_n$, respectively. Then, the test scores associated with the constraints in question may be expressed as

$$\tau_1 = \mu_{ANT\ MOST}(\frac{1}{n}\Sigma_i \mu_{TALL}(h_i)) \tag{5.15}$$

and

$$\tau_2 = \mu_{ANT\ MOST}(\frac{1}{n}\Sigma_i \mu_{SHORT}(h_i)) \quad , \tag{5.16}$$

where *ANT* is an abbreviation for *antonym*, i.e.,

$$\mu_{ANT\ MOST}(u) = \mu_{MOST}(1-u) \quad , \quad u \varepsilon [0,1] \quad , \tag{5.17}$$

and μ_{TALL} and μ_{SHORT} are the membership functions of *TALL* and *SHORT*, respectively. Correspondingly, the overall test score may be expressed as

$$\tau = \tau_1 \; andsign \; \tau_2 \quad .$$

Now, the average height of a Frenchman and hence the answer to the question is given by

$$ans(q) = \frac{1}{n} \Sigma_i h_i \quad . \tag{5.18}$$

Consequently, the possibility distribution of *ans(q)* is given by the solution of the nonlinear program

$$\mu_{ans(q)}(h) = \max_{h_1, \ldots, h_n}(\tau) \tag{5.19}$$

subject to the constraint

$$h = \frac{1}{n} \Sigma_i h_i \quad . \tag{5.20}$$

Alternatively, a simpler but less informative answer may be formulated by forming the intersection of the possibility distributions of *ans(q)* which are induced separately by p_1 and p_2. More specifically, let $\Pi_{ans(q)|p_1}$, $\Pi_{ans(q)|p_2}$, $\Pi_{ans(q)|p_1 \; andsign \; p_2}$ be the possibility distributions of *ans(q)* which are induced by p_1 , p_2 , and the conjunction of p_1 and p_2, respectively. Then, by using the minimax inequality [54], it can readily be shown that

$$\Pi_{ans(q)|p_1} \; cap \; \Pi_{ans(q)|p_2} \; supset \; \Pi_{ans(q)|p_1 \; andsign \; p_2} \quad , \tag{5.21}$$

and hence we can invoke the entailment principle [58] to validate the intersection in question as the possibility distribution of *ans(q)*. For the example under consideration, the desired possibility distribution is readily found to be given by

$$Poss\{ans(q) = h\} = \mu_{ANT \; MOST}(\mu_{TALL}(h)) \; andsign \; \mu_{ANT \; MOST}(\mu_{SHORT}(h)) \quad . \tag{5.22}$$

Chaining of dispositions

As in (5.1) and (5.2), consider two dispositions d_1 and d_2 which upon restoration become propositions of the general form

$$d_1 \rightarrow p_1 \stackrel{\Delta}{=} Q_1 \; A's \; are \; B's$$

$$d_2 \rightarrow p_2 \stackrel{\Delta}{=} Q_2 \; B's \; are \; C's \quad .$$

An ordered pair, $(p_1, \; p_2)$, of propositions of this form will be said to form a *chain*. More generally, an *n-ary chain* may be represented as an ordered n-tuple

$$(Q_1 \; A_1's \; are \; B_1's \; , \; Q_2 \; A_2's \; are \; B_2's \; , \ldots , Q_n \; A_n's \; are \; B_n's) \quad , \tag{5.23}$$

in which $B_1 = A_2, B_2 = A_3, \ldots, B_{n-1} = A_n$.

Now assume that p_1 and p_2 appear as premises in the inference schema

$$\begin{array}{ll} Q_1\ A's\ are\ B's & (major\ premise) \\ \underline{Q_2\ B's\ are\ C's} & (minor\ premise) \\ ?Q\ A's\ are\ C's & (conclusion) \end{array} \tag{5.24}$$

in which *?Q* is a fuzzy quantifier which is to be determined.

A basic rule of inference which is established in [63] and which has a direct bearing - as we shall see presently - on the determination of *Q*, is the *intersection/product syllogism*

$$\begin{array}{l} Q_1\ A's\ are\ B's \\ \underline{Q_2\ (A \text{ and } B)'s\ are\ C's} \\ (Q_1\ citimes\ Q_2)\ A's\ are\ (B \text{ and } C)'s \quad , \end{array} \tag{5.25}$$

in which Q_1 *citimes* Q_2 is a fuzzy number which is the fuzzy product of the fuzzy numbers Q_1 and Q_2. For example, as a special case of (5.25), we may write

$$\begin{array}{l} Most\ students\ are\ single \\ \underline{A\ little\ more\ than\ a\ half\ of\ single\ students\ are\ male} \\ (Most\ citimes\ A\ little\ more\ than\ a\ half\,)\ of\ students\ are\ single\ and\ male \quad . \end{array} \tag{5.26}$$

Since the intersection of *B* and *C* is contained in *C*, the following corollary of (5.25) is its immediate consequence

$$\begin{array}{l} Q_1\ A's\ are\ B's \\ \underline{Q_2\ (A \text{ and } B)'s\ are\ C's} \\ \geq (Q_1\ citimes\ Q_2)\ A's\ are\ C's \quad , \end{array} \tag{5.27}$$

where the fuzzy number $\geq (Q_1\ citimes\ Q_2)$ should be read as *at least* $(Q_1\ citimes\ Q_2)$, with the understanding that $\geq (Q_1\ citimes\ Q_2)$ represents the composition of the binary nonfuzzy relation $\geq$ with the unary fuzzy relation $(Q_1\ citimes\ Q_2)$. In particular, if the fuzzy quantifiers Q_1 and Q_2 are monotone nondecreasing (e.g., when $Q_1 = Q_2 \triangleq most$), then as is stated in [63],

$$\geq (Q_1\ citimes\ Q_2) = Q_1\ citimes\ Q_2 \quad , \tag{5.28}$$

and (5.27) becomes

$$\begin{array}{l} Q_1\ A's\ are\ B's \\ \underline{Q_2\ (A \text{ and } B)'s\ are\ C's} \\ (Q_1\ citimes\ Q_2)\ A's\ are\ C's \quad . \end{array} \tag{5.29}$$

There is an important special case in which the premises in (5.29) form a chain. Specifically, if *B subset A*, then

$$A \cap B = B$$

and (5.29) reduces to what will be referred to as the *product chain rule*, namely,

$$\begin{array}{l} Q_1\ A's\ are\ B's \\ \underline{Q_2\ B's\ are\ C's} \\ (Q_1\ citimes\ Q_2)\ A's\ are\ C's\ \ . \end{array} \tag{5.30}$$

In this case, the chain (Q_1 $A's$ are $B's$, Q_2 $B's$ are $C's$) will be said to be *product transitive.* *

As an illustration of (5.30), we can assert that

$$\begin{array}{l} Most\ students\ are\ undergraduates \\ \underline{Most\ undergraduates\ are\ young} \\ Most^2\ students\ are\ young\ \ , \end{array}$$

where $Most^2$ represents the product of the fuzzy number *Most* with itself.

Chaining under reversibility

An important chaining rule which is approximate in nature relates to the case where the major premise in the inference chain

$$\begin{array}{l} Q_1\ A's\ are\ B's \\ \underline{Q_2\ B's\ are\ C's} \\ Q\ A's\ are\ C's \end{array} \tag{5.31}$$

is *reversible* in the sense that

$$Q_1\ A's\ are\ B's <\!\!\to Q_1\ B's\ are\ A's\ \ , \tag{5.32}$$

where $<\!\!\to$ denotes approximate semantic equivalence [57]. -- For example,

$$Most\ American\ cars\ are\ big <\!\!\to Most\ big\ cars\ are\ American\ \ . \tag{5.33}$$

Under the assumption of reversibility, the following syllogism holds in an approximate sense

$$Q_1\ A's\ are\ B's \tag{5.34}$$

* More generally, an n-ary chain (Q_1 $A_1's$ *are* $B_1's$, . . . , Q_n $A_n's$ *are* $B_n's$) will be said to be *product transitive* if from the premises which constitute the chain it may be inferred that $\geq (Q_1\ citimes \cdots citimes\ Q_n)\ A_1's\ are\ B_n's$.

$$\underline{Q_2\ B's\ are\ C's}$$

$$\geq (0\ circle (Q_1\ circle\ Q_2\ circle\ 1))\ A's\ are\ C's\ \ .$$

We shall refer to this syllogism as the *R-rule*.

To demonstrate the approximate validity of this rule we shall first establish the following lemma.

Lemma.

$$If\ \Sigma Count(A) = \Sigma Count(B) \tag{5.35}$$

and

$$\Sigma Count(B/A) \geq q_1 \tag{5.36}$$

$$\Sigma Count(C/B) \geq q_2 \tag{5.37}$$

then

$$\Sigma Count(C/A) \geq (0\ orsign\ (q_1 + q_2 - 1)\ \ . \tag{5.38}$$

Proof.

We have

$$\Sigma Count(B/A) = \frac{\Sigma Count(A \cap B)}{\Sigma Count(A)}$$

and

$$\Sigma Count(C/B) = \frac{\Sigma Count(B \cap C)}{\Sigma Count(B)}$$

$$= \frac{\Sigma Count(B \cap C)}{\Sigma Count(A)}$$

in virtue of (5.35).

For simplicity, we shall denote $\mu_A(u_i)$, $\mu_B(u_i)$ and $\mu_C(u_i)$, $i = 1, \ldots, n$, by α_i, β_i, and γ_i, respectively. Then,

$$\frac{\Sigma Count(A \cap B)}{\Sigma Count(A)} = \frac{\Sigma_i\ \alpha_i\ andsign\ \beta_i}{\Sigma_i\ \alpha_i} \geq q_1 \tag{5.39}$$

$$\frac{\Sigma Count(B \cap C)}{\Sigma Count(A)} = \frac{\Sigma_i\ \beta_i\ andsign\ \gamma_i}{\Sigma_i\ \alpha_i} \geq q_2 \tag{5.40}$$

and hence

$$\frac{\Sigma_i\ (\alpha_i\ \mathit{andsign}\ \ \beta_i + \beta_i\ \mathit{andsign}\ \gamma_i)}{\Sigma_i\ \alpha_i} \geq q_1 + q_2\ \ , \tag{5.41}$$

and

$$\frac{\Sigma_i\ (\alpha_i\ \mathit{andsign}\ \ \beta_i + \beta_i\ \mathit{andsign}\ \gamma_i - \alpha_i)}{\Sigma_i\ \alpha_i} \geq q_1 + q_2 - 1\ \ . \tag{5.42}$$

Since

$$\frac{\Sigma Count(C \cap A)}{\Sigma Count(A)} = \frac{\Sigma_i\ \alpha_i\ \mathit{andsign}\ \gamma_i}{\Sigma_i\ \alpha_i}\ \ , \tag{5.43}$$

it follows from (5.42) and (5.43) that, to establish (5.38), it will suffice to show that

$$\Sigma_i(\alpha_i\ \mathit{andsign}\ \gamma_i) \geq \Sigma_i(\alpha_i\ \mathit{andsign}\ \beta_i + \beta_i\ \mathit{andsign}\ \gamma_i - \alpha_i)\ \ , \tag{5.44}$$

or, equivalently,

$$\Sigma_i\left[(\alpha_i - \alpha_i\ \mathit{andsign}\ \beta_i) + (\beta_i - \beta_i\ \mathit{andsign}\ \gamma_i)\right] \geq \Sigma_i(\alpha_i - \alpha_i\ \mathit{andsign}\ \gamma_i)\ \ , \tag{5.45}$$

which is a consequence of the equality

$$\Sigma_i\alpha_i = \Sigma_i\beta_i\ \ , \tag{5.46}$$

which in turn follows from (5.35).

Now, to establish (5.45) it is sufficient to show that, for each i, $i = 1, \ldots, n$, we have

$$(\alpha_i - \alpha_i\ \mathit{andsign}\ \beta_i) + (\beta_i - \beta_i\ \mathit{andsign}\ \gamma_i) \geq (\alpha_i - \alpha_i\ \mathit{andsign}\ \gamma_i)\ \ , \tag{5.48}$$

in which the summands as well as the right-hand member are non-negative. To this end, we shall verify that (5.48) holds for all possible values of α_i, β_i and γ_i.

Case 1.

Consider all values of α_i, β_i and γ_i which satisfy the inequality $\alpha_i \leq \gamma_i$. In this case, the right-hand member of (5.48) is zero and thus the inequality is verified.

Case 2.

$\alpha_i > \gamma_i$. In this case, we shall consider four subcases.

(i) $\alpha_i \leq \beta_i$, $\beta_i \leq \gamma_i$, which contradicts $\alpha_i > \gamma_i$.

(ii) $\alpha_i > \beta_i$, $\beta_i \leq \gamma_i$, which verifies the inequality.

(iii) $\alpha_i \leq \beta_i$, $\beta_i > \gamma_i$, which verifies the inequality.

(iv) $\alpha_i > \beta_i$, $\beta_i > \gamma_i$, which verifies the inequality.

This concludes the proof of the lemma.

Now, if condition (5.35), i.e.,

$$\Sigma_i \alpha_i = \Sigma_i \beta_i \quad ,$$

were not needed to prove the lemma, we could invoke the extension principle [56] to extend the inequality

$$\frac{\Sigma Count(C/A)}{\Sigma Count(A)} \geq 0 \; orsign \; (q_1 + q_2 - 1) \quad ,$$

which holds for real numbers, to

$$Q \geq 0 \; circle \; (Q_1 \; circle \; Q_2 \; circle \; 1) \quad , \tag{5.49}$$

which holds for the fuzzy quantifiers in (5.31). As it is, the assumption of reversibility, i.e.,

$$\frac{\Sigma Count(A \cap B)}{\Sigma Count(A)} \; is \; Q_1$$

$$\frac{\Sigma Count(B \cap A)}{\Sigma Count(B)} \; is \; Q_1$$

implies the equality

$$\Sigma Count(A) = \Sigma Count(B)$$

only in an approximate sense. Consequently, as was stated earlier, the *R-rule* (5.34) also holds only in an approximate sense. The question of how this sense could be defined more precisely presents a nontrivial problem which will not be addressed in this paper.

Concluding Remark

The point of departure in this paper is the idea that commonsense knowledge may be regarded as a collection of dispositions. Based on this idea, the representation of commonsense knowledge may be reduced, in most cases, to the representation of fuzzily-quantified propositions through the use of test-score semantics. Then, the rules of inference of fuzzy logic may be employed to deduce answers to questions which relate to the information resident in a knowledge base.

The computational framework for dealing with commonsense knowledge which is provided by fuzzy logic is of relevance to the management of uncertainty in expert systems. The advantage of employing fuzzy logic in this application-area is that it provides a systematic framework for syllogistic reasoning and thus puts on a firmer basis the derivation of combining functions for uncertain evidence. The consequent conjunction syllogism which we established in Section 5 is, in effect, an example of such a combining function. What it demonstrates, however, is that, in general, combining functions cannot be expected to yield real-valued probabilities or certainty factors, as they do in

MYCIN, PROSPECTOR and other expert systems. Thus, in general, the value returned by a combining function should be a fuzzy number or an n-tuple of such numbers.

REFERENCES AND RELATED PUBLICATIONS

1. Adams, E.W. and Levine, H.F., On the uncertainties transmitted from premises to conclusions in deductive inferences. *Synthese 30* (1975), 429-460.

2. Baldwin, J.F. and Zhou, S.Q., A fuzzy relational inference language Report No. EM/FS132, University of Bristol, 1982.

3. Barr, A. and Feigenbaum, E.W., *The Handbook of Artificial Intelligence*, vols. 1, 2 and 3. Los Altos: Kaufmann, 1982.

4. Barwise, J. and Cooper, R., Generalized quantifiers and natural language, *Linguistics and Philosophy* 4 (1981) 159-219.

5. Bellman, R.E. and Zadeh, L.A., Local and fuzzy logics, in: *Modern Uses of Multiple-Valued Logic* , Epstein, G., (ed.), Dordrecht: Reidel, 103-165, 1977.

6. Blanchard, N., Theories cardinales et ordinales des ensembles flou: les multiensembles. Thesis, University of Claude Bernard, Lyon, 1981.

7. Brachman, R.J. and Smith, B.C., *Special Issue on Knowledge Representation*, SIGART 70, 1980.

8. Cooper, W.S., *Logical Linguistics*. Dordrecht: Reidel, 1978.

9. Cresswell, M.J., *Logic and Languages.* London: Methuen, 1973.

10. Cushing, S., *Quantifier Meanings - A Study in the Dimensions of Semantic Competence.* Amsterdam: North-Holland, 1982.

11. Date, C.J., *An Introduction to Database Systems.* Reading: Addison-Wesley, 1977.

12. Davis, R. and Lenat, D.B., *Knowledge-Based Systems in Artificial Intelligence*. New York: McGraw-Hill, 1982.

13. DeLuca, A. and Termini, S., A definition of non-probabilistic entropy in the setting of fuzzy sets theory, *Information and Control* 20 (1972) 301-312.

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

15. Duda, R.O., Hart, P.E., Konolige, K. and Reboh, R., A computer-based consultant for mineral exploration, Tech. Report, SRI International, Menlo Park, CA, 1979.

16. Feigenbaum, E.A., The art of artificial intelligence: Themes and case studies in knowledge engineering, *Proc. IJCAI5,* 1977.

17. Findler, N.V. (ed.), *Associative Networks: Representation and Use of Knowledge by Computer*. New York: Academic Press, 1979.

18. Gaines, B.R., Logical foundations for database systems, *Int. J. Man-Machine Studies 11* (1979) 481-500.

19. Goguen, J.A., The logic of inexact concepts, *Synthese 19* (1969) 325-373 .

20. Hofmann, T.R., Qualitative terms for quantity, In *Proc. 6th European Meeting on Cybernetics and Systems Research*, R. Trappl, (ed.). Amsterdam: North-Holland (1982).

21. Hoover, D.N., Probability logic, *Annals Math. Logic 14*, (1978) 287-313.

22. Ishizuka, M., Fu, K.S. and Yao, J.T.P., A rule-based inference with fuzzy set for structural damage assessment. *Fuzzy Information and Decision Processes*, Gupta, M. and Sanchez, E (eds.) Amsterdam: North-Holland (1982).

23. Keenan, E.L., Quantifier structures in English. *Foundations of Language 7*, (1971) 255-336.

24. Klement, E.P., Schwyhla, W. and Lowen, R., Fuzzy probability measures. *Fuzzy Sets and Systems 5*, (1981) 83-108.

25. Klement, E.P., Operations on fuzzy sets and fuzzy numbers related to triangular norms. *Proc. 11th Conf. Multiple-Valued Logic*, Univ. of Oklahoma, Norman, pp. 218-225, 1981.

26. Mamdani, E.H. and Gaines, B.R., *Fuzzy Reasoning and its Applications.* London: Academic Press, 1981.

27. McCarthy, J., Circumscription: A non-monotonic inference rule. *Artificial Intelligence 13*, (1980) 27-40.

28. McCawley, J.D., *Everything that Linguists have Always Wanted to Know about Logic.* Chicago: University of Chicago Press, 1981.

29. McDermott, D. V. and Doyle, J., Non-monotonic logic, I. *Artificial Intelligence 13*, (1980) 41-72.

30. McDermott, D. V., Non-monotonic logic, II: non-monotonic modal theories. *J. Assoc. Comp. Mach. 29*, (1982) 33--57.

31. Michalski, R., Carbonell, J.G. and Mitchell, T.M., *Machine Learning*. Palo Alto: Tioga Press, 1983.

32. Miller, G.A. and Johnson-Laird, P.N., *Language and Perception*. Cambridge: Harvard University Press, 1976.

33. Mizumoto, M., Umano, M. and Tanaka, K., Implementation of a fuzzy-set-theoretic data-structure system. *3rd Int. Conf. Very Large Data Bases*, Tokyo, 1977.

34. Mizumoto, M.,Fukame, S. and Tanaka, K., Fuzzy reasoning Methods by Zadeh and Mamdani, and improved methods, *Proc. Third Workshop on Fuzzy Reasoning*, Queen Mary College, London, 1978.

35. Moisil, G.C., *Lectures on the Logic of Fuzzy Reasoning*, Bucarest: Scientific Editions, (1975).

36. Noguchi, K., Umano, M., Mizumoto, M. and Tanaka, K., Implementation of fuzzy artificial intelligence language FLOU, *Technical Report on Automation and Language of IECE*, 1976.

37. Orlov, A.I., *Problems of Optimization and Fuzzy Variables* Moscow: Znaniye, 1980.

38. Partee, B., *Montague Grammar*. New York: Academic Press, 1976.

39. Peterson, P., On the logic of *few, many* and *most*. *Notre Dame J. Formal Logic 20* (1979) 155-179.

40. Reiter, R. and Criscuolo, G., Some representational issues in default reasoning, *Computers and Mathematics 9* (1983) 15-28.

41. Rescher, N., *Plausible Reasoning*. Amsterdam: Van Gorcum, 1976.

42. Rich, E. *Artificial Intelligence*. New York: McGraw-Hill, 1983.

43. Rieger, C., Conceptual memory, in *Conceptual Information Processing*, Schank, R. (ed.). Amsterdam: North-Holland, 1975.

44. Schubert, L.K., Goebel, R.G. and Cercone, N., The structure and organization of a semantic net for comprehension and inference, in: *Associative Networks*, Findler, N.V., (ed.). New York: Academic Press, 122-178, 1979.

45. Shortliffe, E.H. and Buchanan, B., A model of inexact reasoning in medicine, *Mathematical Biosciences 23* (1975) 351-379.

46. Shortliffe, E.H., *Computer-Based Medical Consultations: MYCIN*. New York: American Elsevier, 1976.

47. Sugeno, M., Fuzzy measures and fuzzy integrals: a survey, in *Fuzzy Automata and Decision Processes*, Gupta, M.M., Saridis, G.N. and Gaines, B.R., (eds.), pp. 89-102, Amsterdam: North-Holland, 1977.

48. Suppes, P., *A Probabilistic Theory of Causality*. Amsterdam: North-Holland, 1970.

49. Szolovits, P. and Pauker, S.G., Categorical and probabilistic reasoning in medical diagnosis, *Artificial Intelligence 11* (1978) 115-144.

50. Terano, T. and Sugeno, M., Conditional fuzzy measures and their applications. *Fuzzy Sets and Their Applications to Cognitive and Decision Processes* Zadeh, L.A., Fu, K.S., Tanaka, K. and Shimura, M. (eds.), pp. 151-170, New York: Academic Press, 1975.

51. Weiss, S.M. and Kulikowski, C.A., EXPERT: A system for developing consultation models, *Proc. IJCAI6*, 1979.

52. Yager, R.R., Quantified propositions in a linguistic logic, in: *Proceedings of the 2nd International Seminar on Fuzzy Set Theory*, Klement, E.P., (ed.). Johannes Kepler University, Linz,

Austria, 1980.

53. Zadeh, L.A., Probability measures of fuzzy events. *J. Math. Anal. Appl. 23* (1968) 421-427.

54. Zadeh, L.A., Similarity relations and fuzzy orderings, *Inf. Sci. 3* (1971) 177-200.

55. Zadeh, L.A., Fuzzy logic and approximate reasoning (in memory of Grigore Moisil), *Synthese 30* (1975) 407-428.

56. Zadeh, L.A., The concept of a linguistic variable and its application to approximate reasoning. *Inform. Sci. 8*, (1975), 199-249, 301-357, *9*, 43-80.

57. Zadeh, L.A., PRUF -- a meaning representation language for natural languages, *Int. J. Man-Machine Studies 10* (1978) 395-460.

58. Zadeh, L.A., A theory of approximate reasoning, *Electronics Research Laboratory Memorandum M77/58* University of California, Berkeley, 1977. Also in: *Machine Intelligence 9*, Hayes, J.E., Michie, D. and Kulich, L.I., (eds). New York: Wiley 149:194, 1979.

59. Zadeh, L. A., Inference in fuzzy logic. *Proc. 10th Inter. Symp. on Multiple-valued Logic*, Northwestern University, (1980), 124-131.

60. Zadeh, L.A., Possibility theory and soft data analysis. *Electronic Res. Laboratory Memorandum M79/59* University of California, Berkeley, (1980); also in *Mathematical Frontiers of the Social and Policy Sciences* Cobb, L., and Thrall, R.M. (eds.). Boulder: Westview Press, 69-129, 1981.

61. Zadeh, L.A., Test-score semantics for natural languages and meaning-representation via PRUF, *Tech.Note 247, AI Center, SRI International,* Menlo Park, CA, 1981. Also in *Empirical Semantics*, Rieger, B.B., (ed.). Bochum: Brockmeyer, 281-349, 1981.

62. Zadeh, L.A., Fuzzy probabilities and their role in decision analysis, *Proc. 4th MIT/ONR Workshop on Command, Control and Communications* M.I.T., (1981), pp. 159-179.

63. Zadeh, L.A., A computational approach to fuzzy quantifiers in natural languages, *Computers and Mathematics 9* (1983) 149-184.

64. Zimmer, A., Some experiments concerning the fuzzy meaning of logical quantifiers. In *General Surveys of Systems Methodology*, Troncoli, L., (ed.), pp. 435-441. Society for General Systems Research, Louisville, 1982.

65. Zimmermann, H.-J., and Zysno, P., Latent connectives in human decision making, *Fuzzy Sets and Systems 4* (1980), 37-52.