

Copyright © 1972, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

A FUZZY-SET-THEORETIC INTERPRETATION
OF LINGUISTIC HEDGES

by

L. A. Zadeh

Memorandum No. ERL-M335

1 April 1972

(Handwritten signature)

A FUZZY-SET-THEORETIC INTERPRETATION
OF LINGUISTIC HEDGES

by

L. A. Zadeh

Memorandum No. ERL-M335

1 April 1972

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

A Fuzzy-Set-Theoretic Interpretation of Linguistic Hedges*

L. A. Zadeh

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

Summary

A basic idea suggested in this paper is that a linguistic hedge such as very, more or less, much, essentially, slightly, etc. may be viewed as an operator which acts on the fuzzy set representing the meaning of its operand. For example, in the case of the composite term very tall man, the operator very acts on the fuzzy meaning of the term tall man.

To represent a hedge as an operator, it is convenient to define several elementary operations on fuzzy sets from which more complicated operations may be built up by combination or composition. In this way, an approximate representation for a hedge can be expressed in terms of such operations as complementation, intersection, concentration, dilation, contrast intensification, fuzzification, accentuation, etc.

Two categories of hedges are considered. In the case of hedges of Type I, e.g., very, much, more or less, slightly, etc., the hedge can be approximated by an operator acting on a single fuzzy set. In the case

* Research sponsored by the Army Research Office--Durham, Grant DA-ARO-D-31-124-71-G174, and the Joint Services Electronics Program, Contract F44620-71-C-0087.

of hedges of Type II, e.g., technically, essentially, practically, etc., the effect of the hedge is more complicated, requiring a description of the manner in which the components of its operand are modified. If, in addition, the characterization of a hedge requires a consideration of a metric or proximity relation in the space of its operand, then the hedge is said to be of Type IP or IIP, depending on whether it falls into category I or II.

The approach is illustrated by constructing operator representations for several relatively simple hedges such as very, more or less, much, slightly, essentially, etc. More complicated hedges whose effect is strongly context-dependent, require the use of a fuzzy-algorithmic mode of characterization which is more qualitative in nature than the approach described in the present paper.

A Fuzzy-Set-Theoretic Interpretation of Linguistic Hedges^{*}

L. A. Zadeh

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

1. Introduction

Roughly speaking, a fuzzy set is a class with unsharp boundaries, that is, a class in which the transition from membership to non-membership is gradual rather than abrupt. In this sense, the class of tall men is a fuzzy set, as are the classes of beautiful women, young men, red flowers, small cars, etc.

Fuzziness plays an essential role in human cognition because most of the classes encountered in the real world are fuzzy - some only slightly and some markedly so. The pervasiveness of fuzziness in human thought processes suggests that much of the logic behind human reasoning is not the traditional two-valued or even multi-valued logic, but a logic with fuzzy truths, fuzzy connectives and fuzzy rules of inference. Indeed, it may be argued that it is the ability of the human brain to manipulate fuzzy concepts that distinguishes human intelligence from machine intelligence. And yet, despite its fundamental importance, fuzziness has not been accorded much attention in the scientific

* Research sponsored by the Army Research Office--Durham, Grant DA-ARO-D-31-124-71-G174, and the Joint Services Electronics Program, Contract F44620-71-C-0087. This work was inspired and influenced by many discussions with Professor G. Lakoff concerning the meaning of hedges and their interpretation in terms of fuzzy sets. Lakoff's analysis of hedges and related problems may be found in References [1], [2].

literature,^{*} partly because it is antithetic to the deeply entrenched traditions of scientific thinking based on Aristotelian logic and oriented toward exact quantitative analysis, and partly because fuzziness is susceptible of confusion with randomness. In fact, fuzziness and randomness are distinct phenomena which require different modes of treatment and mathematical analysis.

The theory of fuzzy sets^{**} represents an attempt at constructing a conceptual framework for a systematic treatment of fuzziness in both quantitative and qualitative ways. A basic concept in this theory is that of a fuzzy subset A of a universe of discourse U , with A characterized by a membership function $\mu_A(x)$ which associates with each point x in U its "grade of membership" in A . Usually, but not necessarily, $\mu_A(x)$ is assumed to range in the interval $[0,1]$, so that the grade of membership is a number between 0 and 1, with 0 and 1 corresponding to non-membership and full membership, respectively. For example, if U is the set of integers from 0 to 100, then the grade of membership of a man who is 23 years old in the class of young men might be specified to be 0.9. In general, the grades of membership are subjective, in the sense that their specification is a matter of definition rather than objective experimentation or analysis. It should be noted that, although $\mu_A(x)$ may be interpreted as the truth-value of the statement "x belongs to A," it is more natural to view it simply as a grade of membership because the statement "x belongs to A" is not meaningful when A is a fuzzy set.

* Discussions of vagueness and related questions from a philosophical point of view may be found in References [11]-[15].

** Topics in the theory of fuzzy sets which are relevant to the subject of the present paper are discussed in References [3]-[10].

The pervasive fuzziness of the semantics and - to a lesser extent - the syntax of natural languages suggests that some aspects of linguistic theory might be amenable to analysis by techniques derived from the theory of fuzzy sets. A few preliminary steps in this direction were taken in [5], [6] and [7] with the aim of constructing a framework for a quantitative approach to fuzzy semantics and fuzzy syntax. In what follows, the focus of attention will be on a more specific application, namely, the construction of a fuzzy-set-theoretic interpretation of hedges, e.g., "very," "somewhat," "quite," "much," "more or less," "sort of," "essentially," etc., as operators acting on fuzzy subsets of the universe of discourse. Our main concern, however, will be with the basic aspects of this interpretation rather than with detailed analyses of particular hedges.

The possibility of defining hedges as operators acting on fuzzy sets provides a basis for a better understanding of their role in natural languages. More important, it suggests a way of constructing a system of both natural and artificial hedges which could be used to devise algorithmic languages for the description of the behavior of complex systems. Such languages might find significant applications in psychology, sociology, political science, physiology, economics, management science, information retrieval and other fields in which system behavior is frequently too complex or too ill-defined to admit of analysis in conventional mathematical terms.

2. Fuzzy Sets and Languages - Notation and Terminology

Let U be a universe of discourse, i.e., a collection of objects

denoted generically by y . For our purposes, it will be convenient to regard a language, L , as a correspondence between a set of terms, T , and the universe of discourse, U .^{*} This correspondence will be assumed to be defined by a naming relation, N , which associates with each term x in T and each object y in U the degree, $\mu_N(x,y)$, to which x applies to y .^{**} $\mu_N(x,y)$ will be assumed to be a number in the interval $[0,1]$, so that N is, in effect, a fuzzy relation from T to U .^{***}

A term may be atomic, e.g., $x = \underline{\text{red}}$, $x = \underline{\text{barn}}$, $x = \underline{\text{tall}}$, or composite, in which case it is a concatenation of atomic terms, e.g., $x = \underline{\text{red barn}}$, $x = \underline{\text{tall man}}$, $x = \underline{\text{very beautiful woman}}$, $x = \underline{\text{tall and dark}}$, $x = \underline{\text{not very sweet or sour}}$, etc. In either case, when x is chosen to be a particular term in T , say $x = \underline{\text{red}}$, the function $\mu_N(\underline{\text{red}},y)$ defines a fuzzy subset of U whose membership function $\mu_{\underline{\text{red}}}(y)$ is given by

$$\mu_{\underline{\text{red}}}(y) = \mu_N(\underline{\text{red}},y) \quad (1)$$

This fuzzy subset, denoted by $M(\underline{\text{red}})$, is defined to be the meaning^{****}

* A more detailed discussion may be found in [6] and [7]. For simplicity, we assume that T is a non-fuzzy set.

** It should be noted that $\mu_N(x,y)$ may not be defined for all $x \in T$ and $y \in U$. For example, the degree to which the term jealous applies to an inanimate object such as chair may be assumed to be undefined rather than zero.

*** A fuzzy relation, R , from a set X to a set Y is a fuzzy subset of the cartesian product $X \times Y$. ($X \times Y$ is the collection of ordered pairs (x,y) , $x \in X$, $y \in Y$.) R is characterized by a bivariate membership function $\mu_R(x,y)$. For example, if $X = \{\text{Tom, Dick}\}$ and $Y = \{\text{John, Jim}\}$, then the fuzzy relation of resemblance between members of X and Y may be defined by a membership function $\mu_R(x,y)$ whose values might be $\mu_R(\text{Tom, John}) = 0.8$, $\mu_R(\text{Tom, Jim}) = 0.6$, $\mu_R(\text{Dick, John}) = 0.2$ and $\mu_R(\text{Dick, Jim}) = 0.9$. For additional discussion of fuzzy relations see [3], [8] and [9].

**** It will be understood throughout that the meaning of a term depends - to a greater or lesser extent - on the context in which it is used.

of the term red. Equivalently, the term red may be viewed as a label for a fuzzy subset of U which "comprises" (in a fuzzy sense) those elements of U whose color is red.

In short, the meaning, $M(x)$, of a term x is a fuzzy subset of U labeled x and characterized by the membership function

$$\mu_{M(x)}(y) = \mu_N(x,y), \quad x \in T, \quad y \in U \quad (2)$$

where $\mu_N(x,y)$ is the membership function of the naming relation N. To illustrate, suppose that U is the set of integers from 0 to 100 representing ages. Then, the meaning of the term young may be specified to be a fuzzy subset $M(\text{young})$ of U whose membership function is expressed by*

$$\begin{aligned} \mu_{\text{young}}(y) &= 1 \quad \text{for } y \leq 25 \\ &= \left(1 + \left(\frac{y - 25}{5} \right)^2 \right)^{-1} \quad \text{for } y > 25 \end{aligned} \quad (3)$$

and similarly

$$\begin{aligned} \mu_{\text{old}}(y) &= 0 \quad \text{for } y < 50 \\ &= \left(1 + \left(\frac{y - 50}{5} \right)^2 \right)^{-1} \quad \text{for } y \geq 50 \end{aligned} \quad (4)$$

The plots of the membership functions of $M(\text{young})$ and $M(\text{old})$ are shown in Fig. 1.

*The definitions of young and old as expressed by (3) and (4) should be viewed merely as illustrations of (2) rather than as accurate representations of the consensus regarding the meaning of these terms.

A crossover point in a fuzzy set is a point whose grade of membership in the set is 0.5. For example, the crossover point in M(young) is $y = 30$, while in M(old) it is $y = 55$.

The support of a fuzzy set A in U is defined to be the set of points in U for which $\mu_A(y)$ is positive. When the support of A is a finite set, it is frequently convenient to represent A as a linear combination of the elements of the support of A. Thus, if y_i , $i = 1, \dots, n$, is an element of the support of A and μ_i is its grade of membership in A, then A will be expressed as

$$A = \mu_1 y_1 + \mu_2 y_2 + \dots + \mu_n y_n \quad (5)$$

or more compactly

$$A = \sum_{i=1}^n \mu_i y_i \quad (6)$$

In cases where it is necessary to use a separator symbol to differentiate between μ_i and y_i , A will be written as

$$A = \mu_1 / y_1 + \mu_2 / y_2 + \dots + \mu_n / y_n \quad (7)$$

where the separator / serves to identify the μ_i and y_i components of the string $\mu_i y_i$.

A fuzzy singleton is a fuzzy set whose support is a single point in U. If S is a fuzzy singleton, we write

$$S = \mu / y \quad (8)$$

where μ is the grade of membership of y in S . A singleton in the usual (non-fuzzy) sense will be denoted by $1/y$.

The representation of a fuzzy set A in the form of Eq. (7) may be viewed as a decomposition of A into its constituent fuzzy singletons. The plus sign in this representation plays the role of the union (see (38)) of these singletons.

In most instances, the distinction between a term x and its meaning $M(x)$ is implied by the context. Consequently, our notation can be simplified without incurring a significant risk of confusion by writing x for $M(x)$. For example, if U is the set of ages from 0 to 100, i.e.,

$$U = \sum_{i=0}^{100} i \quad (9)$$

then the fuzzy subset* of U corresponding to the term middle-aged may be expressed as

$$\begin{aligned} \text{middle-aged} &= 0.3/40 + 0.5/41 + 0.8/42 + 0.9/43 + 1/44 + 1/45 \\ &+ 1/46 + 1/47 + 1/48 + 0.9/49 + 0.8/50 + 0.7/51 + 0.6/52 + 0.5/53 \\ &+ 0.4/54 + 0.3/55 \end{aligned} \quad (10)$$

with the understanding that the left-hand member of (10) stands for the fuzzy subset $M(\text{middle-aged})$.

When U is a countable set, e.g., the set of positive integers, we write

* A is a subset of B , written as $A \subset B$, if and only if $\mu_A(y) \leq \mu_B(y)$ for all y in U . For example, the fuzzy set $A = 0.8/3 + 0.6/4$ is a subset of $B = 0.9/3 + 0.7/4 + 0.3/5$.

$$U = \sum_{i=1}^{\infty} y_i \quad (11)$$

and

$$A = \sum_{i=1}^{\infty} \mu_i / y_i \quad (12)$$

For example, if U is the set of positive integers, and A is a fuzzy subset of U labeled small integer, we can represent A as

$$\begin{aligned} \underline{\text{small integer}} &= \sum_{i=1}^{\infty} \left(1 + \left(\frac{i-1}{10} \right)^2 \right)^{-1} / i \quad (13) \\ &= 1/1 + 0.99/2 + 0.96/3 + \dots \end{aligned}$$

When the universe of discourse, U, is a continuum, it is convenient to represent U as an "integral"

$$U = \int_U 1/y \quad (14)$$

with a fuzzy subset, A, of U represented as

$$A = \int_U \mu_A(y) / y \quad (15)$$

where μ_A is the membership function of A. It should be emphasized that in (14) and (15) the integral sign is not used in its conventional

sense. Rather, (14) and (15) are merely continuous counterparts of (11) and (12).*

As an illustration of the above notation, (3) can be rewritten more simply as

$$\underline{\text{young}} = \int_0^{25} 1/y + \int_{25}^{100} \left(1 + \left(\frac{y-25}{5}\right)^2\right)^{-1} /y \quad (16)$$

and similarly

$$\underline{\text{old}} = \int_{50}^{100} \left(1 + \left(\frac{y-50}{5}\right)^2\right)^{-1} /y \quad (17)$$

Note If y is an ordered n -tuple with components v_1, \dots, v_n , i.e., $y = (v_1, \dots, v_n)$, $v_i \in V_i \stackrel{\Delta}{=} \text{domain}^{**}$ of v_i , then (15) may be written as

$$A = \int_U \mu_A(v_1, \dots, v_n) / (v_1, \dots, v_n) \quad (18)$$

Using this notation, the relation of resemblance cited earlier (see footnote), may be expressed as

$$R = 0.8/(\text{Tom,John}) + 0.6/(\text{Tom,Jim}) + 0.2/(\text{Dick,John}) + 0.9/(\text{Dick,Jim}) \quad (19)$$

* Like (7), (14) represents the set-theoretic union (see (38)) of an indexed collection of fuzzy singletons $\mu_A(y)/y$. In this sense, the integral representation of a fuzzy set remains valid when U is a countable set rather than a continuum.

** The symbol $\stackrel{\Delta}{=}$ stands for "equal by definition" or "is defined to be" or "denotes."

It should be noted that the grade of membership in a fuzzy set may in itself be a fuzzy set. For example, suppose that the universe of discourse comprises persons named JOHN, TOM, DICK and HARRY, i.e.

$$U = \text{JOHN} + \text{TOM} + \text{DICK} + \text{HARRY} \quad (20)$$

and that I is the fuzzy subset of intelligent men in U. Furthermore, suppose that there are three fuzzy grades of membership labeled high, medium and low, which are defined as fuzzy subsets of the universe V,

$$V = 0 + 0.1 + 0.2 + 0.3 + 0.4 + 0.5 + 0.6 + 0.7 + 0.8 + 0.9 + 1 \quad (21)$$

Thus,

$$\underline{\text{high}} = 0.5/0.7 + 0.7/0.8 + 0.9/0.9 + 1/1 \quad (22)$$

$$\underline{\text{medium}} = 0.5/0.4 + 0.7/0.5 + 1/0.6 + 0.7/0.7 + 0.5/0.8 \quad (23)$$

$$\underline{\text{low}} = 0.5/0.2 + 0.7/0.3 + 1/0.4 + 0.7/0.5 + 0.5/0.6 \quad (24)$$

Then, we may have

$$I \stackrel{\Delta}{=} \underline{\text{intelligent}} = \underline{\text{medium}}/\text{JOHN} + \underline{\text{high}}/\text{TOM} + \underline{\text{low}}/\text{DICK} + \underline{\text{low}}/\text{HARRY} \quad (25)$$

In this way, the meaning of a term can be expressed as a fuzzy subset of the universe of discourse, with the grades of membership being numbers in the interval [0,1] or fuzzy subsets of this interval.

This concludes our brief summary of those aspects of the notation and terminology relating to fuzzy sets which we shall need in the following sections.

3. Operations on Fuzzy Sets

As defined in Section 2, if x_1, x_2, \dots, x_n are atomic terms, then their concatenation

$$x = x_1 x_2 \dots x_n \quad (26)$$

is a composite term. For example, if $x_1 = \text{very}$, $x_2 = \text{tall}$ and $x_3 = \text{man}$, then $x_1 x_2 x_3$ is the composite term very tall man.

In quantitative fuzzy semantics [6], the meaning of term x is a fuzzy subset, $M(x)$, of the universe of discourse (see (2)). From this point of view, one of the basic problems in semantics is that of devising an algorithm for the computation of the meaning of a composite term $x = x_1 \dots x_n$ from the knowledge of the meaning of each of its atomic components, x_i , $i = 1, \dots, n$.

In the present paper, our main concern is with a special case of this problem in which x is of the form $x = hu$, where h is a hedge, e.g., $h = \text{highly}$, and u is a term, e.g., $u = \text{intelligent man}$.

The point of view developed in this paper is that a hedge, h , may be interpreted as an operator, with operand u , which transforms a fuzzy subset $M(u)$ of U into the subset $M(hu)$. To characterize this operator, it is convenient to define several primitive operations on fuzzy sets from which more complicated operators such as hedges may be built up by composition.

We shall begin with the basic set-theoretic operations of complementation, intersection and union, and follow these with several more specialized operations: product, normalization, concentration, dilation,

contrast intensification, convex combination, and fuzzification.

Complementation

Complementation is a unary operation in the sense that it transforms a fuzzy set in U into another fuzzy set in U. More specifically, the complement of a fuzzy set A is denoted by $\neg A$ and is defined by the relation^{*}

$$\mu_{\neg A}(y) = 1 - \mu_A(y) \quad , \quad y \in U \quad (27)$$

Thus, if A is the class of rich men and $\mu_A(\text{John}) = 0.8$, then $\mu_{\neg A}(\text{John}) = 0.2$. Equivalently, if u is a term whose meaning is M(u), then the meaning of M(not u) is given by^{**}

$$M(\text{not } u) = \neg M(u) \quad (28)$$

In short, not(negation) is an operator^{***} which transforms M(u) into $\neg M(u)$.

Intersection

Intersection is a binary operation in the sense that it transforms

* Note that if $\mu_A(y)$ is undefined at y, then the same is true of $\mu_{\neg A}(y)$.

** It should be noted that, in a natural language, the negation not frequently corresponds to a relative complement, that is, to a complement relative to a subset, V, of the universe of discourse, with V implicitly defined by the operand of not and the context in which it appears. For example, in the sentence "Fifi is not a poodle," it might be understood that Fifi is a dog other than a poodle rather than any object in the universe of discourse which is not a poodle.

*** Strictly speaking, \neg acts on a fuzzy set whereas not (negation) acts on its label. Thus, when we write $\neg u$ it should be understood that u stands for M(u).

a pair of fuzzy sets in U into a fuzzy set in U. More specifically, the intersection of two fuzzy sets A and B is a fuzzy set denoted by $A \cap B$ and defined by

$$\mu_{A \cap B}(y) = \mu_A(y) \wedge \mu_B(y), \quad y \in U \quad (29)$$

where, for any real a and b, $a \wedge b$ denotes $\text{Min}(a,b)$, that is,

$$\begin{aligned} a \wedge b &= a && \text{if } a \leq b \\ &= b && \text{if } a > b \end{aligned} \quad (30)$$

To a first approximation,^{*} the conjunctive connective and may be identified with the intersection of fuzzy sets. Thus, if u and v are terms in T, then the meaning of the composite term $x = u \text{ and } v$ is given by^{**}

$$M(u \text{ and } v) = M(u) \cap M(v). \quad (31)$$

For example, if the universe of discourse is the set $U = 1 + 2 + 3 + 4 + 5 + 6$ and the meanings of u and v are expressed in the notation of (12) as

$$u = 0.8/3 + 1/5 + 0.6/6 \quad (32)$$

^{*}This qualification reflects the fact that in a natural language the meaning of and is somewhat context-dependent and is not always expressed by (31).

^{**}It should be noted that, from an algebraic point of view, (31) may be regarded as the definition of a homomorphism from the set of labels of fuzzy subsets of U to the set of fuzzy subsets of U, with the corresponding operations being and and \cap .

and

$$v = 0.7/3 + 1/4 + 0.5/6 \quad (33)$$

then

$$u \text{ and } v = 0.7/3 + 0.5/6 \quad (34)$$

More generally, if

$$u = \int_U \mu_A(y)/y \quad (35)$$

and

$$v = \int_U \mu_B(y)/y \quad (36)$$

then

$$u \text{ and } v = \int_U (\mu_A(y) \wedge \mu_B(y))/y \quad (37)$$

Union

Like the intersection, the union of fuzzy sets is a binary operation. More concretely, the union of two fuzzy sets A and B is a fuzzy set denoted $A \cup B$ - or, more conveniently, $A + B$ - and defined by

$$\mu_{A+B}(y) = \mu_A(y) \vee \mu_B(y), \quad y \in U \quad (38)$$

where $a \vee b$ denotes $\text{Max}(a,b)$, that is

$$\begin{aligned} a \vee b &= a && \text{if } a \geq b \\ &= b && \text{if } a < b \end{aligned} \quad (39)$$

Dual to the correspondence between the conjunctive connective and and \cap is the correspondence between the disjunctive connective or and $+$.

Thus

$$M(\underline{\text{or}}\ v) = M(u) + M(v) \quad (40)$$

or equivalently

$$\underline{\text{or}}\ v = \int_U (\mu_A(y) \vee \mu_B(y)) / y \quad (41)$$

where u and v are defined by (35) and (36).

As an illustration, for the terms defined by (32) and (33), (41) yields

$$\underline{\text{or}}\ v = 0.8/3 + 1/4 + 1/5 + 0.6/6 \quad (42)$$

It can readily be shown [3] that the union and intersection of fuzzy sets are associative as well as distributive operations. Furthermore, they satisfy the De Morgan identity

$$\neg (A \cap B) = \neg A + \neg B \quad (43)$$

which in terms of and, or and not may be stated as

$$\underline{\text{not}}\ (\underline{\text{u and v}}) = \underline{\text{not}}\ \underline{\text{u}}\ \underline{\text{or}}\ \underline{\text{not}}\ \underline{\text{v}} \quad (44)$$

Since $\neg\neg A = A$, (44) entails

$$\underline{\text{not}}\ (\underline{\text{u or v}}) = \underline{\text{not}}\ \underline{\text{u}}\ \underline{\text{and}}\ \underline{\text{not}}\ \underline{\text{v}} \quad (45)$$

Product

The product of two fuzzy sets A and B is denoted by AB and is defined by

$$\mu_{AB}(y) = \mu_A(y) \mu_B(y), \quad y \in U \quad (46)$$

Thus, if

$$A = 0.8/2 + 0.9/5 \quad (47)$$

and

$$B = 0.6/2 + 0.8/3 + 0.6/5 \quad (48)$$

then

$$AB = 0.48/2 + 0.54/5 \quad (49)$$

More generally, using the integral representation of A and B, we can write

$$AB = \int_U \mu_A(y) \mu_B(y) / y \quad (50)$$

For example, if

$$A = \int_0^{\infty} (1 + y^2)^{-1} / y \quad (51)$$

and

$$B = \int_0^{\infty} (1 + y^{-2})^{-1} / y \quad (52)$$

then

$$AB = \int_0^{\infty} y^2(1 + y^2)^{-2}/y \quad (53)$$

An immediate extension of (50) leads to the following definition of the expression A^α , where α is any real number:

$$A^\alpha = \int_U (\mu_A(y))^\alpha / y \quad (54)$$

For example, if $\alpha = 2$ and A is expressed by (47), then

$$A^2 = 0.64/2 + 0.81/5 \quad (55)$$

Similarly, if α is a non-negative real number, then the expression αA is defined by

$$\alpha A = \int_U \alpha \mu_A(y) / y \quad (56)$$

For example, if A is expressed by (47) and $\alpha = 0.5$, then

$$0.5 A = 0.4/2 + 0.45/5 \quad (57)$$

A useful property of the product is that it distributes over the union and the intersection.

Thus,

$$A(B + C) = AB + AC \quad (58)$$

and

$$A(B \cap C) = AB \cap AC \quad (59)$$

Comment It should be noted that, unlike the union and intersection, the product does not correspond to a commonly used connective. However, in some contexts the meaning of and may be more closely approximated by the product than by conjunction.

Comment The product of A and B as defined above differs from the cartesian (or direct product) of A and B. Thus, if A is a fuzzy subset of a universe of discourse U and B is a fuzzy subset of a possibly different universe of discourse V, then (see (18))

$$AxB = \int_{UxV} \mu_A(u) \wedge \mu_B(v)/(u,v) \quad (60)$$

where UxV denotes the cartesian product of the non-fuzzy sets U and V, and $u \in U, v \in V$. Note that when A and B are non-fuzzy, (60) reduces to the conventional definition of the cartesian product of non-fuzzy sets.

The need for differentiation between the product in the sense of (50) and the cartesian product becomes particularly important when $A = B$, since A^2 is commonly used to denote AxA (when A is non-fuzzy), whereas (54) implies that $A^2 = AA$. In what follows, we shall adhere to the latter interpretation except where an explicit statement to the contrary is made.

Normalization

Let $\bar{\mu}_A$ denote the supremum of a membership function μ_A over the

universe of discourse, i.e.,

$$\bar{\mu}_A = \text{Sup}_U \mu_A(y) \quad (61)$$

A fuzzy set A is said to be normal if $\bar{\mu}_A = 1$; otherwise, A is sub-normal. For example, the set

$$A = 1/\text{JOHN} + 0.8/\text{JIM} + 0.6/\text{TOM} \quad (62)$$

is normal, while

$$A = 0.6/\text{JOHN} + 0.8/\text{JIM} + 0.6/\text{TOM} \quad (63)$$

is subnormal.

A subnormal fuzzy set A can be normalized by dividing μ_A by $\bar{\mu}_A$. Using the notation of (56), the operation of normalization may be expressed as

$$\text{NORM}(A) = (\bar{\mu}_A)^{-1} A, \quad \bar{\mu}_A \neq 0 \quad (64)$$

Thus, for the fuzzy set defined by (63), the normalization of A results in

$$\text{NORM}(A) = 0.75/\text{JOHN} + 1/\text{JIM} + 0.75/\text{TOM} \quad (65)$$

Concentration

Like complementation, concentration is a unary operation. As its name implies, the result of applying a concentrator to a fuzzy set A is a fuzzy subset of A such that the reduction in the magnitude of the grade of membership of y in A is relatively small for those y which

have a high grade of membership in A and relatively large for the y with low membership. Thus, if we denote the result of applying a concentrator to A by CON(A), then the relation between the membership function of A and that of CON(A) will typically have the appearance shown in Fig. 2.

To be more specific, we shall assume that the operation of concentration has the effect of squaring the membership function of A. Thus,

$$\mu_{\text{CON}(A)}(y) = \mu_A^2(y), \quad y \in U \quad (66)$$

or, using the definition of A^2 (see (54))

$$\text{CON}(A) = A^2 \quad (67)$$

For example, if the meaning of the term few is defined by

$$\underline{\text{few}} = 1/1 + 1/2 + 0.8/3 + 0.6/4 \quad (68)$$

then

$$\text{CON}(\underline{\text{few}}) = 1/1 + 1/2 + 0.64/3 + 0.36/4 \quad (69)$$

It should be noted that concentration distributes over the union, intersection and product. Thus

$$\text{CON}(A + B) = \text{CON}(A) + \text{CON}(B) \quad (70)$$

$$\text{CON}(A \cap B) = \text{CON}(A) \cap \text{CON}(B) \quad (71)$$

and

$$\text{CON}(AB) = \text{CON}(A) \text{CON}(B) \quad (72)$$

Equation (70) follows from the identities

$$(A + B)^2 = A^2 + AB + BA + B^2 \quad (73)$$

and

$$AB + BA \subset A^2 + B^2 \quad (74)$$

which together imply that

$$(A + B)^2 = A^2 + B^2 \quad (75)$$

Equation (71) follows similarly, with + replaced by \cap .

The operation of concentration can be composed with itself. Thus

$$\text{CON}^2(A) = A^4 \quad (76)$$

and more generally

$$\text{CON}^\alpha(A) = A^{2\alpha} \quad (77)$$

where α is any integer ≥ 2 .

Dilation

The effect of dilation is the opposite of that of concentration. Thus, the result of applying a dilator to a fuzzy set A is a fuzzy set DIL(A) whose membership function is related to that of A as shown in Fig. 3.

More specifically, DIL(A) is defined by

$$\text{DIL}(A) = A^{0.5} \quad (78)$$

which implies that

$$\mu_{\text{DIL}(A)}(y) = \sqrt{\mu_A(y)}, \quad y \in U \quad (79)$$

For example, if few is defined by (68), then

$$\text{DIL}(\text{few}) = 1/1 + 1/2 + 0.9/3 + 0.78/4$$

Contrast intensification

The operation of concentration has the effect of diminishing the value of $\mu_A(y)$ for every y (except where $\mu_A(y) = 1$), with the larger values of $\mu_A(y)$ diminished proportionately less than the smaller values.

The operation of contrast intensification, or simply intensification, differs from that of concentration in that it increases the values of $\mu_A(y)$ which are above 0.5 and diminishes those which are below this threshold. Thus, if the result of applying a contrast intensifier INT to a fuzzy set A is denoted by INT(A), we have

$$\mu_{\text{INT}(A)}(y) \geq \mu_A(y) \quad \text{for} \quad \mu_A(y) \geq 0.5 \quad (80)$$

and

$$\mu_{\text{INT}(A)}(y) \leq \mu_A(y) \quad \text{for} \quad \mu_A(y) \leq 0.5 \quad (81)$$

A simple concrete expression for an operator of this type is the following

$$\mu_{\text{INT}(A)}(y) = 2 \mu_A^2(y) \quad \text{for} \quad 0 \leq \mu_A(y) \leq 0.5 \quad (82)$$

$$\mu_{\text{INT}(A)}(y) = 1 - 2(1 - \mu_A(y))^2 \quad \text{for} \quad 0.5 \leq \mu_A(y) \leq 1$$

The effect of applying this intensifier to a fuzzy set A is shown in Fig. 4.

As in the case of concentration, intensification distributes over the union, intersection and product. Thus

$$\text{INT}(A+B) = \text{INT}(A) + \text{INT}(B) \quad (83)$$

$$\text{INT}(A \cap B) = \text{INT}(A) \cap \text{INT}(B) \quad (84)$$

and

$$\text{INT}(AB) = \text{INT}(A)\text{INT}(B) \quad (85)$$

Note The function defined by (82) is of use also in the representation of membership functions of fuzzy sets (see (170)). For this purpose, it is convenient to define a function S from the real line to [0,1] by the equations

$$\begin{aligned} S(u) &\triangleq 0 && \text{for } u < 0 \\ &\triangleq 2u^2 && \text{for } 0 \leq u \leq 0.5 \\ &\triangleq 1-2(1-u)^2 && \text{for } 0.5 < u \leq 1 \\ &\triangleq 1 && \text{for } u > 1 \end{aligned} \quad (86)$$

This function will be referred to as S-function to stress its resemblance to an S (see Fig. 5).

Putting two S-functions back to back we obtain a pulse-function, $\pi(u)$, which has the appearance shown in Fig. 6. The defining equations for $\pi(u)$ are:

$$\begin{aligned}
\pi(u) &\stackrel{\Delta}{=} 0 && \text{for } u < -1 \\
&\stackrel{\Delta}{=} 2(u+1)^2 && \text{for } -1 \leq u < -0.5 \\
&\stackrel{\Delta}{=} 1-2u^2 && \text{for } -0.5 \leq u \leq 0.5 \\
&\stackrel{\Delta}{=} 2(u-1)^2 && \text{for } 0.5 < u \leq 1 \\
&\stackrel{\Delta}{=} 0 && \text{for } u > 1
\end{aligned} \tag{87}$$

Convex combination

The convex combination [3] is an n-ary operation which combines a set of n fuzzy sets A_1, \dots, A_n into a single fuzzy set A.

The fuzzy set A is a weighted combination of A_1, \dots, A_n in the sense that the membership function of A is related to those of A_1, \dots, A_n by the expression

$$\mu_A(y) = w_1(y)\mu_{A_1}(y) + \dots + w_n(y)\mu_{A_n}(y) \tag{88}$$

where the weights $w_1(y), \dots, w_n(y), 0 \leq w_i(y) \leq 1, i = 1, \dots, n,$ are such that

$$w_1(y) + \dots + w_n(y) = 1 \quad \text{for all } y \text{ in } U. \tag{89}$$

For example, for the sets expressed by (32) and (33), the convex combination of u and v with constant weights $w_1 = 0.8$ and $w_2 = 0.2$ is given by

$$\begin{aligned}
A = &(0.8 \times 0.8 + 0.2 \times 0.7)/3 + 0.8/5 + 0.2/4 + \\
&(0.8 \times 0.6 + 0.2 \times 0.5)/6
\end{aligned} \tag{90}$$

or

$$A = 0.78/3 + 0.2/4 + 0.8/5 + 0.58/6$$

Fuzzification

The operation of contrast intensification has the effect of transforming a fuzzy set A into a fuzzy set A^* which approximates to \sim and is less fuzzy than $\sim A$. As its name implies, the operation of fuzzification has the opposite effect. Thus, its main function is to provide a means of transforming a fuzzy (or non-fuzzy) set A into an approximating set \underline{A} which is more fuzzy than A .

The wavy bar \sim plays the role of a fuzzifier. Thus, if U is the set of real numbers, then $\underline{3}$ represents the fuzzy set of real numbers which are approximately equal to 3. Similarly, if $=$ denotes the relation of equality, then $\underline{=}$ represents approximate equality; and if $>$ denotes greater than, then $\underline{>}$ might be interpreted as more or less greater than.

There are many ways in which fuzzification can be accomplished. In what follows, we shall sketch two approaches which are particularly relevant to the definition of such linguistic hedges as more or less.

At the base of these approaches is a process of point fuzzification which transforms a singleton set $1/u$ in U into a fuzzy set \underline{u} which is concentrated around u . To place in evidence the dependence of \underline{u} on u , \underline{u} will be written as $\underline{u} = K(u)$. Unless stated to the contrary, the grade of membership of u in $K(u)$ will be assumed equal to 1.

The fuzzy set $K(u)$ will be referred to as the kernel of the fuzzification. Usually, $K(u)$ will be taken to be a fuzzy interval,^{*} that is,

* If U is an Euclidean n -space, then a fuzzy interval in U is a convex fuzzy subset of U . (Convex fuzzy sets are defined in [3].)

a fuzzy set whose membership function $\mu_{K(u)}(y)$ is a non-increasing function of the distance between u and y . For example, suppose that the universe of discourse is defined by

$$U = 1 + 2 + 3 + 4 \quad (91)$$

and the singleton set $1/2$ is transformed into the fuzzy set $0.6/1 + 1/2 + 0.8/3 + 0.3/4$, i.e.,

$$1/2 \rightarrow 0.6/1 + 1/2 + 0.8/3 + 0.3/4 \quad (92)$$

Then, in this case

$$K(2) = 0.6/1 + 1/2 + 0.8/3 + 0.3/4 \quad (93)$$

Now consider a fuzzy set represented by

$$A = \mu_1/y_1 + \dots + \mu_n/y_n \quad (94)$$

where μ_i is the grade of membership of y_i in A . If we postulate that fuzzification is a linear transformation, then (94) implies that

$$\tilde{A} = \underbrace{\mu_1/y_1} + \dots + \underbrace{\mu_n/y_n} \quad (95)$$

At this point, it is natural to consider two special cases. In Case I, we hold μ_i constant and set

$$\underbrace{\mu_i/y_i} = \underbrace{\mu_i/y_i}, \quad i = 1, \dots, n \quad (96)$$

On the other hand, in Case II, we hold y_i constant and set

$$\mu_{\tilde{1}}/y_{\tilde{1}} = \mu_{\tilde{1}}/y_{\tilde{1}} \quad (97)$$

Thus, in Case I we fuzzify each point in the support of A, while in Case II we fuzzify the grade of membership.

In Case I, suppose that the transformation which takes $1/y_{\tilde{1}}$ into $y_{\tilde{1}}$ is characterized by the kernel $K(y_{\tilde{1}})$. Then the fuzzification which takes A into \tilde{A} is denoted by $SF(A;K)$ (SF standing for support fuzzification or s-fuzzification, for short) and is defined by

$$\tilde{A} \triangleq SF(A;K) = \mu_{\tilde{1}}K(y_{\tilde{1}}) + \dots + \mu_{\tilde{n}}K(y_{\tilde{n}}) \quad (98)$$

where $\mu_{\tilde{1}}K(y_{\tilde{1}})$ should be interpreted as a fuzzy set which is the product of a scalar constant $\mu_{\tilde{1}}$ and a fuzzy set $K(y_{\tilde{1}})$ (see (56)), and + stands for the union of fuzzy sets.

As a simple illustration, suppose that

$$U = 1 + 2 + 3 + 4 \quad (99)$$

$$A = 0.8/1 + 0.6/2 \quad (100)$$

$$K(1) = 1/1 + 0.4/2 \quad (101)$$

and

$$K(2) = 1/2 + 0.4/1 + 0.4/3 \quad (102)$$

Then

$$SF(A;K) = 0.8(1/1 + 0.4/2) + 0.6(1/2 + 0.4/1 + 0.4/3) \quad (103)$$

$$= 0.8/1 + 0.32/2 + 0.6/2 + 0.24/1 + 0.24/3$$

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

Comment Note that the effect of s-fuzzification is somewhat similar to but not quite the same as that of dilation. One important difference is that a dilation of a non-fuzzy set yields the same non-fuzzy set, whereas an s-fuzzification of a non-fuzzy set will, in general, yield a fuzzy set. However, in the case of the fuzzy set defined by

$$A = 1/1 + 0.8/2 + 0.6/3 \quad (104)$$

we have (see (78))

$$DIL(A) = 1/1 + 0.9/2 + 0.78/3 \quad (105)$$

and the equality

$$DIL(A) = SF(A;K) \quad (106)$$

can be realized with the kernel set

$$K(1) = 1/1 + 0.9/2 \quad (107)$$

$$K(2) = 1/2 + 0.87/3$$

Equation (98) defines the effect of s-fuzzification when A has a finite support. More generally, assume that A is defined by (see (15))

$$A = \int_U \mu_A(y)/y \quad (108)$$

Then, SF(A;K) is given by

$$SF(A;K) = \int_U \mu_A(y)K(y), \quad y \in U \quad (109)$$

where, as in (98), $\mu_A(y)K(y)$ represents the product of the scalar constant $\mu_A(y)$, and \int_U should be interpreted as the union of the family of fuzzy sets $\mu_A(y)K(y)$, $y \in U$.

Comment It is of interest to observe that (109) is analogous to the integral representation of a linear operator. Note that if A is a singleton set $1/y$, then

$$SF(A;K) = K(y) \quad (110)$$

Thus, a singleton set is analogous to a delta-function and $K(y)$ plays the role of impulse response.

As a simple illustration of (109), assume that A is a non-fuzzy set whose membership function is shown in Fig. 7a, and K is a fuzzy set whose membership function is depicted in Fig. 7b. Then, $SF(A;K)$ is a fuzzy set whose membership function has the form shown in Fig. 7c.

It should be noted that s-fuzzification can be employed in an indirect manner to effect a translation of a fuzzy or non-fuzzy set within its universe. As an illustration, suppose that we wish to transform the non-fuzzy set A whose membership function is shown in Fig. 8a into the non-fuzzy set B whose membership function is a translate of μ_A (see Fig. 8b), i.e.,

$$\mu_B(y) = \mu_A(y-a) \quad (111)$$

Let $K(y)$ be the interval $[y-a, y+a]$. Then, we can express B as

$$B = \neg SF(\neg A;K) \quad (112)$$

where \neg is the operation of complementation (see (27)).

Turning to Case II, suppose that the point fuzzification which transforms μ_i into $\underline{\mu}_i$ is defined by the kernel $K(\mu_i)$. Then, denoting by $GF(A;K)$ (GF standing for grade fuzzification or g-fuzzification, for short) the fuzzification which transforms A into \underline{A} , where

$$\underline{A} = \underline{\mu}_1/y_1 + \dots + \underline{\mu}_n/y_n \quad (113)$$

we can write

$$GF(A;K) \triangleq \underline{A} = K(\mu_1)/y_1 + \dots + K(\mu_n)/y_n. \quad (114)$$

For example, if

$$U = 1 + 2 + 3 + 4 \quad (115)$$

and

$$A = \underline{0.8}/1 + \underline{0.6}/2 \quad (116)$$

then

$$GF(A;K) = \underline{0.8}/1 + \underline{0.6}/2 \quad (117)$$

where 0.8 and 0.6 are fuzzy grades of membership defined as, say,

$$K(\underline{0.8}) \triangleq \underline{0.8} = 1/0.8 + 0.6/0.7 + 0.6/0.9 \quad (118)$$

$$K(\underline{0.6}) \triangleq \underline{0.6} = 1/0.6 + 0.6/0.5 + 0.6/0.7$$

More generally, if

$$A = \int_U \mu_A(y)/y \quad (119)$$

then a g-fuzzification of A yields the fuzzy set

$$GF(A;K) = \int_U \underbrace{\mu_A(y)}_{\mu_A(y)/y} /y \quad (120)$$

in which the grade of membership of y in GF(A;K) is the fuzzy set

$$\underbrace{\mu_A(y)}_{\mu_A(y)} = K(y).$$

The various operations on fuzzy sets which we have defined in this section - especially complementation, intersection, product, concentration, convex combination and fuzzification - provide a basis for the characterization of hedges in terms of compositions or combinations of these operations. We turn to this subject in the following section.

4. Representation of Hedges as Operators

Before we can analyze the representation of hedges as operators acting on fuzzy subsets of the universe of discourse, it will be necessary to consider a basic question in the theory of fuzzy sets which relates to the representation of a fuzzy set in terms of other fuzzy sets.

We have already encountered several special instances of such representations in the preceding section. More generally, if A_1, \dots, A_n are fuzzy subsets of U with membership functions $\mu_{A_1}, \dots, \mu_{A_n}$, respectively, then a fuzzy set A in U has A_1, \dots, A_n as its components if the membership function of A is expressible as some (nontrivial) function

of $\mu_{A_1}, \dots, \mu_{A_n}$. Thus, in symbols

$$\mu_A = f(\mu_{A_1}, \dots, \mu_{A_n}) \quad (121)$$

For example, if A is the intersection of A_1 and A_2 , then by (29)

$$\mu_A = \mu_{A_1} \wedge \mu_{A_2} \quad (122)$$

where for simplicity we have omitted the argument y.

Now if A represents the meaning of a fairly complex concept, then, in general, it is expedient to "resolve" A into a set of simpler components A_1, \dots, A_n , so that the membership function of A becomes a function of those of A_1, \dots, A_n . For example, if A is the class of big men, then, roughly, its components might be taken to be

$$A_1 = \text{class of } \underline{\text{tall men}} \quad (123)$$

$$A_2 = \text{class of } \underline{\text{heavy men}} \quad (124)$$

and in terms of these components A may be defined by the equation*

$$\mu_A(y) = 0.6 \mu_{A_1}(y) + 0.4 \mu_{A_2}(y) \quad (125)$$

For example, if

$$\mu_{A_1}(\text{JOHN}) = 0.8 \quad (126)$$

* This definition is used merely as an illustration and is not intended to be an accurate representation of the concept of big man.

and

$$\mu_{A_2}(\text{JOHN}) = 0.5 \quad (127)$$

then

$$\mu_A(\text{JOHN}) = 0.6 \times 0.8 + 0.4 \times 0.5 = 0.68 \quad (128)$$

Our motivation for considering the resolution of a fuzzy set into simpler components has to do with the fact that the representation of certain hedges such as basically, technically, literally, etc., involves in an essential way their effect on the components of the fuzzy sets on which they operate. On the other hand, the effect of simpler hedges such as very, more or less, slightly etc. can be described without resort to the resolution of a set into its components. This suggests that hedges be divided into two somewhat fuzzy categories, which may be defined informally as follows.

Type I. Hedges in this category can be represented as operators acting on a fuzzy set. Typical hedges in this category are: very, more or less, much, slightly, highly.

Type II. Hedges in this category require a description of how they act on the components of the operand. Typical hedges in this category are: essentially, technically, actually, strictly, in a sense, practically, virtually, regular, etc.

Each of the above categories includes a subcategory of hedges - denoted by IP and IIP, respectively - whose effect is influenced by the notion of proximity or ordering in the domain of the operand. Such

hedges - of which slightly is a simple example - are generally more context-dependent than those hedges of Types I and II which do not require the notion of proximity for their characterization.

In what follows, we shall examine in greater detail some of the basic aspects of the representation of hedges of Type I and II. As illustrations of our approach, we shall construct approximate operator representations for several typical hedges of Type I and II. It should be emphasized, however, that these representations are intended mainly to illustrate the approach rather than to provide accurate definitions of the hedges in question. Furthermore, it must be underscored that our analysis and its conclusions are tentative in nature and may require modification in later work.

5. Representation of Hedges of Type I

It will be convenient to begin our discussion by considering a relatively simple and yet very basic hedge, namely, very.

very

Let A be a fuzzy set in U representing the meaning of a term such as

$$x = \underline{\text{old men}} \tag{129}$$

We assume that A is characterized by a membership function of the form shown in Fig. 9.

Now consider the term

$$x^* = \underline{\text{very}} x = \underline{\text{very old men}} \tag{130}$$

and let A^* be the fuzzy set representing the meaning of x^* .

The crux of our idea is to view a hedge such as very as an operator which transforms the fuzzy set A (\triangleq meaning of x) into the fuzzy set A^* (\triangleq meaning of x^*). If we accept this point of view, then the question that arises is: How can the operator very be defined?

Given the richness and complexity of natural languages, it is clear that questions of this kind do not admit of simple and definitive answers. Nevertheless, it is useful to attempt to concretize the meaning of a hedge such as very even if the postulated meaning does not have universal validity and is merely a fixed approximation to a variety of shades of meaning which very can assume in different contexts. It is in this perspective that the definitions of very and other hedges which are formulated in the sequel should be viewed.

Specifically, we assume that if the meaning of a term x is a fuzzy set A , then the meaning of very x , A^* , is given by

$$A^* = \text{CON}(A) \quad (131)$$

or, more explicitly (see (67))

$$A^* = A^2 \quad (132)$$

Using the same symbol to denote a term and its meaning (as in (10)), the definition of very may be expressed more compactly as

$$\text{very } x = x^2 \quad (133)$$

* It should be observed that the operand of very must be a term with a fuzzy meaning, i.e., A must be a fuzzy set. Thus, strictly speaking, it is not correct to say very rectangular or very pregnant, since both rectangular and pregnant have non-fuzzy meaning.

Thus, if

$$x = \mu_1/y_1 + \dots + \mu_n/y_n, \quad y_i \in U, i = 1, \dots, n \quad (134)$$

then

$$\underline{\text{very}} x = \mu_1^2/y_1 + \dots + \mu_n^2/y_n \quad (135)$$

and, more generally, if

$$x = \int_U \mu(y)/y \quad (136)$$

then

$$\underline{\text{very}} x = \int_U \mu^2(y)/y \quad (137)$$

For example, if (see (17))

$$x = \underline{\text{old men}} = \int_{50}^{100} \left(1 + \left(\frac{y-50}{5}\right)^{-2}\right)^{-1} /y \quad (138)$$

then

$$\underline{\text{very old men}} = \int_{50}^{100} \left(1 + \left(\frac{y-50}{5}\right)^{-2}\right)^{-2} /y \quad (139)$$

which implies that if, say, the grade of membership of John in the class of old men is 0.8, then his grade of membership in the class of very old men is 0.64. (See Fig. 9 for illustration.)

Viewed as an operator, very can be composed with itself. Thus

$$\underline{\text{very very old men}} = (\underline{\text{old men}})^4 \quad (140)$$

Using the definitions of not and and, we can compute the meaning of such composite terms as $w = \underline{\text{not very old}}$, $z = \underline{\text{not young and not very old}}$, etc., in which the term men is implicit. Thus

$$\underline{\text{not very old}} = \neg \underline{\text{old}}^2 \quad (141)$$

and

$$\underline{\text{not young and not very old}} = \neg \underline{\text{young}} \cap \neg \underline{\text{old}}^2 \quad (142)$$

For example, suppose that the grade of membership of David in the class of old men is 0.6 and in the class of young men is 0.1. Then his grade of membership in the class of men who are not young and not very old can be computed from (142) to be

$$\begin{aligned} \mu_z(\text{David}) &= (1-0.1) \wedge (1-0.6^2) \\ &= 0.64 \end{aligned} \quad (143)$$

Comment It should be noted that, basically, very modifies the adverb or the adjective which follows it. Consequently, it is not grammatical to write

$$x = \underline{\text{very not exact}} \quad (144)$$

but if not exact is replaced by the single term inexact, then

$$x = \underline{\text{very inexact}} \quad (145)$$

becomes meaningful and we can assert that

$$\begin{aligned} \underline{\text{very inexact}} &= \underline{\text{inexact}}^2 & (146) \\ &= (\underline{\text{not exact}})^2 \end{aligned}$$

On the other hand

$$\begin{aligned} \underline{\text{not very exact}} &= \neg (\underline{\text{very exact}}) \\ &= \neg (\underline{\text{exact}}^2) & (147) \end{aligned}$$

which is different from (146).

Comment Under the definition of very expressed by (133), if the grade of membership of y ($y \in U$) in a fuzzy set labeled x is unity, then the same is true of very x , that is,

$$\mu_x(y) = 1 \quad \text{implies} \quad \mu_{\underline{\text{very } x}}(y) = 1 \quad (148)$$

For example, if the grade of membership of JOHN in the class of old men is 1, then the same is true of the grade of membership of JOHN in the class of very old men. Is this in accord with our intuition?

This basic question does not appear to have a clear-cut answer on purely intuitive grounds. It is easy to show, however, that (148) can be deduced as a consequence of the following two assumptions.

(a) very distributes over the union (e.g., very (tall or fat) = very tall or very fat)

(b) very $x = x$ if x is non-fuzzy. (E.g., very square = square)

Specifically, suppose that there is a non-empty set of points in U whose grade of membership in x is unity. Let this non-fuzzy set be denoted by x_{nf} . Then x can be represented as the union of two disjoint

sets: the non-fuzzy set x_{nf} and a possibly fuzzy set x_f whose support comprises those points in U whose grade of membership in x is less than unity. Thus

$$x = x_{nf} + x_f \quad (149)$$

Now, by assumption (a)

$$\begin{aligned} \underline{\text{very}} x &= \underline{\text{very}} (x_{nf} + x_f) \\ &= \underline{\text{very}} x_{nf} + \underline{\text{very}} x_f \end{aligned} \quad (150)$$

and by assumption (b)

$$\underline{\text{very}} x_{nf} = x_{nf} \quad (151)$$

Consequently

$$\underline{\text{very}} x = x_{nf} + \underline{\text{very}} x_f \quad (152)$$

Now, if the grade of membership of y_1 in x is unity, then by the definition of x_{nf} , y_1 belongs to x_{nf} . But (152) implies that the grade of membership of y_1 in $\underline{\text{very}} x$ is greater than or equal to the grade of membership of y_1 in x_{nf} . This establishes that the grade of membership of y_1 in $\underline{\text{very}} x$ is unity and hence demonstrates that (148) is a consequence of (a) and (b).

As a part of a system of hedges which could be used to characterize the behavior of complex systems, it is convenient to have a hedge whose effect is milder than that of very. To this end, it is helpful to

introduce two artificial hedges plus and minus which are defined below.

Plus and Minus

The artificial hedges plus and minus are, respectively, instances of what might be called accentuators and deaccentuators whose function is to increase the range of shades of meaning of various hedges by providing milder degrees of concentration and dilation than those associated with the operations CON and DIL (see (67) and (78)). Thus, as operators acting on a fuzzy set labeled x, plus and minus are defined as follows

$$\text{plus } x = x^{1.25} \quad (153)$$

and

$$\text{minus } x = x^{0.75} \quad (154)$$

The numerical values of the exponents in (153) and (154) are chosen in such a way as to entail the approximate identity*

$$\text{plus plus } x = \text{minus very } x \quad (155)$$

To illustrate the effects of plus and minus, plots of the membership functions of old man, plus old man, minus very old man and very old man are shown in Fig. 10.

The main use for the artificial hedges plus and minus is likely to lie in applications - e.g., the description of fuzzy algorithms - in which the standards of precision of meaning are higher than in ordinary

* In order that (155) hold precisely, plus x and minus x should be defined as plus x \triangleq x^{1+α} and minus x \triangleq x^{1-α}, where α = √5 - 2.

discourse. In addition, plus and minus can be used to define a natural hedge whose meaning differs slightly from that of some other natural hedge. For example, the hedge highly could be defined as

$$\underline{\text{highly}} = \underline{\text{plus}} \underline{\text{very}} \tag{156}$$

or possibly as

$$\underline{\text{highly}} = \underline{\text{minus}} \underline{\text{very}} \underline{\text{very}} \tag{157}$$

Thus, using (156) we would have the equality

$$\underline{\text{highly}} \underline{\text{intelligent}} = \underline{\text{plus}} \underline{\text{very}} \underline{\text{intelligent}} \tag{158}$$

whereas (157) would entail

$$\underline{\text{highly}} \underline{\text{intelligent}} = \underline{\text{minus}} \underline{\text{very}} \underline{\text{very}} \underline{\text{intelligent}} \tag{159}$$

Much

The hedge much is used mostly in conjunction with relations, e.g., much greater, much more beautiful, much less likely, etc.

Consider the effect of much on the relation greater than defined in the universe of positive real numbers, $U = (0, \infty)$

The meaning of the relation greater than or, more simply, $>$, is a non-fuzzy set in $U \times U$ which may be represented as

$$> = \int_D 1/(v,w) \tag{160}$$

where $v, w \in U$ and D is the set of all points in $U \times U$ at which $v > w$.

The hedge much acts as an s-fuzzifier (see (98)) in the sense that it transforms the non-fuzzy set $>$ into the fuzzy set $>>$ ($>> \triangleq$ much greater than). For simplicity, suppose that $>>$ is defined as

$$>> = \int_D (1 + (v-w)^{-2})^{-1} / (v,w) \quad (161)$$

Then $>>$ can be expressed as the result of acting with an s-fuzzifier on $>$, i.e.,

$$>> = \neg SF(\neg >; K) \quad (162)$$

in which the kernel K is defined by

$$K(v,w) = \int_{D^*} (1 + (r-v-s+w)^2)^{-1} / (r,s) \quad (163)$$

where $r, s \in U$, and $D^* = \{(r,s) | r - v \geq s - w\}$.

The hedge much can be composed with itself by using the rule of composition of fuzzy relations (see [8]). Thus, if we denote much much greater than by $>>>$, then $>>>$ may be expressed as

$$>>> = \int_D \vee_z ((1 + (v-z)^{-2})^{-1} \wedge (1 + (z-w)^{-2})^{-1}) / (v,w) \quad (164)$$

where \vee_z denotes the supremum over z of the parenthesized expression.

Observation It should be noted that the term greater than is sometimes used in a somewhat fuzzy sense rather than in the non-fuzzy sense of (160). For example, greater than may be defined as a fuzzy relation

expressed by

$$\underline{\text{greater than}} = \int_D (0.8 + 0.2 (1 + (\frac{v}{2w})^{-2})^{-1}) / (v,w) \quad (165)$$

In such cases, much greater than may be interpreted as CON(greater than) or, alternatively, as the composition of greater than with itself.

More or less

In such common uses of the hedge more or less as more or less intelligent, more or less rectangular, more or less sweet, more or less plays the role of a fuzzifier - usually an s-fuzzifier. When the operand of more or less is fuzzy, as in more or less intelligent, it may be possible to achieve the same effect by deaccentuation or dilation.

As an illustration, suppose that the term recent is defined by

$$\underline{\text{recent}} = 1/1972 + 0.8/1971 + 0.7/1970 \quad (166)$$

Then, we may define the composite term more or less recent by the expression

$$\underline{\text{more or less recent}} = SF(\underline{\text{recent}}; K) \quad (167)$$

in which the kernel set K is defined by (168)

$$K(1972) = 1/1972 + 0.9/1971$$

$$K(1971) = 1/1971 + 0.9/1970$$

$$K(1970) = 1/1970 + 0.8/1969$$

Thus, on substituting (168) in (167), we obtain

$$\begin{aligned} \underline{\text{more or less recent}} &= 1/1972 + 0.9/1971 + 0.72/1970 \\ &+ 0.56/1969 \end{aligned} \tag{169}$$

Note that (169) cannot be obtained from (166) by dilation because the singleton 1/1969 is absent in (166).

As another example, suppose that the term tall (applying to man) is defined in terms of the S-function (see (86)) as follows

$$\underline{\text{tall}} = \int_{62}^{\infty} S\left(\frac{h - 62}{12}\right)/h \tag{170}$$

where h denotes the height in inches. (Note that, as defined by (170), the membership function of tall, $\mu_{\underline{\text{tall}}}(h)$, is zero for $h < 62$; is equal to 0.5 at $h = 68$; and is equal to 1 for $h > 74$.)

Now, if we define more or less tall by

$$\underline{\text{more or less tall}} = \neg SF(\neg \underline{\text{tall}}; K) \tag{171}$$

where the kernel set $K(h)$ is the interval $[h - 2, h + 2]$, then (171) yields

$$\underline{\text{more or less tall}} = \int_{60}^{\infty} S\left(\frac{h - 60}{12}\right)/h \tag{172}$$

which is equivalent to a translation (see Fig. 11) of the membership function of tall, i.e.

$$\mu_{\text{more or less tall}}(h) = \mu_{\text{tall}}(h + 2), \quad h > 0 \quad (173)$$

Slightly

Basically, slightly is a hedge of Type IP, that is, its effect is dependent on the definition of proximity or ordering in the domain of its operand. There are cases, however, in which the meaning of slightly can be defined in terms of hedges of Type I, usually under the tacit assumption that the domain of the operand is a linearly ordered set.

In such cases, various shades of the meaning of slightly might be defined by the following expressions

$$\text{slightly } x = \text{NORM}(x \text{ and not very } x) \quad (174)$$

$$\text{slightly } x = \text{INT}(\text{NORM}(\text{plus } x \text{ and not very } x)) \quad (175)$$

$$\text{slightly } x = \text{INT}(\text{NORM}(\text{plus } x \text{ and not plus very } x)) \quad (176)$$

in which the operation of normalization (see (64)) is needed because the fuzzy sets within the parentheses are, in general, subnormal, and the function of INT (see (82)) is to intensify the contrast between these elements that have a high grade of membership in its operand and those whose grade of membership is low. Essentially, (174), (175), and (176) represent various degrees of approximation to an operator which transforms a fuzzy set x such as shown in Fig. 12a into a fuzzy set of the form shown in Fig. 12b.

As an illustration of the use of (174), suppose that x (e.g., $x = \text{tall men}$) is defined by

$$x = \int_0^{\infty} (1 + (\frac{y}{a})^{-2})^{-1} / y \quad (177)$$

where a is the crossover point of x , i.e., the value of y at which

$\mu_x(y) = 0.5$. Then

$$\text{very } x = \int_0^{\infty} (1 + (\frac{y}{a})^{-2})^{-2} / y \quad (178)$$

$$\text{not very } x = \int_0^{\infty} (1 - (1 + (\frac{y}{a})^{-2})^{-2}) / y \quad (179)$$

and

$$x \text{ and not very } x = \int_0^{\infty} (1 + (\frac{y}{a})^{-2})^{-1} \wedge (1 - (1 + (\frac{y}{a})^{-2})^{-2}) / y \quad (180)$$

The value of y at which the membership function of x and not very x attains its maximum value is easily computed to be

$$y_{\max} = a \sqrt{\frac{2}{\sqrt{5} - 1}} \approx 1.3 \quad (181)$$

and the value of the maximum is

$$\mu_{\max} = \frac{2}{\sqrt{5} + 1} \approx 0.6 \quad (182)$$

Thus

$$\begin{aligned} \text{NORM}(\underline{x \text{ and not very } x}) = & \int_0^{1.3a} 1.6(1 + (\frac{y}{a})^{-2})^{-1}/y \\ & + \int_{1.3a}^{\infty} 1.6(1 - (1 + (\frac{y}{a})^{-2})^{-2})/y \quad (183) \end{aligned}$$

The membership function described by the right-hand side of (183) does not provide a good approximation to the desired shape of slightly x as depicted in Fig. 12b. More important, none of the expressions for slightly x given above is suitable for operation on terms having non-fuzzy meaning since in the case of such terms plus x = very x = x and hence slightly x yields an empty set.

To illustrate this point, suppose that $x = \underline{\text{positive number}}$, with the requirement that the membership function of slightly positive number should have the form shown in Fig. 13. Then, using the definition of the pulse-function $\pi(u)$ (see (87)), we can approximate to slightly x by

$$\underline{\text{slightly } x} = \text{SF}(x ; K) \quad (184)$$

where

$$\begin{aligned} K(y) = & \int_0^{\infty} \pi\left(\frac{u-a}{w}\right)/u \quad \text{for } y = a, \quad u \in U \quad (185) \\ & = 0 \quad \text{for } y \neq a \end{aligned}$$

Note that the values of the parameters a and w are determined by the

context in which the term slightly positive number is used.* In general, the context-dependence of the meaning of slightly x makes it difficult to construct a definition for the hedge slightly which has a broad validity.

Sort of

Sort of is a member of a family of hedges which have the effect of reducing the grade of membership of those objects which are in the "center" of a class x and increasing those which are on its periphery. Fig. 14 shows in graphical terms the relation between x and sort of x when x is a fuzzy set of the form

$$x = \int_0^{\infty} (1 + (\frac{y}{a})^2)^{-1}/y \quad (186)$$

in which the parameter a is the crossover value of y .

For the fuzzy set defined by (186), sort of x can be approximated to by the following expression

$$\underline{\text{sort of}} x = \text{NORM}(\neg \text{CON}^2(x) \cap \text{DIL}(x)) \quad (187)$$

in which the term $\neg \text{CON}^2(x)$ serves to reduce the grade membership of those points which are close to zero, while $\text{DIL}(x)$ (see (78)) increases the grade of membership of points which are remote from zero. Thus, if x is interpreted as small, then (187) might read

* The parameter a defines the location of the peak of the pulse while w defines its width, i.e., the separation between the crossover points. Note that for $y \neq a$, the grade of membership of y in $K(y)$ is not unity (see the definition of the kernel set).

$$\text{sort of small} = \text{NORM} (\text{more or less small but not very very small}) \quad (188)$$

where DIL is interpreted as more or less and but plays the role of and.

6. Representation of Hedges of Type II

As was stated in Section 4, the distinguishing feature of hedges of Type II is that their characterization as operators involves a description of the manner in which they affect the components of the operand.

For this reason, the characterization of hedges of Type II is a considerably more complex problem than that of hedges of Type I. Thus, in general, the definition of a hedge of Type II has to be formulated as a fuzzy algorithm involving hedges of Type I. Definitions of this kind will be discussed in a subsequent paper.

In what follows, we shall merely touch upon a relatively simple case in which the effect of a hedge can be described - to a first approximation - as a modification in the weighting coefficients of a convex combination.

Specifically, consider the hedge essentially and assume that its operand is a term x whose components are denoted by x_1, \dots, x_n . For example, $x = \text{decent}$, with the components of x assumed to be $x_1 = \text{kind}$, $x_2 = \text{honest}$, $x_3 = \text{polite}$ and $x_4 = \text{attractive}$. We assume, further, that the fuzzy set x is a convex combination of its components (see (88)), that is,

$$\mu = w_1\mu_1 + w_2\mu_2 + w_3\mu_3 + w_4\mu_4 \quad (189)$$

where the w_i , $i = 1, \dots, n$ are non-negative weights whose sum is unity, μ is the grade of membership of an individual y in x , and μ_i , $i = 1, \dots, 4$ is the grade of membership of y in x_i . For concreteness, suppose that $w_1 = 0.4$, $w_2 = 0.3$, $w_3 = 0.2$ and $w_4 = 0.1$. Then, if $\mu_1(\text{JOHN}) = 0.9$, $\mu_2(\text{JOHN}) = 0.8$, $\mu_3(\text{JOHN}) = 0.9$ and $\mu_4(\text{JOHN}) = 0.2$, we have

$$\begin{aligned} \mu(\text{JOHN}) &= 0.4 \times 0.9 + 0.3 \times 0.8 + 0.2 \times 0.9 + 0.1 \times 0.2 \\ &= 0.8 \end{aligned} \tag{190}$$

Now the magnitude of w_i is a measure of the importance of the attribute labeled x_i . Intuitively, the hedge essentially has the effect of increasing the weights of the important attributes and diminishing those that are relatively unimportant. To achieve this effect, let us first normalize the weights w_i so that if w_1 is the largest weight, then after normalization we have

$$w'_1 = 1 \tag{191}$$

$$w'_2 = \frac{w_2}{w_1} = 0.75$$

$$w'_3 = \frac{w_3}{w_1} = 0.5$$

and

$$w'_4 = \frac{w_4}{w_1} = 0.25$$

On squaring the normalized weights, forming their sum and dividing each normalized weight by the sum, we obtain (approximately)

$$w_1^* = \frac{1}{1.87} = 0.53 \quad (192)$$

$$w_2^* = \frac{0.56}{1.87} = 0.3$$

$$w_3^* = \frac{0.25}{1.87} = 0.13$$

$$w_4^* = \frac{0.06}{1.87} = 0.04$$

In terms of these weights, the effect of essentially on decent can be approximately characterized by

$$\mu^* = 0.53 \mu_1 + 0.3 \mu_2 + 0.13 \mu_3 + 0.04 \mu_4 \quad (193)$$

where μ^* denotes the membership function of essentially decent. Thus, in the case of JOHN, we have

$$\begin{aligned} \mu^*(\text{JOHN}) &= 0.53 \times 0.9 + 0.3 \times 0.8 + 0.13 \times 0.9 + 0.04 \times 0.2 \\ &= .91 \end{aligned} \quad (194)$$

which is higher than $\mu(\text{JOHN})$ since the low grade of membership of JOHN in the fuzzy set attractive is weighed less heavily in essentially decent than it is in decent.

In short, in the very approximate characterization of the hedge essentially given above, the effect of essentially is described as a change in the coefficients of the convex combination, with the "important" components of x increased in magnitude in relation to the less important ones. The same general approach can be used to characterize the effect

of hedges such as technically, regular, etc. provided the membership function of x can be expressed as a convex combination of its components. When this is not the case, the characterization of the effect of the hedge becomes a more complicated problem which, in general, requires a fuzzy-algorithmic approach to its solution.

Concluding Remarks

In the foregoing discussion, we have concerned ourselves with relatively simple hedges whose effect can be described in terms of combinations or compositions of the elementary operations of complementation, concentration, intersection, dilation, accentuation, contrast intensification and convex combination.

We have not considered hedges which are context-dependent to a degree where it is not feasible to represent them, even very approximately, in terms of a set of elementary operations such as those defined in this paper. Such hedges require the use of a fuzzy-algorithmic mode of characterization which is more qualitative than the approach described here. Examples of such hedges and their fuzzy-algorithmic definitions will be presented in subsequent papers.

References

1. G. Lakoff, "Linguistics and Natural Logic," in Semantics of Natural Languages, D. Davidson and G. Harman (eds.), D. Reidel, Dordrecht, the Netherlands, 1971.
2. G. Lakoff, "Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts," Proc. 8th Regional Meeting of Chicago Linguistic Society, University of Chicago Linguistics Dept., April, 1972.
3. L. A. Zadeh, "Fuzzy Sets," Information and Control, Vol. 8, pp. 338-353, 1965.
4. L. A. Zadeh, "Fuzzy Algorithms," Information and Control, Vol. 12, pp. 94-102, 1968.
5. E. T. Lee and L. A. Zadeh, "Note on Fuzzy Languages," Information Sciences, Vol. 1, pp. 421-434, 1969.
6. L. A. Zadeh, "Quantitative Fuzzy Semantics," Information Sciences, Vol. 3, pp. 159-176, 1971.
7. L. A. Zadeh, "Fuzzy Languages and Their Relation to Human and Machine Intelligence," Proc. of Conference on Man and Computer, Bordeaux, France, 1970. Memorandum M-302, 1971, Electronics Research Laboratory, University of California, Berkeley, Calif. 94720.
8. L. A. Zadeh, "Similarity Relations and Fuzzy Orderings," Information Sciences, Vol. 3, pp. 177-200, 1971.
9. S. Tamura, S. Higuchi and K. Tanaka, "Pattern Classification Based on Fuzzy Relations," Trans. on Systems, Man and Cybernetics, Vol. 1, pp. 61-66, 1971.
10. J. A. Goguen, "The Logic of Inexact Concepts," Synthese, 19, pp. 325-373, 1969.

11. L. Wittgenstein, "Logical Form" Proc. of the Aristotelian Society, Vol. 9, pp. 162-171, 1929.
12. M. Black, "Reasoning with Loose Concepts," Dialogue, Vol. 2, pp. 1-12, 1963.
13. M. Scriven, "The Logic of Criteria," Jour. of Philosophy, Vol. 56, pp. 857-868, 1959.
14. H. Khatchadourian, "Vagueness, Meaning and Absurdity," American Philosophical Quarterly, Vol. 2, pp. 119-129, 1965.
15. R. R. Verma, "Vagueness and the Principle of Excluded Middle," Mind, Vol. 79, pp. 67-77, 1970.

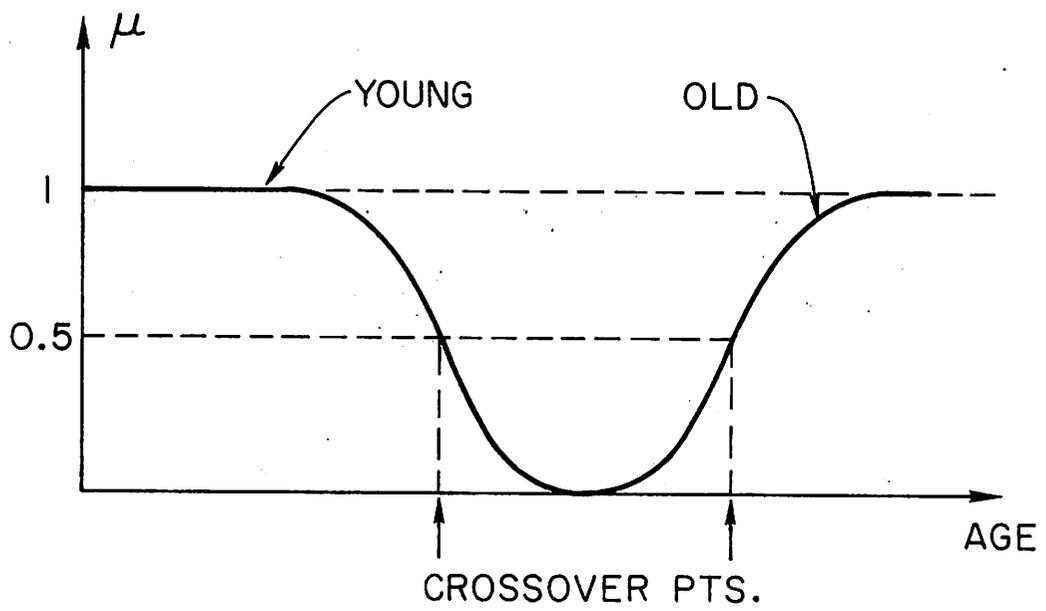


Fig. 1. Membership functions of young and old.

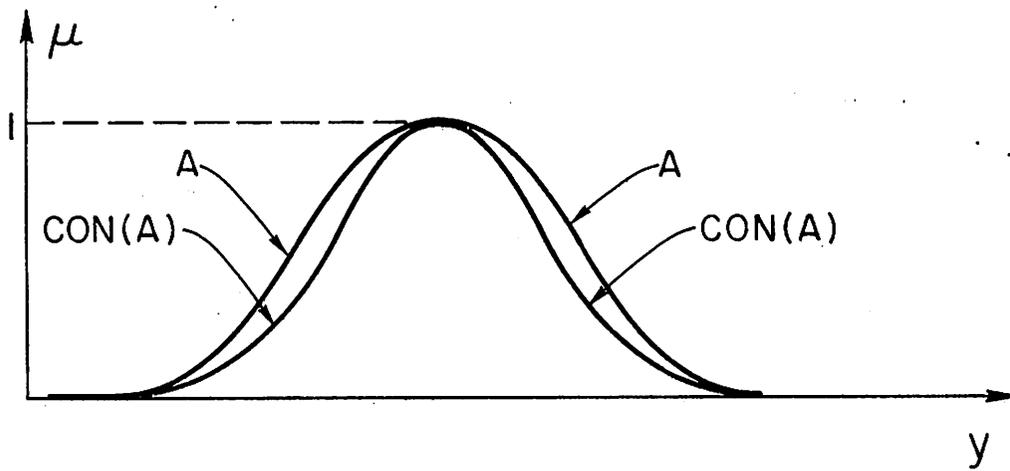


Fig. 2. The effect of concentration on a fuzzy set A.

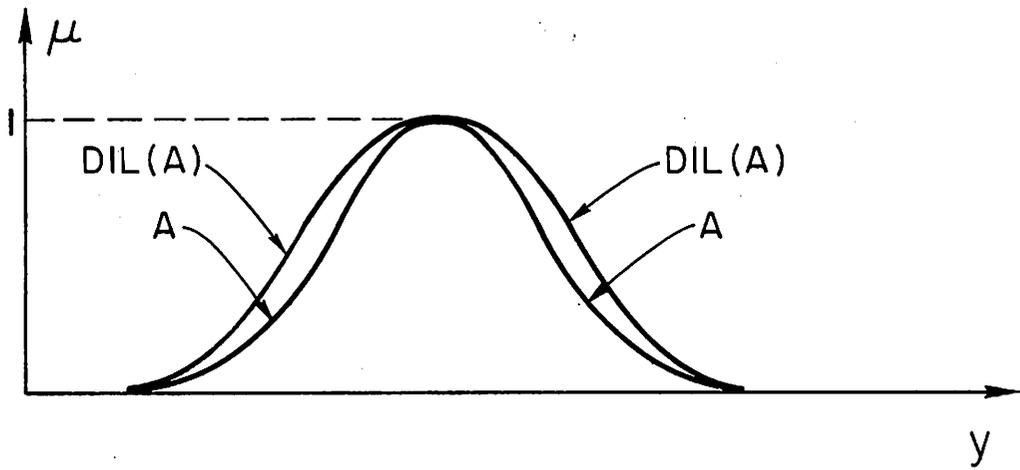


Fig. 3. The effect of dilation on the operand A.

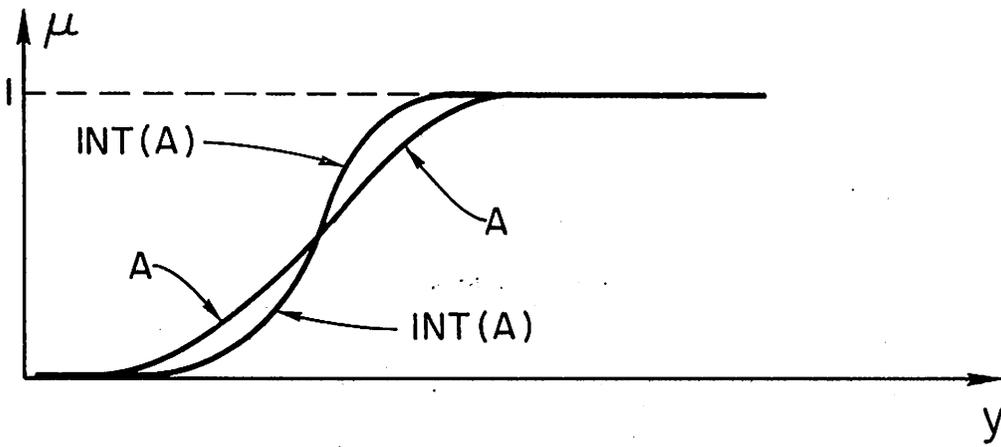


Fig. 4. The effect of intensification on the operand A.

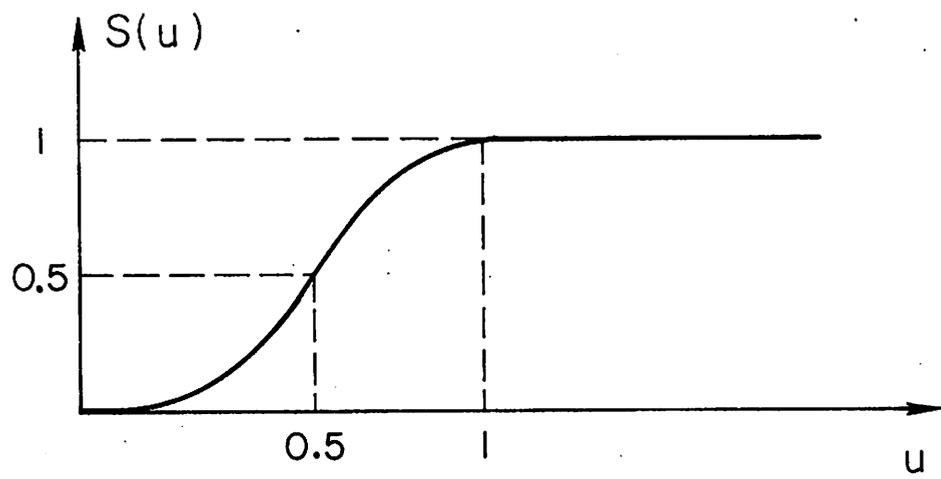


Fig. 5. The form of S-function.

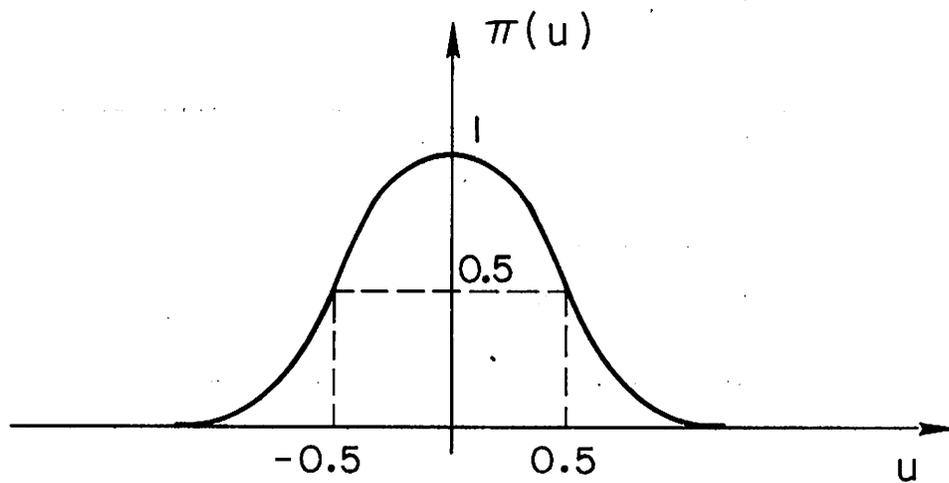


Fig. 6. The form of π -function.

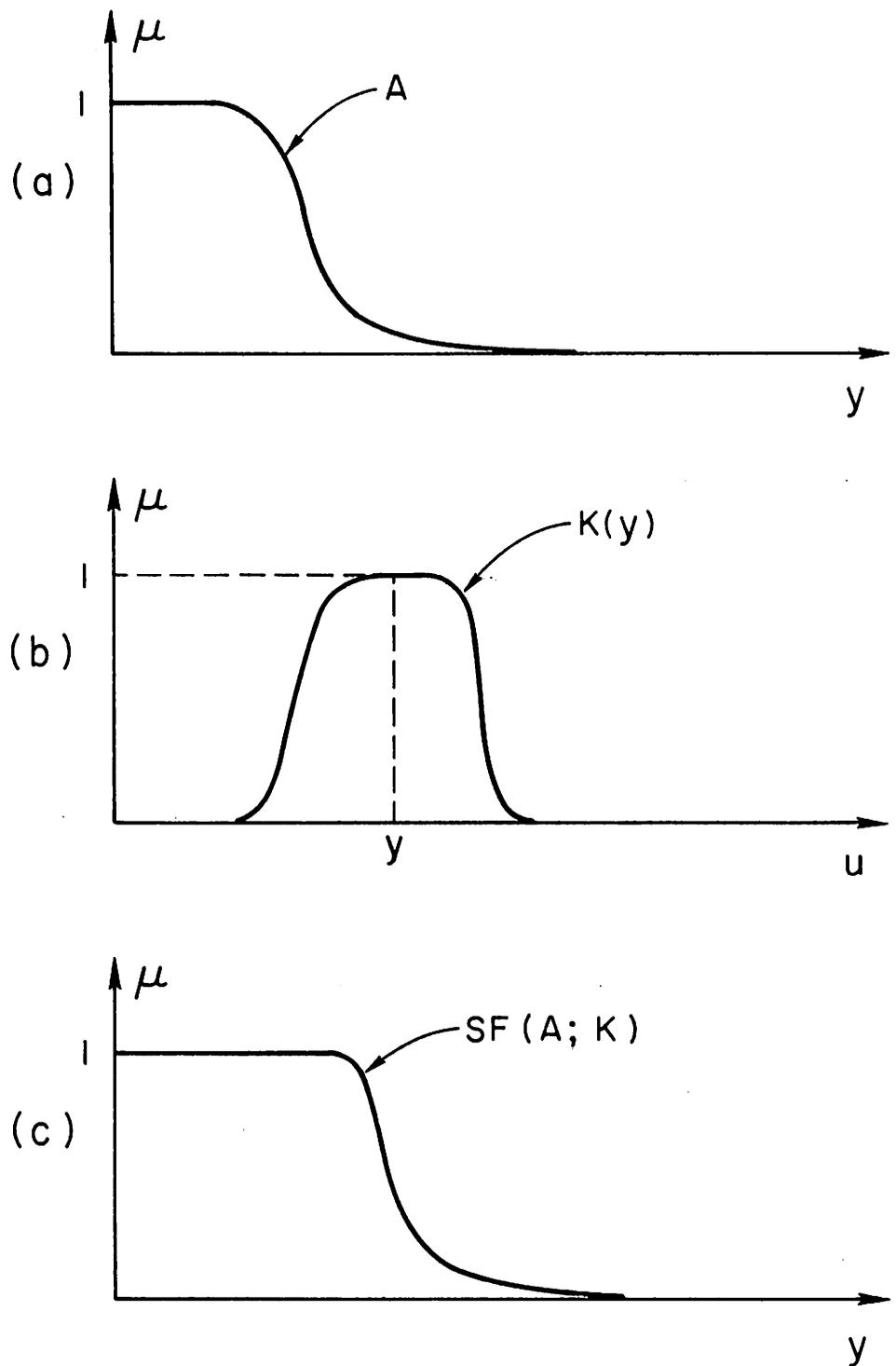


Fig. 7. The effect of fuzzification of a fuzzy set A with kernel set K .

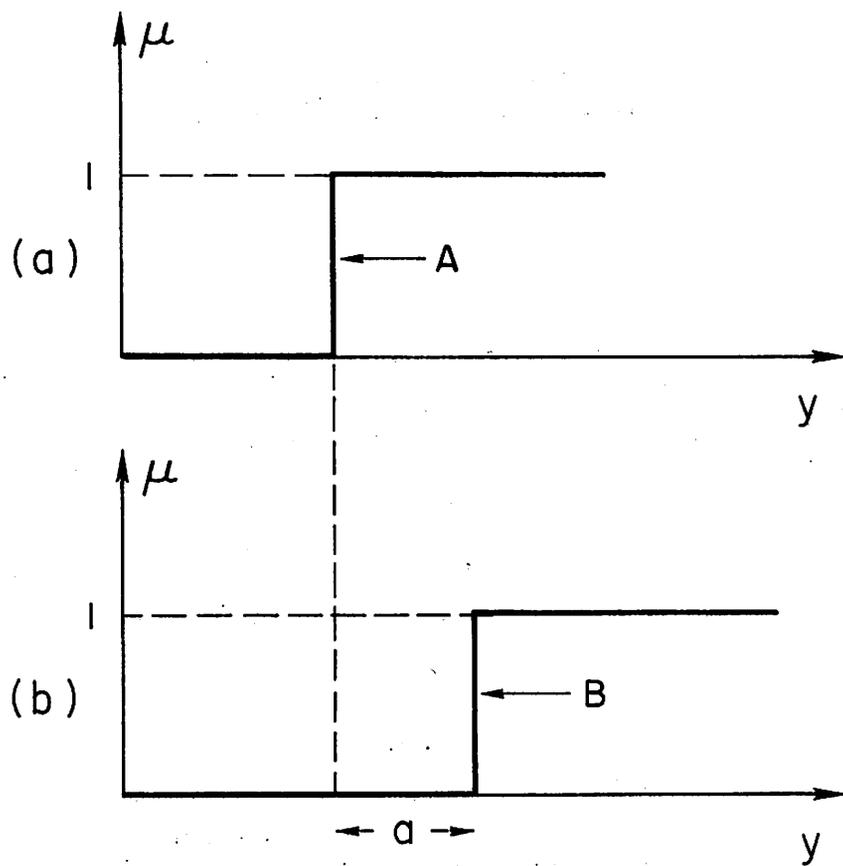


Fig. 8. Translation of A by means of fuzzification.

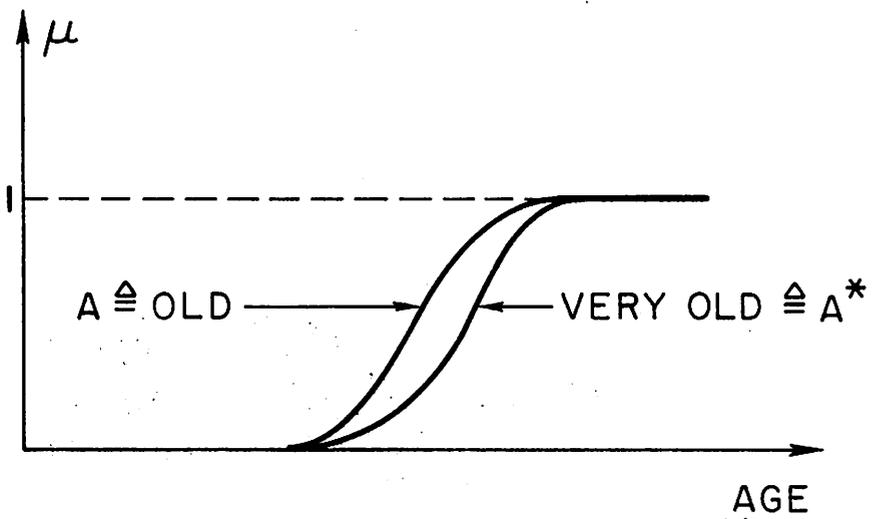


Fig. 9. The effect of the hedge very.

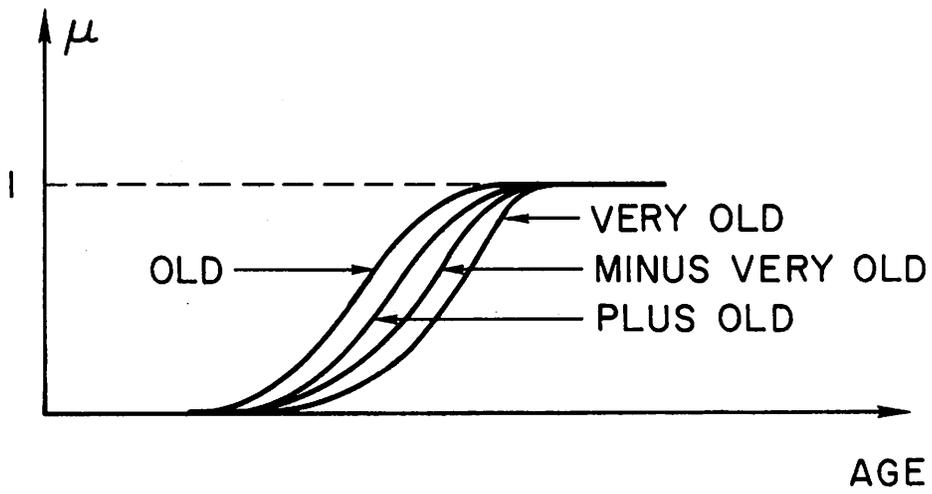


Fig. 10. The effect of the hedges plus, minus and very.

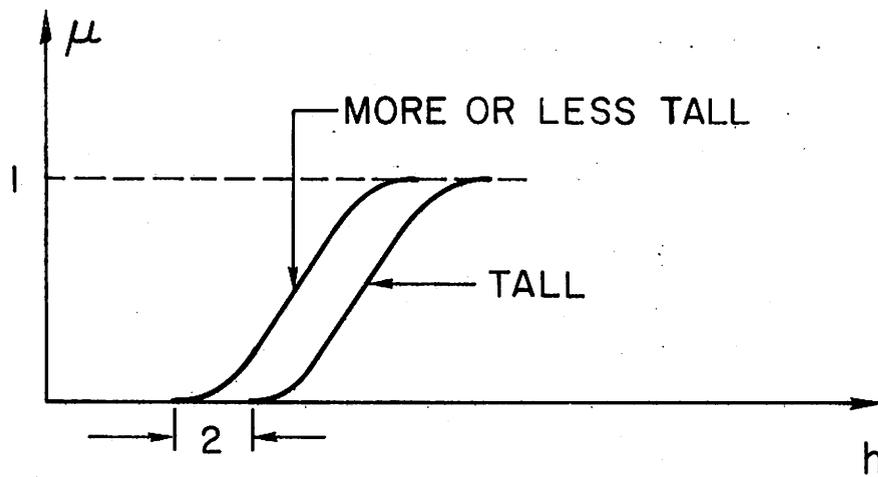


Fig. 11. The effect of the hedge more or less.

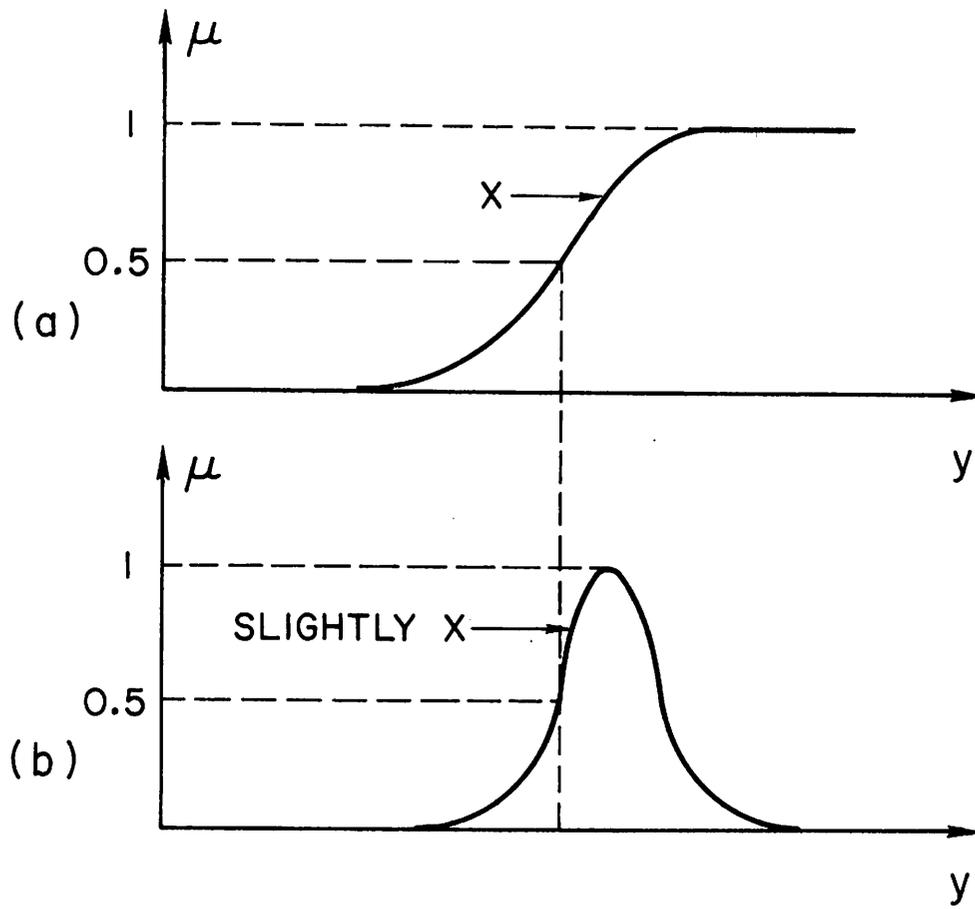


Fig. 12. The effect of the hedge slightly.

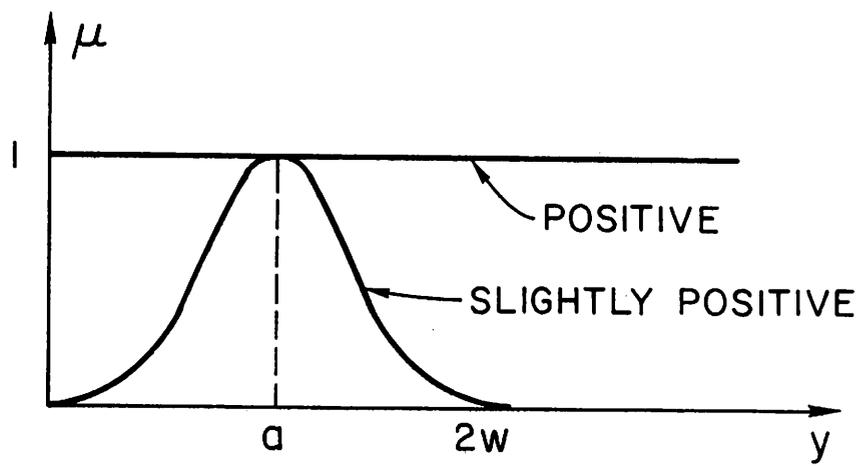


Fig. 13. The effect of slightly achieved by fuzzification.

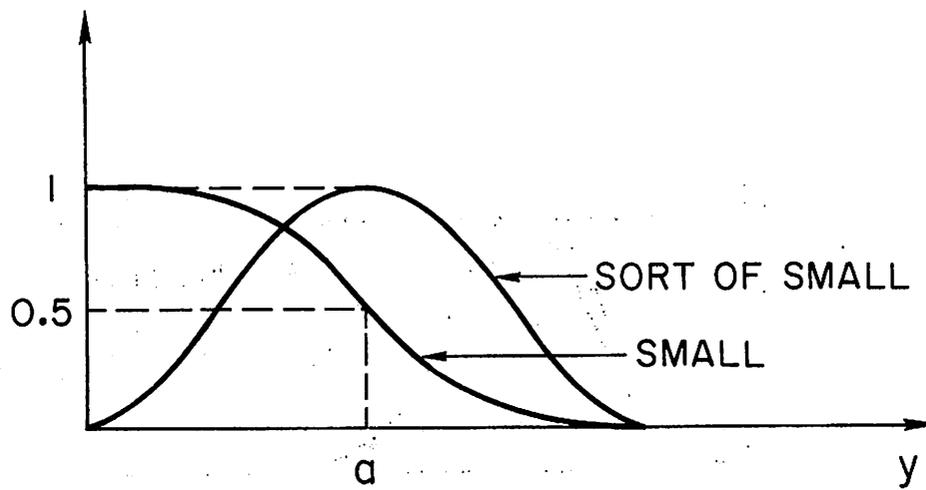


Fig. 14. The effect of the hedge sort of.