DISCRIMINATIVE EFFECT OF INFORMATION

--THE THIRD MEASURE OF INFORMATION--

by

Y. Enta

Memorandum No. UCB/ERL M78/62

May 1978

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

TABLE OF CONTENTS

ABSTRACT

In decision-making, discrimination of fuzzy elements is as important as dealing with randomness. The problems of the latter have been investigated primarily based on probability theory developed since Pascal's ponderings in the 17th century. But the study of the former has not yet been fully explored due to the fact that it has been less than thirteen years since L. A. Zadeh initiated the serious study of fuzziness in 1965.

This paper presents a measure by which the effect of information on fuzziness is estimated through discrimination of fuzzy elements. This paper follows the concept of possibility distribution which is a key concept in possibility theory derived from fuzzy set theory. Through the degree of separation among some possibility distributions, we attempt to devise a measure for fuzziness, which will be referred to as the discriminative effect of information.

The structure of this measure for fuzziness is similar to the structures of the measures for randomness in that what is called amount of information uses the concept of entropy and what is called value of information uses the concept of expected utility. Both of the measures for randomness are grounded in probability theory which is derived from random theory.

Then, several basic methods of information-processing are discussed from the point of view of the measure of the discriminative effect of information.

This measure may be regarded as the third measure of information, and it appears to be of relevance to the problems of pattern-recognition as well as decision-making.

INTRODUCTION

The quantity of information to which we are subjected today is so immense that it defies any comparison with the past. Indeed, we live in the age of the Information Society. Yet, when we take a closer look at information that surrounds us, we realize that a certain aspect of it has not changed over the years. That aspect is uncertainty. Most information contains:

1. Uncertainty as probabilistic event (to be referred to as randomness), and
2. Uncertainty as vagueness of the meaning of the event itself (to be referred to as fuzziness).

Information such as "John will graduate from the University in June" is an example of information containing the first category of uncertainty. Likewise, "Wendy is rich" exemplifies uncertainty of the second category. "Tom will marry an attractive girl" is an example that contains both kinds of uncertainty. An explanation of the first category is unnecessary, as it is obvious to most people. However, the meanings of words such as rich and attractive are not fixed, but they are rather relative and therefore permit vagueness.

The field of information-processing investigates how information with all its uncertainty can best be economically utilized for better decision-making. The inevitable task here is to measure the effect of the change in magnitude of uncertainty as the result of some information-processing.

C. E. Shannon developed the measure of amount of information from entropy; and in the area of statistical decision theory, the measure of value of information was developed from the concept of expected utility.

These two measures are very important tools today in the field of information-processing. But these measures are only applied to the uncertainty of the first category, i.e., randomness. This paper attempts to present a measure for the second category, i.e., fuzziness.[1/]

Here are two simple decision problems.

Example A. A choice must be made between two alternatives, a_1 and a_2. The available information reads as follows:

"The alternative a_1 seems to have a payoff of near 100, while the alternative a_2 seems to have a payoff of near 101."

Does this information invite an immediate decision to choose alternative a_2? Not likely. Provided the research cost is not too high, the decision-maker will require further information. He is more likely to make a choice and decide on a_2 after confirming that, for instance, "the payoff of a_1 is very near 100, while that of a_2 is very near 101."

Example B. In the case of an enraged husband whose wife has just been taken hostage by a kidnapper, he would not start shooting at the kidnapper based on the information that "the person on the right seems to be the criminal and the one on the left seems to be his wife." He has to wait for further information, no matter how costly it may be, which tells him that "it is almost certain that the person on the right is the criminal,"

---

[1]Regarding a measure for fuzziness, there is an excellent study done by A. De Luca and S. Termini. The theme of their paper is to apply the concept of entropy to any fuzzy set in order to directly measure the degree of fuzziness of that set. Therefore, it is similar in its method and implication to the measure of amount of information which is now well known.

In contrast, the theme of this paper is to indirectly measure the degree of effect of information on fuzziness through discrimination of fuzzy objects. This concept of measure is similar to the concept of resolving power in optics.

It is easily anticipated that we will find a close and interesting relation between the measure presented by A. De Luca and S. Termini and the measure proposed in this paper. The writer would like to discuss this point as the subject of another paper in the near future.

Cf. De Luca, A., and S. Termini [2].

and/or "it is now confirmed that the person on the right is your wife," before he can make his move.

These two examples illustrate that the degree of discrimination of the fuzzy elements is heightened by further information. In both cases, while the degree of discrimination is low, the decision was made to gather further information, and the choice was made only after the higher degree of discrimination had been obtained by further information. The degree of discrimination plays an important role in decision-making.

However, there have been few serious studies which investigate the effects of information-processing at the stages of discriminating the fuzzy elements. Because investigations in the area of fuzziness have started only recently. wait until L. A. Zadeh published the theory of "fuzzy Sets" in 1965.[2/]

This paper follows the concept of possibility distribution which is a key concept in possibility theory which, in turn, is based on fuzzy set theory. Through the degree of separation among some possibility distributions, we attempt to present a measure for fuzziness, which will be referred to as discriminative effect of information in this paper. The structure of this measure is similar to the structures of the measures for randomness, i.e., amount of information and value of information, which are developed from probability theory based on random theory by using the concepts of entropy and expected utility, respectively (Table 1).

In this paper I would like to discuss the characteristics of the discriminative effect of basic information-processing. In the concluding remarks of my earlier paper, "Fuzzy Choice Models" (1976),[3/] I touched on the possibility of "...the third measure of information which is related

---

[2] Cf. Zadeh, L. A. [6].

[3] Cf. Enta, Y. [3].

to discrimination among the fuzzy objects might be developed based on the fuzzy set theory." I am now attempting to present that possibility.

Object	Basic theory	Instrumental measure	Measure
Randomness	Probability theory	Entropy	Amount of information
		Expected utility	Value of information
Fuzziness	Possibility theory	Degree of separation among possibility distributions	Discriminative effect of information

Table 1

Frames of three information-measures

I. POSSIBILITY DISTRIBUTION

What is the substance of the statement (or proposition), "alternative a_1 has a payoff of near 100"? L. A. Zadeh's answer to this question is that the substance is the possibility distribution derived from the statement. Let us explain the concept of possibility distribution, which is not yet well known, by quoting his explanation:[1/]

> What is a possibility distribution? It is convenient to answer this question in terms of another concept, namely, that of a fuzzy restriction to which the concept of a possibility distribution bears a close relation.
>
> Let X be a variable which takes values in a universe of discourse U, with the generic element of U denoted by u and
>
> $$X = u, \tag{1.1}$$
>
> signifying that X is assigned the value u, $u \in U$.
>
> Let F be a fuzzy subset of U which is characterized by a membership function μ_F. Then F is a fuzzy restriction on X (or associated with X) if F acts as an elastic constraint on the values that may be assigned to X --in the sense that the assignment of a value u to X has the form
>
> $$X = u: \mu_F(u), \tag{1.2}$$
>
> where $\mu_F(u)$ is interpreted as the degree to which the constraint represented by F is satisfied when u is assigned to X. Equivalently, (1.2) implies that $1 - \mu_F(u)$ is the degree to which the constraint in question must be stretched in order to allow the assignment of u to X.[2/]

---

[1] Cf. Zadeh, L. A. [9, pp. 5-7].

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

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

$$R(X) = F. \tag{1.3}$$

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

To illustrate the concept of a fuzzy restriction, consider a proposition of the form p $\triangleq$ X is F,[3] where X is the name of an object, a variable or a proposition, and F is the name of a fuzzy subset of U, as in "Jessie is very intelligent," "X is a small number," "Harriet is blond is quite true," etc. In more detail, the translation of such a proposition may be expressed as

$$R(A(X)) = F, \tag{1.4}$$

where A(X) is an implied attribute of X which takes values in U, and (1.4) signifies that the proposition p $\triangleq$ X is F has the effect of assigning F to the fuzzy restriction on the values of A(X).

As a simple example of (1.4), let p be the proposition "the payoff of alternative a_1 is near 100," in which near 100 is a fuzzy subset of real numbers, i.e., $U = R^1$ characterized by the membership function illustrated in Figure 1.1. In this case, the implied attribute A(X) is Payoff (alternative a_1) and the translation of "The payoff of alternative a_1 is near 100" assumes the form:

---

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

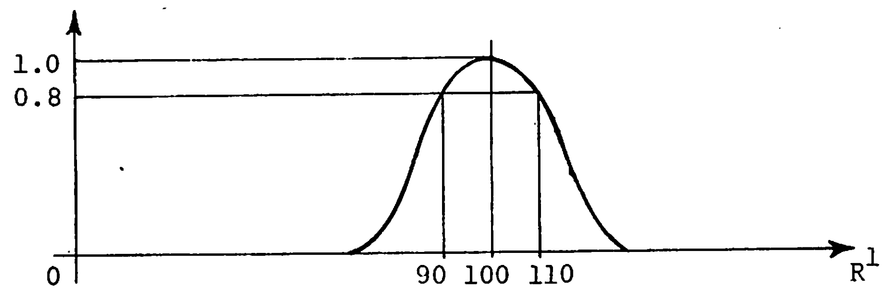

Figure 1.1: Fuzzy set; near 100.

The payoff of alternative a_1 is near 100 $\rightarrow$ Payoff (alternative a_1) = near 100. (1.5)

To relate the concept of a fuzzy restriction to that of a possibility distribution, we interpret the right-hand member of (1.5) in the following manner.

Consider a payoff, say u = 90, whose grade of membership in the fuzzy set near 100 (as defined by Figure 1.1) is 0.8. First, we interpret 0.8 as the degree of compatibility of 90 with the concept labeled near 100. Then, we postulate that the proposition "The payoff of alternative a_1 is near 100" converts the meaning of 0.8 from the degree of compatibility of 90 with near 100 to the degree of possibility that the payoff of alternative a_1 is 90 given the proposition "The payoff of alternative a_1 is near 100." In short, the compatibility of a value of u with near 100 becomes converted into the possibility of that value of u given "The payoff of alternative a_1 is near 100."

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

Definition 1.1. Let F be a fuzzy subset of a universe of discourse U which is characterized by its membership function μ_F, with the grade of membership, $\mu_F(u)$, interpreted as the compatibility of u with the concept labeled F.

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

$$R(X) = F, \tag{1.6}$$

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

$$\Pi_X = R(X). \tag{1.7}$$

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

$$\pi_X \triangleq \mu_F. \tag{1.8}$$

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

In view of (1.7), the relational assignment equation (1.6) may be expressed equivalently in the form

$$\Pi_X = F, \tag{1.9}$$

placing in evidence that the proposition p $\triangleq$ X is F has the effect of associating X with a possibility distribution Π_X which, by (1.7), is equal to F. When expressed in the form of (1.9), a relational assignment equation will be referred to as a possibility association equation, with the understanding that Π_X is induced by p.

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

$$\text{small integer} = 1/1 + 1/2 + 1/3 + 0.8/4 + 0.5/5 + 0.2/6. \quad (1.10)$$

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

$$\Pi_X = 1/1 + 1/2 + 1/3 + 0.8/4 + 0.5/5 + 0.2/6,$$

in which a term such as 0.8/4 signifies that the possibility that X is 4, given that X is a small integer, is 0.8.

II. DEGREE OF SEPARATION AMONG POSSIBILITY DISTRIBUTIONS

We assume that we have two different objects (variables or propositions), X and X', to be distinguished from each other. We then assume the following two propositions concerning X and X', respectively:

$$p \triangleq X \text{ is } F$$

and

$$p' \triangleq X' \text{ is } F'.$$

When there is no other information, we discriminate X and X' based only on the two possibility distributions, i.e., Π_X (= F) and $\Pi_{X'}$ (= F'), which are derived from the propositions above.

Even when we are able to distinguish that X is X and X' is X', there are differences in the degree of discrimination depending on the degrees of fuzziness of the propositions. We have to guess that the relative position and shape of the two curves that represent each possibility distribution, which we will refer to as relative relation, differentiate the degree of discrimination. Therefore, the clue to the measure for fuzziness must be found in the degree of discrimination.

For instance, as in Example A above, when we have the following two propositions about alternatives a_1 and a_2,

$$p \triangleq \text{The payoff of } a_1 \text{ is near 100} \tag{2.1}$$

and

$$p' \triangleq \text{The payoff of } a_2 \text{ is near 101,} \tag{2.2}$$

two alternatives, a_1 and a_2, are discriminated by the relative relation between the two fuzzy subsets of real numbers R^1, near 100 and near 101.

Now, there are many aspects of a relative relation. We face the problem of making a choice of the most suitable aspect which expresses the clarity of discrimination. The most suitable and meaningful choice for that purpose seems to be the degree of separation among possibility distributions, which we will explain later. It is analogous that entropy is chosen as the randomness measure for the purpose of the efficient signal transmission and expected utility is devised as the randomness measure for better decision.

When some information-processing changes the degree of separation among possibility distributions, the amount of the change is regarded as the discriminative effect of the information-processing. It is analogous that amount of information and value of information are considered as the amounts of the change of instrumental measures such as entropy and expected value.

Stated in more general terms, the concept of discriminative effect of information may be defined as follows:

Definition 2.1. Let

$$p^{(1)} \triangleq X^{(1)} \text{ is } F^{(1)}$$

$$\vdots$$

$$p^{(n)} \triangleq X^{(n)} \text{ is } F^{(n)}$$

be propositions about n objects, variables, or propositions, $X^{(1)},\ldots,X^{(n)}$. Let the degree of separation, DS, among n possibility distributions $\Pi_{X^{(1)}}(= F^{(1)}),\ldots,\Pi_{X^{(n)}}(= F^{(n)})$ derived from n propositions, be α, i.e.,

$$DS(F^{(1)},\ldots,F^{(n)}) = \alpha. \tag{2.3}$$

As the result of new information, I, let the proposition about $X^{(i)}$ be changed to

$$q^{(i)} = X^{(i)} \text{ is } G, \qquad i = 1,\ldots,n$$

and the degree of separation among n possibility distributions $\Pi_{X^{(1)}}(= F^{(1)}),\ldots,\Pi_{X^{(i)}}(= G),\ldots,\Pi_{X^{(n)}}(= F^{(n)})$ about $X^{(1)},\ldots,X^{(i)},\ldots,X^{(n)}$ be changed to β, i.e.,

$$DS(F^{(1)},\ldots,G,\ldots,F^{(n)}) = \beta. \tag{2.4}$$

Then, the discriminative effect of the information I, DE(I), is defined as $-(\alpha - \beta)$, i.e.,

$$\begin{aligned} DE(I) &= -(\alpha - \beta) \\ &= -[DS(F^{(1)},\ldots,F^{(i)},\ldots,F^{(n)}) - DS(F^{(1)},\ldots,G,\ldots,F^{(n)})]. \end{aligned} \tag{2.5}$$

As a simple illustration, let us assume that we have two propositions about two alternatives, such as (2.1) and (2.2). Let the degree of separation between two possibility distributions, near 100 and near 101, be 0.1, i.e.,

$$DS(\underline{\text{near } 100}, \underline{\text{near } 101}) = 0.1.$$

By adding another information about alternative a_1, I, let proposition (2.1) be changed to

$$q \triangleq \text{The payoff of } a_1 \text{ is very near } 100. \tag{2.6}$$

Then the degree of separation between very near 100 and near 101 becomes 0.4, i.e.,

$$DS(\underline{\text{very near } 100}, \underline{\text{near } 101}) = 0.4.$$

Then, the discriminative effect of the information I, DE(I), is $-(0.1 - 0.4) = 0.3$.

Definition 2.2.[1] Let n possibility distributions $\Pi_{x^{(i)}}$ (= $F^{(i)}$), i = 1,...,n, be convex fuzzy subsets[2] of a universe of discourse U. For each possibility distribution, we introduce the set $\Gamma_{F_\alpha^{(i)}}$ defined by

$$\Gamma_{F_\alpha^{(i)}} = \{u \mid \mu_{F^{(i)}}(u) > \alpha\}, \qquad \alpha \in [0, 1]; \quad i = 1,\ldots,n \tag{2.6}$$

Let C be Inf α such that $\Gamma_{F_\alpha^{(i)}}$, i = 1,...,n, are disjoint.[3] The number 1 - C will be called the degree of separation of n possibility distributions, $DS(\Pi_{x^{(1)}},\ldots,\Pi_{x^{(n)}})$, i.e.,

$$DS(\Pi_{x^{(1)}},\ldots,\Pi_{x^{(n)}}) = 1 - C. \tag{2.7}$$

---

[1] The definition of a degree of separation in this paper differs from the definition by L. A. Zadeh used in his paper, "Fuzzy Sets," 1965. The definition by Zadeh is suitable when there are only two fuzzy sets. But the definition set forth in this paper is applicable when there are more than two fuzzy sets. The definition by Zadeh, however, has its merit in that the fuzzy sets to be separated are not limited to be convex. Cf. Zadeh, L. A. [6, p. 351].

[2] A fuzzy subset F of U is convex if and only if the sets defined by

$$\Gamma_\alpha = \{u \mid \mu_F(u) \geqq \alpha\}$$

are convex for all α in the interval (0, 1].

The equivalent definition of convexity is as follows: F is convex if and only if

$$\mu_F(\lambda u + (1 - \lambda)u') \geqq \mu_F(u) \wedge \mu_F(u')$$

for all u and u' in U and all λ in [0, 1]. But note that this definition does not imply that $\mu_F(u)$ must be a convex function of u. Cf. Zadeh, L. A. [6, p. 347].

[3] As in the case of ordinary sets, two fuzzy sets F and F' are disjoint, if $F \cap F'$ is empty.

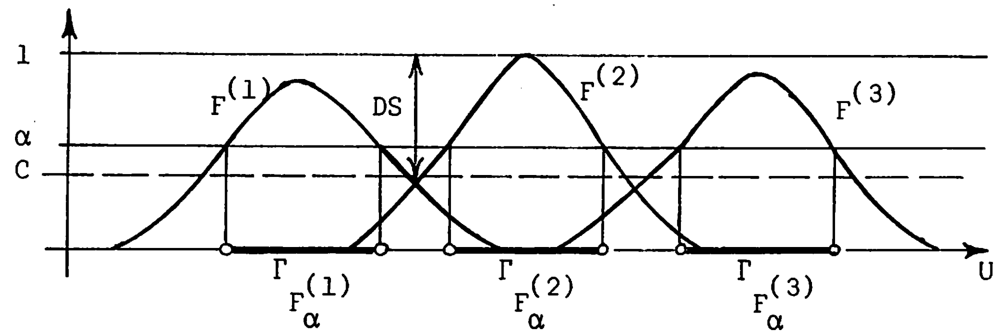

Figure 2.1: Illustration of degree of separation for n = 3.

From this, we deduce the following theorem.

Theorem 2.1. Let n possibility distributions $\Pi_{x^{(i)}}$ $(= F^{(i)})$, $i = 1,\ldots,n$, be convex fuzzy subsets of a universe of discourse U, and the maximal grade for each intersection, $F^{(i)} \cap F^{(j)}$, $i \neq j$, $i,j = 1,\ldots,n$, be M, and $\bar{M}$ be $\sup_{i,j} M$, i.e.,

$$\bar{M} = \sup_{i,j} \{\sup_{U} [\mu_{F^{(i)}}(u) \wedge \mu_{F^{(j)}}(u)]\}, \qquad i \neq j, \quad i,j = 1,\ldots,n \tag{2.8}$$

where $\wedge$ stands for the infix form of min and $\vee$ stands for the infix form of max as well. Then, the degree of separation of these possibility distributions is the number $1 - \bar{M}$, i.e.,

$$DS(\Pi_{x^{(1)}},\ldots,\Pi_{x^{(n)}}) = 1 - \bar{M}. \tag{2.9}$$

Proof. Let $\bar{M}$ be the maximal grade of intersection $F^{(l)} \cap F^{(k)}$, $l,k \leq n$. Then the two sets,

$$\Gamma_{F_{\bar{M}}^{(\ell)}} = \{u \mid \mu_{F^{(\ell)}}(u) > \bar{M}\}$$

and

$$\Gamma_{F^{(k)}_{\bar{M}}} = \{u |\ \mu_{F^{(k)}}(u) > \bar{M}\}$$

are disjoint and contiguous to each other at a point, say, u^0, such that $\mu_{F^{(\ell)}}(u^0) = \mu_{F^{(k)}}(u^0) = \bar{M}$, because if they were not, there would be a point u' that $\mu_{F^{(\ell)}}(u') > \bar{M}$ and $\mu_{F^{(k)}}(u') > \bar{M}$, and hence, $\mu_{F^{(\ell)}}(u') \wedge \mu_{F^{(k)}}(u') > \bar{M}$, which contradicts the assumption that $\bar{M} = \sup_U [\mu_{F^{(\ell)}}(u) \wedge \mu_{F^{(k)}}(u)]$.

Moreover, these two sets are disjoint from any other set, say, $\Gamma_{F^{(m)}_{\bar{M}}} = \{u |\ \mu_{F^{(m)}}(u) > \bar{M}\}$, $m < n$, because if they were not, say, $\Gamma_{F^{(k)}_{\bar{M}}}$ and $\Gamma_{F^{(m)}_{\bar{M}}}$ were not disjoint, there would be a point, say, u'', such that $\mu_{F^{(m)}}(u'') > \bar{M}$ and $\mu_{F^{(k)}}(u'') > \bar{M}$, and hence, $\mu_{F^{(m)}}(u'') \wedge \mu_{F^{(k)}}(u'') > \bar{M}$, which contradicts the assumption that $M = \sup_{i,j} \{\sup_U [\mu_{F^{(i)}}(u) \wedge \mu_{F^{(j)}}(u)]\}$, $i \neq j$, $i = 1,\ldots,n$.

On the other hand, let ε be any small positive number. Then,

$$\Gamma_{F^{(\ell)}_{\bar{M}-\varepsilon}} \supset \Gamma_{F^{(\ell)}_{\bar{M}}}$$

and

$$\Gamma_{F^{(k)}_{\bar{M}-\varepsilon}} \supset \Gamma_{F^{(k)}_{\bar{M}}}.$$

Thus, the two sets, $\Gamma_{F^{(\ell)}_{\bar{M}-\varepsilon}}$ and $\Gamma_{F^{(k)}_{\bar{M}-\varepsilon}}$ are not disjoint. Therefore, from Definition 2.2, the number $1 - \bar{M}$ is the degree of separation among them. Q.E.D.

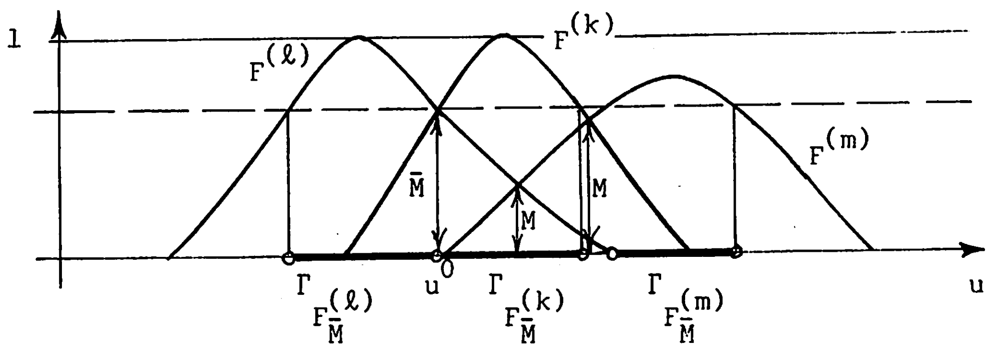

Figure 2.2

The theorem states that in the case of n = 2, the degree of separation of two convex possibility distributions is one minus the maximal grade in their intersection. For simplicity of calculating the degree of separation, we postulate that all possibility distributions which we will treat below are convex.

After all, the degree of separation can be interpreted as the index of the degree of ease in separating the possibility distributions disjointly. Therefore, it is natural to use the degree of separation as an instrument which expresses the degree of discriminating fuzzy things and hence, measures the discriminative effect of information. It is also meaningful from the points of view described below.

Obviously, from Definition 2.1, the degree of discrimination (hence, the discriminative effect of information) is in the interval between [0, 1], i.e.,

$$DS(\Pi_{x^{(1)}},\ldots,\Pi_{x^{(n)}}) \in [0,\ 1]. \tag{2.10}$$

Furthermore, it has the following characteristics:

1. The degree of separation between two possible distributions can be raised by making a sharper increase on either curve. Of course, it is

raised more drastically when both curves are increased. Translating it into terms of discrimination, in order to raise the degree of discrimination between two fuzzy things, it is sufficient to make one of the fuzzy things clarified. Of course, making both things clearer raises the degree of discrimination even further.

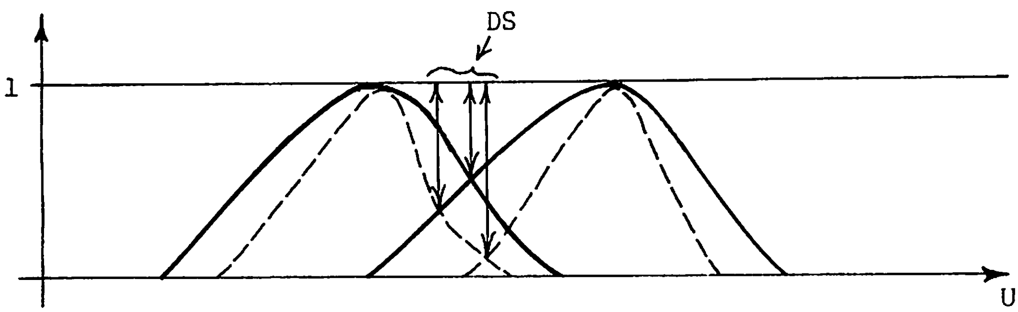

Figure 2.3

2. If a distance between two objects X and X' in U becomes larger, then the degree of separation between them becomes larger even if the shapes of their possibility distributions remain the same. If the distance becomes smaller, the degree of separation becomes smaller. In brief, when we discriminate between similar things, we have to make the propositions about them more restrictive to get the given degree of discrimination.

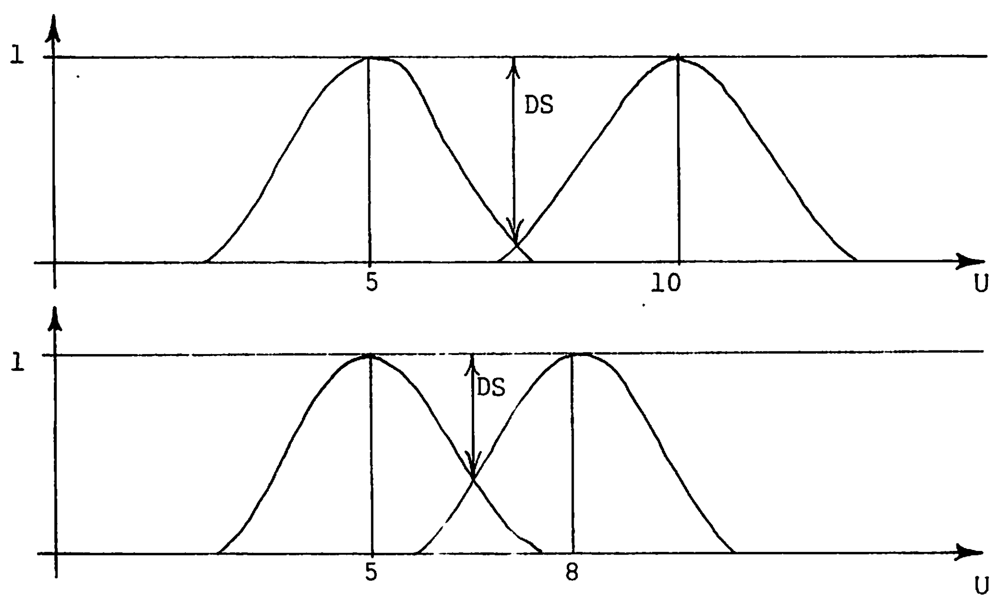

Figure 2.4

3. The degree of separation among the given number of possibility distributions is not raised by adding another possibility distribution, i.e.,

$$DS(\Pi_{x(1)},\ldots,\Pi_{x(n)}) \geq DS(\Pi_{x(1)},\ldots,\Pi_{x(n)},\Pi_{x(n+1)}). \tag{2.11}$$

This is expressed in terms of discrimination that increasing the number of fuzzy things does not result in raising the degree of discrimination. As shown in Figure 2.5, the degree of discrimination might be lowered by adding a new fuzzy thing. This relation is similar to the entropy-measure for randomness, where entropy is raised by increasing the number of transmitted signals.

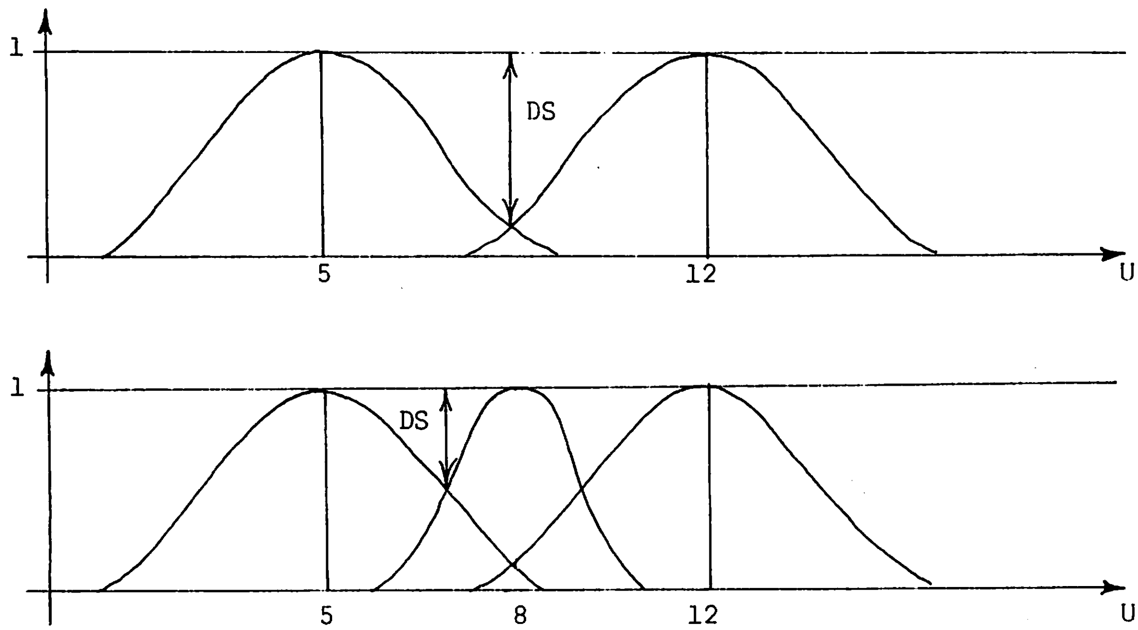

Figure 2.5

The points we described above are familiar to us in our daily experience when we make discriminations of one kind or another. Therefore, it is a very proper choice to adopt the degree of separation as the instrument which expresses the degree of discrimination and hence, measures the discriminative effect of information.

III. GENERAL CHARACTERISTICS OF DISCRIMINATIVE EFFECT OF INFORMATION

We often combine two simple propositions, such as "Paul is young" and "Paul is fat," into one composite proposition, "Paul is young and fat." On the other hand, from the point of information-economy, we often reduce a proposition which includes more than two implied attributes regarding some object to another proposition which implies fewer attributes than the original proposition does. For example, the proposition "Kazuko is a charming girl," which expresses her attributes of intelligence and looks, can be reduced to another proposition, "Kazuko is beautiful," which expresses her attribute of looks only. We know and manage many other ways of information processing. What effects do these information-processing ways have on discrimination? That is, what are the discriminative effects of these methods of information-processing? In this section, we shall discuss some important methods of information-processing.

1. Informative proposition

In the following two sets of propositions, the second set gives a clearer discrimination between Gareth and David than the first: "Gareth is tall" and "David is short"; "Gareth is tall" and "David is very short." This increased discrimination was obtained through more information-gathering about David.[1/] Here, we shall treat such a problem regarding the relation between the discrimination and the proposition change.

In general, if p is a proposition of the form p $\triangleq$ X is F, which is translated into a possibility association equation

[1] Of course, we cannot deny the possibility that information-gathering concerning David results in a proposition, e.g., "David is tall." But this kind of change of proposition from "David is short" to "David is tall" belongs to the field of randomness. In this paper, we do not dare treat this kind of change.

$$\Pi_{A(X)} = F, \tag{3.1}$$

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

$$I(p) \triangleq \Pi_{A(X)}, \tag{3.2}$$

where

$$\Pi_{A(X)} = R(A(X)) = F. \tag{3.3}$$

For example, if the proposition p $\triangleq$ David is short, then the information conveyed by p is

$$I(\text{David is short}) = \Pi_{(\text{Height(David)})},$$

where

$$\Pi_{(\text{Height(David)})} = \text{short}.$$

in which the fuzzy subset of R^1, short, might be given by Figure 3.1.

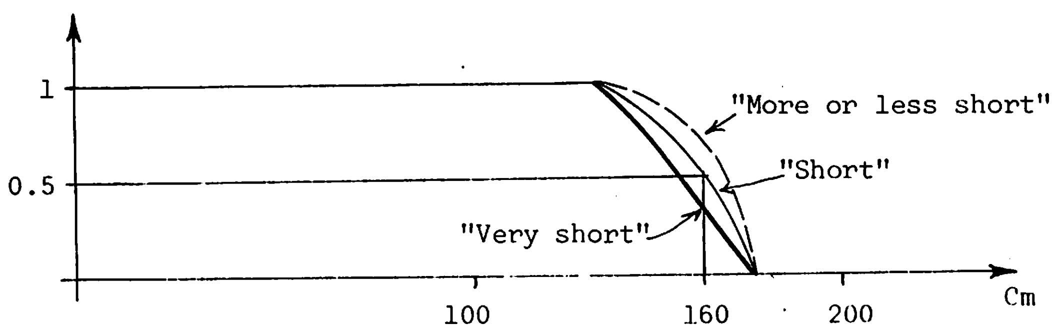

Figure 3.1

Now, if p $\triangleq$ X is F and q $\triangleq$ X is G and G $\subset$ F, then q implies p, denoted by q $\Rightarrow$ p. In that sense, q is at least informative as p, expressed as I(q) $\geqq$ I(p).

For example, if p $\triangleq$ The weather is bad today and q $\triangleq$ The weather is rainy today and rainy weather $\subset$ bad weather, then q is more informative than p, i.e.,

I(The weather is rainy today) > I(The weather is bad today).

In conclusion, the concept of informative proposition may be defined as follows:

Definition 3.1. When proposition p $\triangleq$ X is F and proposition q $\triangleq$ X is G and G $\subset$ F, then proposition q is called at least as informative as p, expressed as I(q) $\geqq$ I(p).

This definition means that if the curve of a possibility distribution derived from the proposition is sharper, then the proposition is more informative.

From the outset, getting more informative propositions or making the propositions more informative is one of the fundamental bases of information processing. Here we shall discuss the discriminative effect of the more informative proposition.

(i) Basic proposition

A basic statement concerning the discriminative effect of a more informative proposition is as follows:

Proposition 3.1. The degree of discrimination given by a more informative proposition is at least as high as the degree given by a less informative proposition. That is, the discriminative effect of a more informative proposition is nonnegative.

Note that the converse is not true. This proposition can be obtained from the following theorem:

Theorem 3.1. We assume that we have two different objects X and X', and the following three propositions concerning them:

$$p \triangleq X \text{ is } F \rightarrow \Pi_{A(X)} = F, \tag{3.4}$$

$$p' \triangleq X' \text{ is } F' \rightarrow \Pi_{A(X')} = F', \tag{3.4'}$$

$$q \triangleq X \text{ is } G \rightarrow \Pi_{A(X)} = G. \tag{3.5}$$

If $G \subset F$, then

$$DS(F, F') \leqq DS(G, F'), \tag{3.6}$$

where F, F', and G are convex fuzzy subsets of U.

Proof. The maximal grades of $F \cap F'$ and $G \cap F'$ are, respectively,

$$\bar{M}_{F\cap F'} = \sup_U [\mu_F(u) \wedge \mu_{F'}(u)], \tag{3.7}$$

$$\bar{M}_{G\cap F'} = \sup_U [\mu_G(u) \wedge \mu_{F'}(u)]. \tag{3.8}$$

From the premise, since $G \subset F$, we obtain

$$\mu_G(u) \leqq \mu_F(u), \qquad \forall\, u \in U. \tag{3.9}$$

Thus,

$$\bar{M}_{F\cap F'} \geqq \bar{M}_{G\cap F'}.$$

Moreover, F, F', and G are assumed to be convex (see p. 16). Hence, from Theorem 2.1,

$$DS(F, F') \leqq DS(G, F'). \tag{3.10}$$

Q.E.D.

Proposition 3.1 then means that, for example, when we discriminate between two persons by their wealth, it is more certain to do it by q $\triangleq$ Douglas is a millionaire and p' $\triangleq$ Peter is poor than by p $\triangleq$ Douglas is rich and p'; moreover, it is more certain to do it by q and q' $\triangleq$ Peter is destitute than by q and p'.

A more informative proposition concerning X contains more restrictive information about X. And it means that the fuzziness which is the obstacle in discrimination is less. Considering this fact, the previous proposition is self-evident.

(ii) Linguistic hedges

As was pointed out in Section I, the fuzzy set F in the proposition p $\triangleq$ X is F is acting as a fuzzy constraint on X, R(A(X)).

In natural language which we use in daily life, we have many linguistic hedges, such as very, more or less, much, lightly, etc., which act on the fuzzy constraint, then modify the fuzzy set F to F^+. That is, if

$$p \triangleq X \text{ is } F \rightarrow R(A(X)) = F$$
$$\Pi_{A(X)} = F, \tag{3.11}$$

then

$$q \triangleq X \text{ is } mF \rightarrow R(A(X)) = F^{+}$$
$$\Pi_{A(X)} = F^{+}, \quad (3.12)$$

where m is any linguistic hedge functioning as modifier on F, and F^{+} is a modification of F defined by m.

Here, we shall treat only the representative ones very and more or less, and discuss the discriminative effects about them briefly.[2/]

According to L. A. Zadeh, in (3.12), if

m = very,

then

$$F^{+} = F^{2},$$

where

$$\mu_{F^2}(u) = \mu_F(u)^2. \quad (3.13)$$

Likewise, if

m = more or less,

then

$$F^{+} = F^{1/2} \quad (3.14)$$

where

$$\mu_{F^{1/2}}(u) = \mu_F(u)^{1/2}.$$

---

[2] A more detailed discussion of linguistic hedges may be found in [5] and [7].

As an illustration, if

$$\text{David is short} \rightarrow \Pi_{\text{Height(David)}} = \text{short},$$

then

$$\text{David is very short} \rightarrow \Pi_{\text{Height(David)}} = \text{short}^2,$$

where these two possibility distributions are expressed in Figure 3.1.

From (3.13) and (3.14),

$$F^2 \subset F \subset F^{1/2}. \tag{3.15}$$

Moreover, if F is convex, then F^2 and $F^{1/2}$ are convex, too. Hence, from Theorem 3.1, we can get the following statement.

<u>Proposition 3.2</u>. A proposition with the linguistic hedge <u>very</u> is more informative than one without it; thus, its discriminative effect is non-negative. Contrarily, a proposition with <u>more or less</u> is not more informative than one without it; thus, its discriminative effect is nonpositive.

For example, the degree of discrimination between Gareth and David by the proposition "Gareth is tall" and "David is short" is lower than in the case "Gareth is very tall" and "David is short." The difference between the two degrees of discrimination above is the discriminative effect of the proposition or information with the addition of the linguistic hedge of <u>very</u>.

<u>(iii) Composition--same dimension</u>

We assume that there are n simple propositions concerning X,

$$P_{(1)} \triangleq X \text{ is } F_{(1)} \rightarrow \Pi_{A(X)} = F_{(1)} \tag{3.16}$$

$$\vdots$$

$$P_{(n)} \triangleq X \text{ is } F_{(n)} \rightarrow \Pi_{A(X)} = F_{(n)}, \tag{3.17}$$

where $F_{(1)},\ldots,F_{(n)}$ are fuzzy subsets of a universe of discourse U.

When we combine these simple propositions into one composite proposition, there are three forms, depending on the degree to which we believe each of them:

(1) $q \triangleq X$ is $F_{(1)}$ and...and $F_{(n)} \rightarrow \Pi_{A(X)} = F_{(1)} \cap \cdots \cap F_{(n)} = \bigcap_i F_{(i)}$

and

$$\pi_{A(X)}(u) = \mu_{F_{(1)}}(u) \wedge \cdots \wedge \mu_{F_{(n)}}(u) = \bigwedge_i \mu_{F_{(i)}}(u) \quad (3.18)$$

(2) $q \triangleq X$ is $F_{(1)}$ or...or $F_{(n)} \rightarrow \Pi_{A(X)} = F_{(1)} \cup \cdots \cup F_{(n)} = \bigcup_i F_{(i)}$

and

$$\pi_{A(X)}(u) = \mu_{F_{(1)}}(u) \vee \cdots \vee \mu_{F_{(n)}}(u) = \bigvee_i \mu_{F_{(i)}}(u) \quad (3.19)$$

(3) $q \triangleq X$ is $F_{(1)}$ with the degree of belief ω_1 and...

and $F_{(n)}$ with the degree of belief ω_n

$\rightarrow \Pi_{A(X)} = \omega_1 F_{(1)} + \cdots + \omega_n F_{(n)} = \sum_i \omega_i F_{(i)}$

and

$$\pi_{A(X)}(u) = \omega_1 \mu_{F_{(1)}}(u) + \cdots + \omega_n \mu_{F_{(n)}}(u) = \sum_i \omega_i \mu_{F_i}(u). \quad (3.20)$$

where + denotes the arithmetic addition, and where $\omega_i \in [0, 1]$ and $\sum_i \omega_i = 1$.

Let us explain briefly. In case (1), i.e., the conjunctive composition, all of the simple propositions are believed as true, but in case (2), i.e., the disjunctive one , at least one of the propositions is believed as true but which one is true is not known. In case (3), we shall call the composition a <u>weighted sum composition</u>; the simple propositions $p_{(1)},\ldots,p_{(n)}$ are respectively believed as true with the degrees of belief of $\omega_1,\ldots,\omega_n$. In this connection, when the $p_{(i)}$ is believed as true, with $\omega_i = 1$, the weighted sum composition becomes $p_{(i)}$ itself.

As a simple illustration,[3/] if

$$p_{(1)} \triangleq X \text{ is a small integer,} \tag{3.21}$$

$$p_{(2)} \triangleq X \text{ is a near 3 integer,} \tag{3.22}$$

where small integer is defined by

$$\text{small integer} = 1/1 + 1/2 + 1/3 + 0.8/4 + 0.5/5 + 0.2/6 \tag{3.23}$$

and where

$$\text{near 3 integer} = 0.6/1 + 0.9/2 + 1/3 + 0.9/4 + 0.6/5 + 0.2/6, \tag{3.24}$$

then the conjunctive proposition, i.e., "X is a small and near 3 integer" induces the following possibility distribution:

$$0.6/1 + 0.9/2 + 1/3 + 0.8/4 + 0.5/5 + 0.2/6; \tag{3.25}$$

the disjunctive proposition, i.e., "X is a small or near 3 integer" induces

$$1/1 + 1/2 + 1/3 + 0.9/4 + 0.6/5 + 0.2/6; \tag{3.26}$$

and the weighted sum proposition, such as "X is a small integer with 0.6 degree of belief and near 3 integer with 0.4 degree of belief," induces

$$0.84/1 + 0.96/2 + 1/3 + 0.84/4 + 0.54/5 + 0.2/6. \tag{3.27}$$

Now, from (3.18), (3.19), and (3.20),

$$\bigwedge_i \mu_{F_{(i)}}(u) \leq \sum_i \omega_i \mu_{F_{(i)}}(u) \leq \bigvee_i \mu_{F_{(i)}}(u), \qquad \forall u \in U. \tag{3.28}$$

That is,

$$\bigcap_i F_{(i)} \subset \sum_i \omega_i F_{(i)} \subset \bigcup_i F_{(i)}. \tag{3.29}$$

Of course,

$$\bigcap_i F_{(i)} \subset F_{(i)} \subset \bigcup_i F_{(i)}. \tag{3.30}$$

Then, from Theorem 3.1, we obtain the following statement:

Proposition 3.3. A conjunctive proposition is at least as informative as a weighted sum proposition, and the latter is at least as informative as the disjunctive proposition; then the magnitude of discriminative effect decreases in the following order: the conjunctive proposition, the weighted sum proposition, and the disjunctive proposition, provided that the possibility distributions of these composite propositions are convex.

The difference among these composite propositions are caused by the difference of the degrees to which we believe each of the simple propositions which compose these composite propositions. After all, when the degree of belief, that is, the degree of randomness is low, the composite proposition is more informative than in the case in which the degree of randomness is not low; and its discriminative effect becomes higher. This is one example of the influence of randomness on discrimination.

2. Composition--different dimensions

So far, we have the premise that all possibility distributions derived from the propositions are the fuzzy subsets of the same universe of discourse (of the same dimension). Here we take away the premise and discuss the method of information-processing, such as composition, more generally.

Let $p_1 \triangleq$ X's attribute A_1 is F_1 be a simple proposition concerning the attribute A_1 of $X_1, \ldots,$ and $p_n \triangleq$ X's attribute A_n is F_n be a simple proposition concerning the attribute A_n of X. There are three forms of composition of these simple propositions, such as conjunctive, disjunctive, and weighted sum composition, also.

Here, we shall discuss the discriminative effects of composition, covering multi-dimensions.

(i) Conjunctive composition

$$\text{If } p_1 \triangleq \text{X's attribute } A_1 \text{ is } F_1 \rightarrow \Pi_{A_1(X)} = F_1 \tag{3.31}$$

$$\vdots$$

$$p_n \triangleq \text{X's attribute } A_n \text{ is } F_n \rightarrow \Pi_{A_n(X)} = F_n, \tag{3.32}$$

and if also the attributes $A_1, \ldots, A_n$ are noninteractive, then

$$q \triangleq \text{X's attribute } A_1 \text{ is } F_1 \text{ and}, \ldots, \text{and}$$

$$\text{X's attribute } A_n \text{ is } F_n \rightarrow \Pi_{(A_1(X), \ldots, A_n(X))} = F_1 \times \cdots \times F_n$$

and

$$\pi_{(A_1(X), \ldots, A_n(X))}(u_1, \ldots, u_n) = \mu_{F_1}(u_1) \wedge \cdots \wedge \mu_{F_n}(u_n) = \bigwedge_i \mu_{F_i}(u_i), \tag{3.33}$$

where $F_1, \ldots, F_n$ are the fuzzy subsets of $U_1, \ldots, U_n$ with membership functions $\mu_{F_1}(u_1), \ldots, \mu_{F_n}(u_n)$, respectively, $\Pi_{(A_1(X), \ldots, A_n(X))}$ is an n-ary possibility distribution of an n-ary variable $(A_1(X), \ldots, A_n(X))$ in $U = U_1 \times \cdots \times U_n$ with its membership function $\pi_{(A_1(X), \ldots, A_n(X))}$ and $F_1 \times \cdots \times F_n$ is a cartesian product of n unary relations of $F_1, \ldots, F_n$.[4/]

[4] The discussion of the rationale for identifying noninteraction with set intersection can be seen in [1, p. 31].

It should be noted that $F_1 \times \cdots \times F_n$ may be expressed equivalently as

$$F_1 \times \cdots \times F_n = \bar{F}_1 \cap \cdots \cap \bar{F}_n = \bigcap_i \bar{F}_i \tag{3.34}$$

where $\bar{F}_1,\ldots,\bar{F}_n$ are the cylindrical extensions[5/] of $F_1,\ldots,F_n$, respectively.

For example, the proposition,

q = Paul is young and fat

conveys the possibility distribution

$$\Pi_{(Age(Paul), Weight(Paul))} = young \times fat,$$

where young and fat are the fuzzy subsets of the integers from 0 to 150 and the real numbers from 0 to 400, respectively.

We can obtain the following statement concerning a noninteractive conjunction.

Proposition 3.4. The degree of discrimination by a noninteractive conjunctive proposition is the same as the highest degree of discrimination attained by one of the simple propositions which compose the noninteractive conjunctive proposition.

It follows from the theorem below:

Theorem 3.2. Let $\Pi_{A_1(X)},\ldots,\Pi_{A_n(X)}$ be convex possibility distributions in $U_1,\ldots,U_n$, respectively, and be noninteractive, and let $\Pi_{A_1(X')},\ldots,\Pi_{A_n(X')}$ be the same. Then,

$$DS(\Pi_{(A_1(X),\ldots,A_n(X))}, \Pi_{(A_1(X'),\ldots,A_n(X'))})$$
$$= \max \{DS(\Pi_{A_1(X)}, \Pi_{A_1(X')}),\ldots DS(\Pi_{A_n(X)}, \Pi_{A_n(X')})\}. \tag{3.35}$$

---

[5] If F_i is a fuzzy subset of U_i, then its cylindrical extension in $U_1 \times \cdots \times U_n$ is a fuzzy subset of $U_1 \times \cdots \times U_n$ defined by $\bar{F}_i = \int_{U_1\times\cdots\times U_n} \mu_{F_i}(u_i)/(u_1,\ldots,u_n)$.

<u>Lemma</u>. If the possibility distributions $\Pi_{A_1(X)},\ldots,\Pi_{A_n(X)}$ in $U_1,\ldots,U_n$, respectively, be convex; then a cartesian product of $\Pi_{A_1(X)},\ldots,\Pi_{A_n(X)}$, i.e., $\Pi_{A_1(X)} \times \cdots \times \Pi_{A_n(X)}$ in $U = U_1 \times \cdots \times U_n$ is convex too.

<u>Proof of lemma</u>. If

$$\Pi_{A_1(X)} = \int_{U_1} \pi_{A_1}(u_1)/u_1 \tag{3.36}$$

$$\vdots$$

$$\Pi_{A_n(X)} = \int_{U_n} \pi_{A_n}(u_n)/u_n \tag{3.37}$$

where the integral $\int_{U_i}$ denotes the union of fuzzy singleton $\pi_{A_i}(u_i)/u_i$ over the universe of discourse U_i. Then,

$$\Pi_{A_1(X)} \times \cdots \times \Pi_{A_n(X)} = \int_{U_1\times\cdots\times U_n} \pi_{A_1}(u_1) \wedge \cdots \wedge \pi_{A_n}(u_n)/(u_1,\ldots,u_n). \tag{3.38}$$

Let any two fixed points in U be $u = (u_1,\ldots,u_n)$ and $u' = (u_1',\ldots,u_n')$ and let $\lambda \in [0, 1]$. Then the grade of $\Pi_{A_1(X)} \times \cdots \times \Pi_{A_n(X)}$ at the point $\lambda u + (1 - \lambda)u'$ is

$$\pi_{A_1}(\lambda u_1 + (1 - \lambda)u_1') \wedge \cdots \wedge \pi_{A_n}(\lambda u_n + (1 - \lambda)u_n'). \tag{3.39}$$

Since $\Pi_{A_1(X)},\ldots,\Pi_{A_n(X)}$ are convex, we obtain

$$\pi_{A_1}(\lambda u_1 + (1 - \lambda)u_1') \geqq \pi_{A_1}(u_1) \wedge \pi_{A_1}(u_1') \tag{3.40}$$

$$\vdots$$

$$\pi_{A_n}(\lambda u_n + (1 - \lambda)u_n') \geqq \pi_{A_n}(u_n) \wedge \pi_{A_n}(u_n'). \tag{3.41}$$

Thus,

$$\pi_{A_1}(\lambda u_1 + (1-\lambda)u_1') \wedge \cdots \wedge \pi_{A_n}(\lambda u_n + (1-\lambda)u_n')$$
$$\geq (\pi_{A_1}(u_1) \wedge \pi_{A_1}(u_1')) \wedge \cdots \wedge (\pi_{A_n}(u_n) \wedge \pi_{A_n}(u_n'))$$
$$= (\pi_{A_1}(u_1) \wedge \cdots \wedge \pi_{A_n}(u_n)) \wedge (\pi_{A_1}(u_1') \wedge \cdots \wedge \pi_{A_n}(u_n')). \quad (3.42)$$

Q.E.D.

Proof of Theorem 3.2. If

$$\Pi_{A_1(X)} = \int_{U_1} \pi_{A_1}(u_1)/u_1 \quad (3.43)$$
$$\vdots$$
$$\Pi_{A_n(X)} = \int_{U_n} \pi_{A_n}(u_n)/u_n \quad (3.44)$$

$$\Pi_{A_1(X')} = \int_{U_1} \pi'_{A_1}(u_1)/u_1 \quad (3.43')$$
$$\vdots$$
$$\Pi_{A_n(X')} = \int_{U_n} \pi'_{A_n}(u_n)/u_n \quad (3.44')$$

Since (3.43),...,(3.44) are noninteractive, as are (3.43'),...,(3.44'), we obtain

$$\Pi_{(A_1(X),\ldots,A_n(X))} = \Pi_{A_1(X)} \times \cdots \times \Pi_{A_n(X)}$$
$$= \int_{U_1\times\cdots\times U_n} \pi_{A_1}(u_1) \wedge \cdots \wedge \pi_{A_n}(u_n)/(u_1,\ldots,u_n), \quad (3.45)$$

$$\Pi_{(A_1(X'),\ldots,A_n(X'))} = \Pi_{A_1(X')} \times \cdots \times \Pi_{A_n(X')}$$
$$= \int_{U_1\times\cdots\times U_n} \pi'_{A_1}(u_1) \wedge \cdots \wedge \pi'_{A_n}(u_n)/(u_1,\ldots,u_n). \quad (3.45')$$

Now, the maximal grade of $\Pi_{A_1(X)} \cap \Pi_{A_1(X')}$, i.e., $\bar{M}_{\Pi_{A_1(X)},\Pi_{A_1(X')}}$, is

$$\bar{M}_{\Pi_{A_1(X)},\Pi_{A_1(X')}} = \operatorname*{Sup}_{U_1} [\pi_{A_1}(u_1) \wedge \pi'_{A_1}(u_1)]. \tag{3.46}$$

In the same way,

$$\vdots$$

$$\bar{M}_{\Pi_{A_n(X)},\Pi_{A_n(X')}} = \operatorname*{Sup}_{U_n} [\pi_{A_n}(u_n) \wedge \pi'_{A_n}(u_n)]. \tag{3.46'}$$

And the maximal grade of $\Pi_{(A_1(X),\ldots,A_n(X))} \cap \Pi_{(A_1(X'),\ldots,A_n(X'))}$ is

$$\begin{aligned}
&\bar{M}_{\Pi_{(A_1(X),\ldots,A_n(X))},\Pi_{(A_1(X'),\ldots,A_n(X'))}} \\
&= \operatorname*{Sup}_{U_1\times\cdots\times U_n} [(\pi_{A_1}(u_1) \wedge \cdots \wedge \pi_{A_n}(u_n)) \wedge (\pi'_{A_n}(u_n) \wedge \cdots \wedge \pi'_{A_n}(u_n))] \\
&= \operatorname*{Sup}_{U_1\times\cdots\times U_n} [(\pi_{A_1}(u_1) \wedge \pi'_{A_1}(u_1)) \wedge \cdots \wedge (\pi_{A_n}(u_n) \wedge \pi'_{A_n}(u_n))] \\
&= \operatorname*{Sup}_{U_1} \{\ldots\{\operatorname*{Sup}_{U_n} [(\pi_{A_1}(u_1) \wedge \pi'_{A_1}(u_1)) \wedge \cdots \wedge (\pi_{A_n}(u_n) \wedge \pi'_{A_n}(u_n))]\}\ldots\} \\
&= \bar{M}_{\Pi_{A_1(X)},\Pi_{A_1(X')}} \wedge \cdots \wedge \bar{M}_{\Pi_{A_n(X)},\Pi_{A_n(X')}}.
\end{aligned} \tag{3.47}$$

Furthermore, since (3.43),...,(3.44) and (3.43'),...,(3.44') are convex, from the lemma, $\Pi_{(A_1(X),\ldots,A_n(X))}$ and $\Pi_{(A_1(X'),\ldots,A_n(X'))}$ are convex too. Thus, from Theorem 2.1,

$$DS(\Pi_{(A_1(X),\ldots,A_n(X))},\Pi_{(A_1(X'),\ldots,A_n(X'))})$$

$$= 1 - \bar{M}_{\Pi_{(A_1(X),\ldots,A_n(X))},\Pi_{(A_1(X'),\ldots,A_n(X'))}}$$

$$= 1 - (\bar{M}_{\Pi_{A_1(X)},\Pi_{A_1(X')}} \wedge \cdots \wedge \bar{M}_{\Pi_{A_n(X)},\Pi_{A_n(X')}})$$

$$= \max \{DS(\Pi_{A_1(X)},\Pi_{A_1(X')}),\ldots,DS(\Pi_{A_n(X)},\Pi_{A_n(X')})\}. \qquad \text{Q.E.D.}$$

As a simple illustration, if $U_1 = U_2 =$ the space of integers,

$$\Pi_{A_1(X)} = 0.4/-1 + 1/0 + 0.6/1, \tag{3.48}$$

$$\Pi_{A_2(X)} = 0.2/-1 + 1/0 + 0.2/1, \tag{3.49}$$

$$\Pi_{A_1(X')} = 0.6/0 + 1/1 + 0.5/2, \tag{3.48'}$$

$$\Pi_{A_2(X')} = 0.3/0 + 1/1 + 0.4/2, \tag{3.49'}$$

and if also, $\Pi_{A_1(X)}$ and $\Pi_{A_2(X)}$ are noninteractive, as are $\Pi_{A_1(X')}$ and $\Pi_{A_2(X')}$, then,

$$\Pi_{(A_1(X),A_2(X))} = 0.2/(-1,-1) + 0.4/(-1,0) + 0.2/(-1,1) + 0.2/(0,-1) + 1/(0,0) + 0.2/(0,1) + 0.2/(1,-1) + 0.6/(1,0) + 0.2/(1,1) \tag{3.50}$$

$$\Pi_{(A_1(X'),A_2(X'))} = 0.3/(0,0) + 0.6/(0,1) + 0.4/(0,2) + 0.3/(1,0) + 1/(1,1) + 0.4/(1,2) + 0.3/(2,0) + 0.5/(2,1) + 0.4/(2,2). \tag{3.50'}$$

Hence,

$$\bar{M}_{\Pi_{A_1(X)},\Pi_{A_1(X')}} = \text{Sup }[1 \wedge 0.6,\ 0.6 \wedge 1] = 0.6, \tag{3.51}$$

$$\bar{M}_{\Pi_{A_2(X)},\Pi_{A_2(X')}} = \text{Sup }[1 \wedge 0.3,\ 0.2 \wedge 1] = 0.3, \tag{3.51'}$$

$$\bar{M}_{\Pi_{(A_1(X),A_2(X))},\Pi_{(A_1(X'),A_2(X'))}} = \text{Sup }[1 \wedge 0.3,\ 0.2 \wedge 0.6,\ 0.6 \wedge 0.3,\ 0.2 \wedge 1] = 0.3 \tag{3.52}$$

From the premise of convexity,

$$DS(\Pi_{A_1(X)},\Pi_{A_1(X')}) = 1 - 0.6 = 0.4, \tag{3.53}$$

$$DS(\Pi_{A_2(X)},\Pi_{A_2(X')}) = 1 - 0.3 = 0.7, \tag{3.53'}$$

$$DS(\Pi_{(A_1(X),A_2(X))},\Pi_{(A_1(X'),A_2(X'))}) = 1 - 0.3 = 0.7. \tag{3.54}$$

Proposition 3.4 may be explained in the following way.

For example, assume that we have the following propositions concerning Paul and Jeff:

$p_1 \triangleq$ Paul is young, (3.55)

$p_2 \triangleq$ Paul is fat, (3.56)

$p_1' \triangleq$ Jeff is more or less young, (3.55')

$p_2' \triangleq$ Jeff is very thin, (3.56')

If the attributes age and fatness are noninteractive, then it is sufficient to discriminate between them by the propositions including the information about their fatness. Then, it is not necessary to discriminate between them by composing the following conjunctive propositions:

q $\underline{\Delta}$ Paul is young and fat,

q' $\underline{\Delta}$ Jeff is more or less young and very thin.

But when the attributes $A_1,\ldots,A_n$ are not noninteractive, obviously from the process of the proof of Theorem 3.2, we cannot obtain a simple and generalized statement concerning the discriminative effect of a conjunctive composition. But when the attributes $A_1,\ldots,A_n$ have some interactive relation, say, the relation in which increases in the grades of the possibilities in some attributes, e.g., A_i's can be compensated for by decreases in the grades of the possibilities in the other attributes, e.g., A_j's--such as the relation between looks and voice--the possibility distribution $\Pi_{(A_1(X),\ldots,A_n(X))}$ derived from a compensatory conjunctive proposition,

q $\underline{\Delta}$ X's attribute A_1 is F_1 and*...and* attribute A_n is F_n[6/]

may be

$$\Pi_{(A_1(X),\ldots,A_n(X))} = F_1 \cdot \cdots \cdot F_n,$$

$$\pi_{(A_1(X),\ldots,A_n(X))}(u_1,\ldots,u_n) = \mu_{F_1}(u_1) \cdot \cdots \cdot \mu_{F_n}(u_n), \qquad (3.57)$$

where • denotes arithmetic multiplication. In this case, we can obtain the following statement which is somewhat different from Proposition 3.4.

Proposition 3.5. The degree of discrimination by a compensate conjunctive proposition is at least as high as the highest degree of discrimination attained by one of the simple propositions, which compose the composite proposition, provided that the possibility distribution derived from the

---

[6]The conjunction and* denotes any interactive conjunction, to differentiate it from the noninteractive conjunction and.

compensate conjunctive proposition is convex.

Proof. Let the possibility distributions derived from the simple propositions concerning the attributes $A_1,\ldots,A_n$ of objects X and X' be (3.43),...,(3.44) and (3.43'),...,(3.44'), respectively. From (3.57), the possibility distributions derived from the compensate conjunctive propositions concerning X and X' are, respectively,

$$\begin{aligned}\Pi_{(A_1(X),\ldots,A_n(X))} &= \Pi_{A_1(X)} \cdot \cdots \cdot \Pi_{A_n(X)} \\ &= \int_{U_1\times\cdots\times U_n} \pi_{A_1}(u_1) \cdot \cdots \cdot \pi_{A_n}(u_n)/(u_1,\ldots,u_n),\end{aligned} \tag{3.58}$$

$$\begin{aligned}\Pi_{(A_1(X'),\ldots,A_n(X'))} &= \Pi_{A_1(X')} \cdot \cdots \cdot \Pi_{A_n(X')} \\ &= \int_{U_1\times\cdots\times U_n} \pi'_{A_1}(u_1) \cdot \cdots \cdot \pi'_{A_n}(u_n)/(u_1,\ldots,u_n)\end{aligned} \tag{3.58'}$$

The maximal grades of $\Pi_{A_1(X)} \cap \Pi_{A_1(X')},\ldots,\Pi_{A_n(X)} \cap \Pi_{A_n(X')}$ are (3.46),...,(3.46'), respectively. And the maximal grade of $\Pi_{(A_1(X),\ldots,A_n(X))} \cap \Pi_{(A_1(X'),\ldots,A_n(X'))}$ is

$$\begin{aligned}&\bar{M}_{\Pi_{(A_1(X),\ldots,A_n(X))},\Pi_{(A_1(X'),\ldots,A_n(X'))}} \\ &= \sup_{U_1\times\cdots\times U_n} [(\pi_{A_1}(u_1) \cdot \cdots \cdot \pi_{A_n}(u_n)) \wedge (\pi'_{A_1}(u_1) \cdot \cdots \cdot \pi'_{A_n}(u_n))] \\ &= \sup_{U_1} \{\ldots\{\sup_{U_n} [(\pi_{A_1}(u_1) \cdot \cdots \cdot \pi_{A_n}(u_n)) \wedge (\pi'_{A_1}(u_1) \cdot \cdots \cdot \pi'_{A_n}(u_n))]\}\ldots\}.\end{aligned} \tag{3.59}$$

Since $\pi_{A_2}(u_2), \pi'_{A_2}(u_2), \ldots, \pi_{A_n}(u_n), \pi'_{A_n}(u_n) \in [0, 1]$, we obtain

$$\sup_{U_2} \{\ldots\{\sup_{U_n} [(\pi_{A_1}(u_1) \cdot \cdots \cdot \pi_{A_n}(u_n)) \wedge (\pi'_{A_1}(u_1) \cdot \cdots \cdot \pi'_{A_n}(u_n))]\}\ldots\}$$

$$\leqq \pi_{A_1}(u_1) \wedge \pi'_{A_1}(u_1), \qquad \forall u_1 \in U_1. \qquad (3.60)$$

Hence,

$$\bar{M}_{\Pi_{(A_1(X),\ldots,A_n(X))}, \Pi_{(A_1(X'),\ldots,A_n(X'))}} \leqq \bar{M}_{\Pi_{A_1(X)}, \Pi_{A_1'(X')}}. \qquad (3.61)$$

Likewise,

$$\vdots$$

$$\bar{M}_{\Pi_{(A_1(X),\ldots,A_n(X))}, \Pi_{(A_1(X'),\ldots,A_n(X'))}} \leqq \bar{M}_{\Pi_{A_n(X)}, \Pi_{A_n(X')}}. \qquad (3.61')$$

Moreover, if $\Pi_{(A_1(X),\ldots,A_n(X))}$ and $\Pi_{(A_1(X'),\ldots,A_n(X'))}$ are convex, then, from Theorem 2.1,

$$DS(\Pi_{(A_1(X),\ldots,A_n(X))}, \Pi_{(A_1(X'),\ldots,A_n(X'))})$$

$$\geqq DS(\Pi_{A_1(X)}, \Pi_{A_1(X')}), \ldots, DS(\Pi_{A_n(X)}, \Pi_{A_n(X')}). \qquad (3.62)$$

Q.E.D.

Let us use the example described above and the four possibility distributions be (3.48),...,(3.49'). The possibility distributions derived from the compensate conjunctive propositions concerning X and X' are, respectively,

$$\Pi_{(A_1(X),A_2(X))} = 0.08/(-1,1) + 0.4/(-1,0) + 0.08/(-1,1) + 0.2/(0,-1) + 1/(0,0) + 0.2/(0,1) + 0.12/(1,-1) + 0.6/(1,0) + 0.12/(1,1), \quad (3.63)$$

$$\Pi_{(A_1(X'),A_2(X'))} = 0.18/(0,0) + 0.6/(0,1) + 0.24/(0,2) + 0.3/(1,0) + 1/(1,1) + 0.4/(1,2) + 0.15/(2,0) + 0.5/(2,1) + 0.2/(2,2). \quad (3.63')$$

Thus,

$$\bar{M}_{\Pi_{(A_1(X),A_2(X))},\Pi_{(A_1(X'),A_2(X'))}} = \text{Sup}\,[1 \wedge 0.18,\ 0.2 \wedge 0.6,\ 0.6 \wedge 0.3,\ 0.12 \wedge 1] = 0.3 \quad (3.64)$$

From (3.51) and (3.51'), $\bar{M}_{\Pi_{A_1(X)},\Pi_{A_1(X')}}$ and $\bar{M}_{\Pi_{A_2(X)},\Pi_{A_2(X')}}$ are 0.6 and 0.3, respectively. In this case, since $\Pi_{(A_1(X),A_2(X))}$ and $\Pi_{(A_1(X'),A_2(X'))}$ are convex, from Theorem 2.1,

$$DS(\Pi_{(A_1(X),A_2(X))},\Pi_{(A_1(X'),A_2(X'))}) = 1 - 0.3 = 0.7. \quad (3.65)$$

Of course, from (3.53) and (3.53'),

$$DS(\Pi_{A_1(X)},\Pi_{A_1(X')}) = 0.4,$$

$$DS(\Pi_{A_2(X)},\Pi_{A_2(X')}) = 0.7.$$

Finally, when we consider the implication of Proposition 3.4, after getting Proposition 3.5, there is an interesting analogy between that and a basic statement of probability theory. That is, Proposition 3.4 implies that in discrimination, if the attributes A_1 and A_2 are noninteractive, then the discrimination by the proposition containing the information of attribute with high discriminability is sufficient and the other proposition about another attribute becomes useless. It is analogous to that in probability theory, if the events A_1 and A_2 are independent, the information that event A_1 has occurred does not change the estimate of the probability that A_2 will occur, then the information is worthless, although there is a difference in which the former is related to fuzziness and the latter to randomness.

(ii) Disjunctive composition

Disjunction is dual of conjunction. That is, if simple propositions are (3.31),...,(3.32) and the attributes $A_1,\ldots,A_n$ be noninteractive, then

$$q \triangleq \text{X's attribute } A_1 \text{ is } F_1 \text{ or},\ldots,\text{or attribute } A_n \text{ is } F_n \rightarrow \Pi_{(A_1(X),\ldots,A_n(X))} = \bar{F}_1 \cup \cdots \cup \bar{F}_n = \bigcup_i \bar{F}_i,$$

where $\cup$ denotes union and

$$\pi_{(A_1(X),\ldots,A_n(X))}(u_1,\ldots,u_n) = \mu_{F_1}(u_1) \vee \cdots \vee \mu_{F_n}(u_n) = \bigvee_i \mu_{F_i}(u_i). \tag{3.67}$$

We can obtain the following statement concerning a noninteractive disjunction.

Proposition 3.6. The degree of discrimination by a noninteractive disjunctive proposition is at most as high as the lowest degree of discrimination attained by one of the simple propositions, which compose the disjunctive proposition, provided that the possibility distribution derived from the noninteractive disjunctive proposition is convex.

Proof. Let the possibility distributions derived from the simple propositions concerning the attributes $A_1,\ldots,A_n$ of objects X and X' be (3.43),...,(3.44) and (3.43'),...,(3.44'), respectively. From (3.67), the possibility distribution derived from the noninteractive disjunctive propositions concerning X and X' are, respectively,

$$\Pi_{(A_1(X),\ldots,A_n(X))} = \bar{\Pi}_{A_1(X)} \cup \cdots \cup \bar{\Pi}_{A_n(X)} = \int_{U_1\times\cdots\times U_n} \pi_{A_1}(u_1) \vee \cdots \vee \pi_{A_n}(u_n)/(u_1,\ldots,u_n), \tag{3.68}$$

$$\Pi_{(A_1(X'),\ldots,A_n(X'))} = \bar{\Pi}_{A_1(X')} \cup \cdots \cup \bar{\Pi}_{A_n(X')} = \int_{U_1\times\cdots\times U_n} \pi'_{A_1}(u_1) \vee \cdots \vee \pi'_{A_n}(u_n)/(u_1,\ldots,u_n). \tag{3.68'}$$

Now the maximal grades of $\Pi_{A_1(X)} \cap \Pi_{A_1(X')},\ldots,\Pi_{A_n(X)} \cap \Pi_{A_n(X')}$ are given by (3.46),...,(3.46'), respectively. And the maximal grade of $\Pi_{(A_1(X),\ldots,A_n(X))} \cap \Pi_{(A_1(X'),\ldots,A_n(X'))}$ is

$$\begin{aligned} &\bar{M}_{\Pi_{(A_1(X),\ldots,A_n(X))},\Pi_{(A_1(X'),\ldots,A_n(X'))}} \\ &= \sup_{U_1\times\cdots\times U_n} [(\pi_{A_1}(u_1) \vee \cdots \vee \pi_{A_n}(u_n)) \wedge (\pi'_{A_1}(u_1) \vee \cdots \vee \pi'_{A_n}(u_n))] \\ &= \sup_{U_1} \{\ldots\{\sup_{U_n} [(\pi_{A_1}(u_1) \vee \cdots \vee \pi_{A_n}(u_n)) \wedge (\pi'_{A_1}(u_1) \vee \cdots \vee \pi'_{A_n}(u_n))]\}\ldots\}. \end{aligned} \tag{3.69}$$

Since

$$\sup_{U_2}\{\ldots\{\sup_{U_n}[(\pi_{A_1}(u_1)\vee\cdots\vee\pi_{A_n}(u_n))\wedge(\pi'_{A_1}(u_1)\vee\cdots\vee\pi'_{A_n}(u_n))]\}\ldots\}$$

$$\geqq \pi_{A_1}(u_1)\wedge\pi'_{A_1}(u_1), \qquad \forall u_1\in U_1, \qquad (3.70)$$

we can obtain

$$\bar{M}_{\Pi_{(A_1(X),\ldots,A_n(X))},\Pi_{(A_1(X'),\ldots,A_n(X'))}} \geqq \bar{M}_{\Pi_{(A_1(X)},\Pi_{A_1(X')}}. \qquad (3.71)$$

Likewise,

$$\vdots$$

$$\bar{M}_{\Pi_{(A_1(X),\ldots,A_n(X))},\Pi_{(A_1(X'),\ldots,A_n(X'))}} \geqq \bar{M}_{\Pi_{A_n(X)},\Pi_{A_n(X')}}. \qquad (3.71')$$

Moreover, if $\Pi_{(A_1(X),\ldots,A_n(X))}$ and $\Pi_{(A_1(X'),\ldots,A_n(X'))}$ are convex, then, from Theorem 2.1,

$$DS(\Pi_{(A_1(X),\ldots,A_n(X))},\Pi_{(A_1(X'),\ldots,A_n(X'))}$$

$$\leqq DS(\Pi_{A_1(X)},\Pi_{A_1(X')}),\ldots,DS(\Pi_{A_n(X)},\Pi_{A_n(X')}). \qquad (3.72)$$

Q.E.D.

Using the example described above, let the four possibility distributions be (3.48),...,(3.49'). The possibility distributions derived from the noninteractive disjunctive propositions concerning X and X' are, respectively,

$$\Pi_{(A_1(X),A_2(X))} = 0.4/(-1,1) + 1/(-1,0) + 0.4/(-1,1) + 1/(0,-1) + 1/(0,0) + 1/(0,1) + 0.6/(1,-1) + 1/(1,0) + 0.6/(1,1), \quad (3.73)$$

$$\Pi_{(A_1(X'),A_2(X'))} = 0.6/(0,0) + 1/(0,1) + 0.6/(0,2) + 1/(1,0) + 1/(1,1) + 1/(1,2) + 0.5/(2,0) + 1/(2,1) + 0.5/(2,2). \quad (3.73')$$

Hence,

$$\bar{M}_{\Pi_{(A_1(X),A_2(X))},\Pi_{(A_1(X'),A_2(X'))}} = \text{Sup}\,[1 \wedge 0.6,\ 1 \wedge 1,\ 1 \wedge 1,\ 0.6 \wedge 1] = 1.$$

From (3.51) and (3.51'),

$$\bar{M}_{\Pi_{A_1(X)},\Pi_{A_2(X')}} = 0.6$$

and

$$\bar{M}_{\Pi_{A_2(X)},\Pi_{A_2(X')}} = 0.3.$$

In this case, since $\Pi_{(A_1(X),A_2(X))}$ and $\Pi_{(A_1(X'),A_2(X'))}$ are convex, from Theorem 2.1,

$$DS(\Pi_{(A_1(X),A_2(X))},\Pi_{(A_1(X'),A_2(X'))}) = 1 - 1 = 0.$$

Of course,

$$DS(\Pi_{A_1(X)},\Pi_{A_1(X')}) = 0.4$$

and

$$DS(\Pi_{A_2(X)},\Pi_{A_2(X')}) = 0.7.$$

(iii) Weighted sum composition

While we are treating the propositions concerning the attributes of different dimensions, we can also make an interpretation about the weighted sum composition which is different from the one made in 1(iii). Although, as shown in 1(iii), we can interpret a nonnegative number ω_i as the magnitude to which we believe the i-th proposition, it is also possible to interpret ω_i as the magnitude of contribution made by the attribute A_i in discrimination.

In either interpretation, if the simple propositions are (3.31),...,(3.32), then the weighted sum composite proposition becomes

$q \triangleq$ X's attribute A_1 is F_1 with the weight of ω_1 and,...,and X's attribute A_n is F_n with the weight of ω_n

$$\rightarrow \Pi_{(A_1(X),\ldots,A_n(X))} = \omega_1\bar{F}_1 + \cdots + \omega_n\bar{F}_n = \sum_i \omega_i\bar{F}_i,$$

$$\pi_{(A_1(X),\ldots,A_n(X))}(u_1,\ldots,u_n) = \omega_1\mu_{F_1}(u_1) + \cdots + \omega_n\mu_{F_n}(u_n) = \sum_i \omega_i\mu_{F_i}(u_i), \quad (3.74)$$

where + denotes arithmetic addition, and where $\omega_i \in [0, 1]$ and $\sum_i \omega_i = 1$.

For example, if the four possibility distributions are (3.48),...,(3.49') and $\omega_1 = 0.6$, $\omega_2 = 0.4$, then the possibility distributions of the weighted sum composite proposition are, respectively,

$$\Pi_{(A_1(X),A_2(X))} = 0.32/(-1,1) + 0.64/(-1,0) + 0.68/(-1,1) + 0.68/(0,-1) + 1/(0,0) + 0.68/(0,1) + 0.44/(1,-1) + 0.76/(0,1) + 0.44/(1,1), \quad (3.75)$$

$$\Pi_{(A_1(X'),A_2(X'))} = 0.48/(0,0) + 0.76/(0,1) + 0.52/(0,2) + 0.72/(1,0) + 1/(1,1) + 0.76/(1,2) + 0.42/(2,0) + 0.7/(2,1) + 0.46/(2,2). \quad (3.75')$$

Hence,

$$\bar{M}_{\Pi_{(A_1(X),A_2(X))},\Pi_{(A_1(X'),A_2(X'))}}$$

$$= \text{Sup } [1 \wedge 0.48,\ 0.68 \wedge 0.76,\ 0.76 \wedge 0.72,\ 0.44 \wedge 1] \qquad (3.76)$$

$$= 0.72.$$

In this case, since $\Pi_{(A_1(X),A_2(X))}$ and $\Pi_{(A_1(X'),A_2(X'))}$ are convex, from Theorem 2.1,

$$DS(\Pi_{(A_1(X),A_2(X))},\Pi_{(A_1(X'),A_2(X'))}) = 1 - 0.72 = 0.28. \qquad (3.77)$$

From (3.33), (3.67), and (3.74), we can obtain a statement analogous to the one in 1(iii), as below. That is,

$$\bigcap_i F_i \subset \sum_i \omega_i F_i \subset \bigcup_i F_i \qquad (3.78)$$

and

$$\bigcap_i F_i \subset F_i \subset \bigcup_i F_i. \qquad (3.79)$$

Hence, even where these three compositions are made from the propositions concerning the different attributes of different dimensions, we see the same relation as obtained in 1(iii), where we treated the compositions made from the propositions concerning same attributes of same dimensions. This means that the conjunctive propositions are more informative than the weighted sum propositions, and the latter are more informative than the disjunctive propositions. Namely, the magnitude of discriminative effect decreases in the following order: conjunctive propositions, the weighted sum propositions, and disjunctive propositions. Although the range of dimensions treated here are different from the range used in 1(iii), we can recognize the influence of randomness on the discrimination.

(iv) Conditional composition

When n = 2, meaning the number of different attributes are limited to two, we can devise another kind of composition, i.e., conditional composition. That is, if n = 2 in (3.31),...,(3.32) and also if attributes A_1 and A_2 are noninteractive, then

$$q \triangleq \text{If X's attribute } A_1 \text{ is } F_1 \text{ then X's attribute } A_2 \text{ is } F_2$$
$$\rightarrow \Pi_{(A_1(X),A_2(X))} = \bar{F}_1' \oplus \bar{F}_2, \tag{3.80}$$

where F' is the complement of F_1 and $\oplus$ denotes the bounded sum, of which 1 is the maximum, then

$$\pi_{(A_1(X),A_2(X))}(u_1,u_2) = 1 \wedge (1 - \mu_{F_1}(u_1) + \mu_{F_2}(u_2)), \tag{3.81}$$

in which + and - denote the arithmetic addition and subtraction.[7]

As it may easily be guessed, from the processes of the proofs of Theorem 3.2 and Proposition 3.5, we cannot obtain a simple and generalized statement concerning the discriminative effect of a noninteractive conditional composition.

3. Unconditional dimension reduction

The proposition "Roger is big" expresses two attributes, height and weight. But considering the economics of information or the ability to process information, we often reduce that proposition into a proposition which expresses only one attribute, say, "Roger is tall." Now we shall

[7] In the form defined by (3.80), it is consistent with the definition of implication in $L_{\text{aleph-1}}$ logic.

discuss the discriminative effect of this type of dimension reduction of a proposition.

Let us formulate a problem. The proposition $p \triangleq X$ is F contains statements about n attributes $A_1,\ldots,A_n$ with $A_i(X)$ taking values in U_i, $i = 1,\ldots,n$, i.e.,

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

$$\pi_{(A_1(X),\ldots,A_n(X))}(u_1,\ldots,u_n) = \mu_F(u_1,\ldots,u_n). \tag{3.82}$$

When we reduce the proposition above to another proposition which includes the information about k (< n) attributes $A_{i_1},\ldots,A_{i_k}$ of X, say,

$$P_{(k)} \triangleq X\text{'s attributes } A_{i_1},\ldots,A_{i_k} \text{ are } F_{(k)},$$

through some kind of procedure without using any other propositions. What is the discriminative effect of this unconditional dimension reduction of the proposition? That is the problem we are going to discuss here.

The concept of the unconditional dimension reduction has a close relation to the concept of the marginal possibility distribution, which in turn is analogous to the concept of the marginal probability distribution.

We shall explain it in detail. Let $X = (A_1(X),\ldots,A_n(X))$ be an n-ary fuzzy variable taking values in $U = U_1 \times \cdots \times U_n$, and let Π_X $(= \Pi_{(A_1(X),\ldots,A_n(X))})$ be a possibility distribution associated with X $(= (A_1(X),\ldots,A_n(X))$, with $\pi_X(u_1,\ldots,u_n)$ $(= \pi_X(u), u \triangleq (u_1,\ldots,u_n))$ denoting the possibility distribution function of Π_X.

Let $k \triangleq (i_1,\ldots,i_k)$ be a subsequence of the index sequence $(1,\ldots,n)$ and let $X_{(k)}$ be the k-ary fuzzy variable $X_{(k)} \triangleq (A_{i_1}(X),\ldots,A_{i_k}(X))$. The <u>marginal possibility distribution</u> $\Pi_{X_{(k)}}$ $(= \Pi_{(A_{i_1}(X),\ldots,A_{i_k}(X))})$ is a possibility distribution associated with $X_{(k)}$ which is induced by Π_x as the projection (or shadow) of Π_x on $U_{(k)} \triangleq U_{i_1} \times \cdots \times U_{i_k}$. Thus,

$$\Pi_{x_{(k)}} = \Pi_{(A_{i_1}(X),\ldots,A_{i_k}(X))} \triangleq \underset{U_{(k)}}{\text{Proj}}\ \Pi_x, \tag{3.83}$$

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

$$\pi_{X_{(k)}}(u_{(k)}) = \underset{u_{(k')}}{\text{Sup}}\ [\pi_x(u)], \tag{3.84}$$

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

Then, we can consider the possibility distribution which will be derived from the unconditionally reduced k (< n)-ary proposition which is induced by the n-ary original proposition as a k-ary marginal possibilith distribution, $\Pi_{x_{(k)}} = \Pi_{(A_{i_1}(X),\ldots,A_{i_k}(X))}$.

Here we can obtain the following statement concerning an unconditional dimension reduction.

Proposition 3.6. By using the unconditional dimension reduction, the degree of discrimination of the reduced proposition is at most as high as the degree of discrimination by the original proposition. That is, the discriminative effect of unconditional dimension reduction is nonpositive. Both statements are true provided that the possibility distribution derived from the reduced proposition is convex.

Proof. Let the Π_x and $\Pi_{x'}$ be the convex possibility distributions in $U = U_1 \times \cdots \times u_n$ concerning X and X', respectively, and $\pi_x(u)$ and $\pi'_{X'}(u)$ be the possibility distributions of Π_x and $\Pi_{x'}$, respectively. And let $u = (u_{(k)}, u_{(k')})$ and let u, $u_{(k)}$, and $u_{(k')}$ be any point in U, $u_{(k)} = U_{i_1} \times \cdots \times U_{i_k}$ and $U_{(k')} = U_{j_1} \times \cdots \times U_{j_{k'}}$, respectively. From (3.84), the possibility distribution functions of the marginal possibility distributions on $U_{(k)}$ induced by Π_x and $\Pi_{x'}$ are, respectively,

$$\pi_{x_{(k)}}(u_{(k)}) = \operatorname{Sup}_{u_{(k')}} [\pi_x(u)] = \operatorname{Sup}_{u_{(k')}} [\pi_x(u_{(k)}, u_{(k')})] \quad (3.85)$$

$$\pi'_{x'_{(k)}}(u_{(k)}) = \operatorname{Sup}_{u_{(k')}} [\pi'_{x'}(u)] = \operatorname{Sup}_{u_{(k')}} [\pi'_{x'}(u_{(k)}, u_{(k')})]. \quad (3.85')$$

Now, the maximal grade of $\Pi_x \cap \Pi_{x'}$ is

$$\bar{M}_{\Pi_x, \Pi_{x'}} = \operatorname{Sup}_{u} [\pi_x(u) \wedge \pi'_{x'}(u)] = \operatorname{Sup}_{u} [\pi_x(u_{(k)}, u_{(k')}) \wedge \pi'_{x'}(u_{(k)}, u_{(k')})] = \operatorname{Sup}_{u_{(k)}} \{\operatorname{Sup}_{u_{(k')}} [\pi_x(u_{(k)}, u_{(k')}) \wedge \pi'_{x'}(u_{(k)}, u_{(k')})]\}. \quad (3.86)$$

On the other hand, the maximal grade of $\Pi_{x_{(k)}} \cap \Pi_{x'_{(k)}}$ is

$$\bar{M}_{\Pi_{x_{(k)}}, \Pi_{x'_{(k)}}} = \operatorname{Sup}_{u_{(k)}} \{\operatorname{Sup}_{u_{(k')}} [\pi_x(u_{(k)}, u_{(k')})] \wedge \operatorname{Sup}_{u_{(k')}} [\pi'_{x'}(u_{(k)}, u_{(k')})]\}. \quad (3.87)$$

Since,

$$\sup_{u_{(k')}} [\pi_x(u_{(k)},u_{(k')}) \wedge \pi'_{x'}(u_{(k)},u_{(k')})]$$

$$\leqq \sup_{u_{(k')}} [\pi_x(u_{(k)},u_{(k')})] \wedge \sup_{u_{(k')}} [\pi'_{x'}(u_{(k)},u_{(k')})], \qquad \forall\, u_{(k)} \in U_{(k)},$$

we obtain

$$\bar{M}_{\Pi_x,\Pi_{x'}} \leqq \bar{M}_{\Pi_{x_{(k)}},\Pi_{x'_{(k)}}} .$$

Moreover, if $\Pi_{x_{(k)}}$ and $\Pi_{x'_{(k)}}$ are convex, then, from Theorem 2.1,

$$DS(\Pi_x,\Pi_{x'}) \geqq DS(\Pi_{x_{(k)}},\Pi_{x'_{(k)}}). \tag{3.88}$$

Q.E.D.

As a simple illustration, let

$U_1 = U_2 =$ Integers and Π_x and $\Pi_{x'}$ be expressed in the following tables, respectively,

Π_x	$A_1(X)$	$A_2(X)$	π
	-1	0	0.4
	0	-1	0.3
	0	0	1
	0	1	0.4
	1	0	0.6

Table 3.1: Π_x

$\Pi_{x'}$	$A_1(X')$	$A_2(X')$	π
	0	1	0.6
	1	0	0.2
	1	1	1
	1	2	0.6
	2	1	0.5

Table 3.1': $\Pi_{x'}$

From Tables 3.1 and 3.1', the marginal possibility distributions on U_1 concerning X and X' are, respectively,

$$\Pi_{X_1} = \Pi_{A_1(X)} = \underset{U_1}{\text{Proj}}\ \Pi_X = \int_{U_1} \underset{U_2}{\text{Sup}}\ (\pi(u_1,u_2))/u_1$$

$$= 0.4/{-1} + \text{Sup}\ [0.3,\ 1,\ 0.4]/0 + 0.6/1$$

$$= 0.4/{-1} + 1/0 + 0.6/1, \quad (3.89)$$

$$\Pi_{X_1'} = \Pi_{A_1(X')} = 0.6/0 + 1/1 + 0.5/2. \quad (3.89')$$

Hence,

$$\bar{M}_{\Pi_x,\Pi_{x'}} = \text{Sup}\ [0.4 \wedge 0.6,\ 0.6 \wedge 0.2] = 0.4. \quad (3.90)$$

On the otherhand,

$$\bar{M}_{\Pi_{A_1(X)},\Pi_{A_1(X')}} = \text{Sup}\ [1 \wedge 0.6,\ 0.6 \wedge 1] = 0.6. \quad (3.91)$$

Thus,

$$DS(\Pi_x,\Pi_{x'}) = 1 - 0.4 = 0.6, \quad (3.92)$$

and, since $\Pi_{A_1(X)}$ and $\Pi_{A_1(X')}$ are convex,

$$DS(\Pi_{A_1(X)},\Pi_{A_1(X')}) = 1 - 0.6 = 0.4. \quad (3.93)$$

The meaning and the reasoning of Proposition 3.6 is clear. As a simple illustration, in the case of discriminating between Roger and Jack, it is better to do so by propositions such as "Roger is big" and "Jack is not big" than by the propositions which express one attribute, Height, only, such as, "Roger is tall" and "Jack is more or less tall." The reason why this is so is that such unconditional, dimensionally reduced propositions neglect some parts of information which were contained in the original proposition.

Using the concept of the marginal possibility distribution, we can define the concept of noninteractivity of fuzzy variables, which is analogous to the concept of independence of random variables.

Definition 3.2. The fuzzy variables $X_{(k)} = (X_{i_1},\ldots,X_{i_k})$ and $X_{(k')} = (X_{j_1},\ldots,X_{j_{k'}})$ are noninteractive if and only if the possibility distribution associated with $X = (X_1,\ldots,X_n)$ is the cartesian product of the marginal possibility distributions associated with $X_{(k)}$ and $X_{(k')}$, i.e.,

$$\Pi_x = \Pi_{x_{(k)}} \times \Pi_{x_{(k')}}, \tag{3.94}$$

where

$$\Pi_{x_{(k)}} = \operatorname{Proj}_{U_{(k)}} \Pi_x \quad \text{and} \quad \Pi_{x_{(k')}} = \operatorname{Proj}_{U_{(k')}} \Pi_x .$$

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

$$\Pi_x = \Pi_{x_1} \times \cdots \times \Pi_{x_n}, \tag{3.95}$$

where

$$\Pi_{x_i} = \operatorname{Proj}_{U_i} \Pi_x, \qquad i = 1,\ldots,n.$$

Therefore, if the attributes $A_1,\ldots,A_n$ implied are noninteractive and $\operatorname{Proj}_{U_i} \Pi_x$ and $\operatorname{Proj}_{U_i} \Pi_{x'}$, $i = 1,\ldots,n$, are convex, then, from Theorem 3.2 and (3.95),

$$DS(\Pi_x, \Pi_{x'}) = \max\left(DS(\operatorname{Proj}_{U_1} \Pi_x, \operatorname{Proj}_{U_1} \Pi_{x'}), \ldots, DS(\operatorname{Proj}_{U_n} \Pi_x, \operatorname{Proj}_{U_n} \Pi_{x'})\right). \tag{3.96}$$

Thus, we arrive at the following statement:

Proposition 3.7. If the attributes $A_1,\ldots,A_n$ implied are noninteractive, then the degree of discrimination by the proposition containing the information about all the attributes is the same as the maximal grade of discrimination attained by one of the n unconditionally reduced unary propositions derived from the original proposition, provided that the possibility distributions derived from the reduced propositions are convex.

When we consider the implications of Propositions 3.4, 3.5, 3.6, and 3.7 together, we can draw the following conclusion, as shown in Figure 3.2. That is, there is a symmetrical relation between the nonnegative discriminative effect of a conjunctive proposition and the nonpositive discriminative effect of a nonconditional, dimensionally reduced proposition on the axis where the attributes $A_1,\ldots,A_n$ are noninteractive.

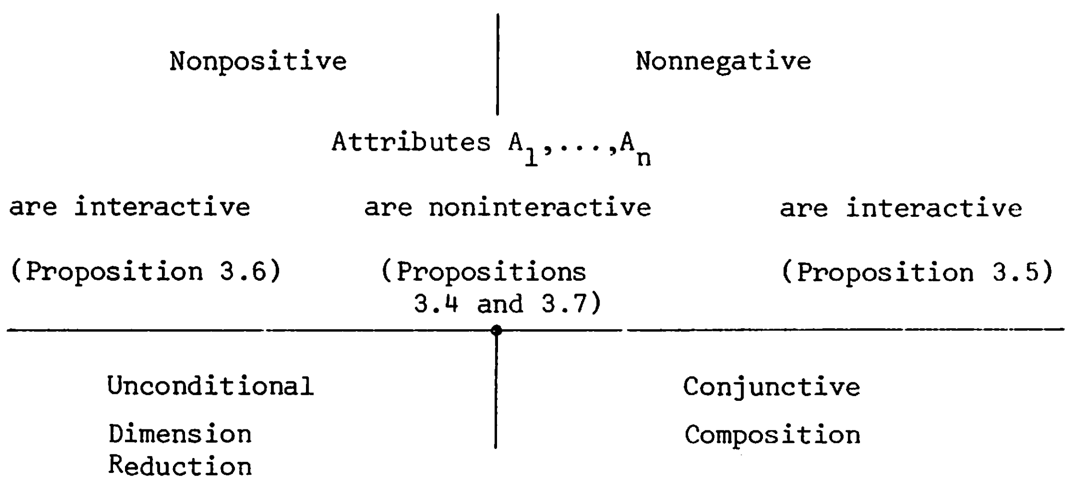

Figure 3.2

4. Conditional dimension reduction

We have discussed the discriminative effect of unconditional dimension reduction in 3. Here, we shall discuss a problem in which we can use some information in dimension reduction. We call this kind of reduction conditional dimension reduction.

More specifically, let the proposition $p \triangleq X$ is F express the information concerning n attributes $A_1,\ldots,A_n$ of X, taking values in $U_1,\ldots,U_n$, respectively, i.e.,

$$X \text{ is } F \rightarrow \Pi_X = \Pi_{(A_1(X),\ldots,A_n(X))} = F,$$
$$\pi_{(A_1(X),\ldots,A_n(X))}(u_1,\ldots,u_n) = \mu_F(u_1,\ldots,u_n). \tag{3.82}$$

And let another proposition $p_{(k')} \triangleq$ X's attributes $A_{j_1},\ldots,A_{j_{k'}}$ are $F_{k'}$ express the information concerning attributes $A_{j_1},\ldots,A_{j_{k'}}$ $(k' < n)$ of X, taking values in $U_{j_1},\ldots,U_{j_{k'}}$, respectively, i.e.,

$$X\text{'s attributes } A_{j_1},\ldots,A_{j_{k'}} \text{ are } F_{k'} \rightarrow \Pi_{X_{(k')}} = \Pi_{(A_{j_1}(X),\ldots,A_{j_{k'}}(X))} = F_{k'},$$

$$\pi_{(A_{j_1}(X),\ldots,A_{j_k}(X))}(u_{j_1},\ldots,u_{j_{k'}}) = \mu_{F_{k'}}(u_{j_1},\ldots,u_{j_{k'}}). \tag{3.97}$$

When we use the proposition $p_{(k')}$ to reduce the original proposition p and obtain a reduced proposition $p_{(k)}$, what is the discriminative effect of this conditional dimension reduction? That is the problem we discuss here.

(i) Comparison of conditional and unconditional dimension reduction

A difference between conditional and unconditional dimension reduction is the availability of the possibility distribution of a conditional proposition, i.e.,

$$\Pi_{(A_{j_1}(X),\ldots,A_{j_k}(X))} = F'_k.$$

Upon reconsidering the matter, the possibility distribution derived from an unconditional, dimensionally reduced proposition is

$$\Pi_{(A_{i_1}(X),\ldots,A_{i_k}(X))} \triangleq \underset{U_{(k)}}{\text{Proj}}\ \Pi_x. \tag{3.83}$$

Therefore, it is consistent that we consider the possibility distribution derived from a conditional, dimensionally reduced proposition, as

$$\Pi_{(A_{i_1}(X),\ldots,A_{i_k}(X))} \triangleq \underset{U_{(k)}}{\text{Proj}}\ (\Pi_x \cap \bar{F}_{k'}). \tag{3.98}$$

Equivalently, the right-hand member of (3.98) may be regarded as the composition of Π_x and $F_{k'}$, i.e.,

$$\underset{U_{(k)}}{\text{Proj}}\ (\Pi_x \cap \bar{F}_k) = \Pi_x \circ F_{k'}. \tag{3.99}$$

Now,

$$\Pi_x \cap \bar{F}_{k'} \subset \Pi_x;$$

therefore,

$$\underset{U_{(k)}}{\text{Proj}}\ (\Pi_x \cap \bar{F}_{k'}) \subset \underset{U_{(k)}}{\text{Proj}}\ \Pi_x. \tag{3.100}$$

Hence, if $\underset{U_{(k)}}{\text{Proj}}\ (\Pi_x \cap \bar{F}_{k'})$ and $\underset{U_{(n)}}{\text{Proj}}\ \Pi_x$ are convex, then, from Theorem 3.1, we can obtain the following statement:

<u>Proposition 3.8</u>. A conditional, dimensionally reduced proposition is at least as informative as an unconditional, dimensionally reduced proposition;

the discriminative effect of the former is at least as high as the one of the latter, provided that the possibility distributions derived from the reduced propositions are convex.

As a simple illustration, let the possibility distributions of X and X' be Table 3.1 and 3.1', respectively. Moreover, let the possibility distributions of the conditional propositions concerning the attribute A_2 of X and X' be, respectively,

$$\Pi_{A_2(X)} = 0.2/-1 + 1/0 + 0.6/1, \tag{3.101}$$

$$\Pi_{A_2(X')} = 0.1/0 + 1/1 + 0.4/2. \tag{3.101'}$$

From (3.101) and (3.101'), we can obtain the cylindrical extensions of $\Pi_{A_2(X)}$ and $\Pi_{A_2(X')}$ as follows:

$A_2(X)$ \ $A_1(X)$	-1	0	1
-1	0.2	0.2	0.2
0	1	1	1
1	0.6	0.6	0.6

Table 3.2: $\bar{\Pi}_{A_2}(X)$

$A_2(X')$ \ $A_1(X')$	0	1	2
0	0.1	0.1	0.1
1	1	1	1
2	0.4	0.4	0.4

Table 3.2': $\bar{\Pi}_{A_2}(X)$

where the value of each element of these matrices expresses the grade of possibility of each combination of the variables. Tables 3.1 and 3.1' are transformed into the following matrices:

$A_2(X)$ \ $A_1(X)$	-1	0	1
-1	0	0.3	0
0	0.4	1	0.6
1	0	0.4	0

Table 3.3: Π_X

$A_2(X')$ \ $A_1(X')$	0	1	2
0	0	0.2	0
1	0.6	1	0.5
2	0	0.6	0

Table 3.3': $\Pi_{X'}$

$\Pi_X \cap \bar{\Pi}_{A_2(X)}$ and $\Pi_{X'} \cap \bar{\Pi}_{A_2(X')}$ are obtained from the combinations of Tables 3.2 and 3.3 and Tables 3.2' and 3.3', respectively, as follows:

$A_2(X)$ \ $A_1(X)$	-1	0	1
-1	0	0.2	0
0	0.4	1	0.6
1	0	0.4	0

Table 3.4: $\Pi_X \cap \bar{\Pi}_{A_2}(X)$

$A_2(X')$ \ $A_1(X')$	0	1	2
0	0	0.1	0
1	0.6	1	0.5
2	0	0.4	0

Table 3.4': $\Pi_{X'} \cap \bar{\Pi}_{A_2}(X')$

From Tables 3.4 and 3.4',

$$\underset{U_1}{\text{Proj}}\,(\Pi_X \cap \bar{\Pi}_{A_2}(X)) = 0.4/1 + 1/0 + 0.6/1, \tag{3.102}$$

$$\underset{U_1}{\text{Proj}}\,(\Pi_{X'} \cap \bar{\Pi}_{A_2}(X')) = 0.6/0 + 1/1 + 0.5/2. \tag{3.103}$$

Proposition 3.8 stands to reason. Because a conditional, dimensionally reduced proposition has used more information than an unconditional, dimensionally reduced proposition, then the former becomes more informative and more discriminative.

(ii) Comparison of conditional dimension reductions

From (3.98), which is the definition-equation for the possibility distribution of the conditional, dimensionally reduced proposition, we can obtain the following statement:

Proposition 3.9. In conditional dimension reduction, if a conditional proposition $q_{(k')}$ is at least as informative as the other conditional proposition $p_{(k')}$, then the conditional, dimensionally reduced proposition using $q_{(k')}$ is at least as informative as the other one using $p_{(k')}$; thus, the discriminative effect of the former is at least as high as the one of the latter, provided that these possibility distributions are convex.

Proof. Let $F_{k'}$ and $G_{k'}$ be the possibility distributions derived from the conditional propositions $p_{(k')}$ and $q_{(k')}$, respectively. Since

$$I(q_{(k')}) \geqq I(p_{(k')}),$$

we obtain

$$G_{k'} \subset F_{k'}.$$

From (3.98),

$$\operatorname*{Proj}_{U_{(k)}} (\Pi_X \cap \bar{G}_{k'}) \subset \operatorname*{Proj}_{U_{(k)}} (\Pi_X \cap \bar{F}_{k'}). \tag{3.104}$$

Q.E.D.

The implication of the proposition is self-evident.

5. Qualified proposition

In daily life, we often get the propositions such as "It is more or less true that X is a small integer" and "It is likely that X is a small integer." In natural language, there are many qualifiers, such as more or less and likely, which are attached to the propositions, e.g., "X is a small integer," and act on fuzzy constraints and modify the fuzzy set F to F^+.

There are three types of qualifiers:

I. τ, where τ is a linguistic truth-value, e.g., true, very true, more or less true, false, etc.

II. λ, where λ is a linguistic probability-value (or likelihood), e.g., likely, very likely, very unlikely, etc.

III. π, where π is a linguistic possibility-value, e.g., possible, quite possible, slightly possible, impossible, etc.

Here we shall discuss the discriminative effect of the truth-qualifier τ, because the treatment of τ provides the basis for the others, λ and π.[8]

In general, the proposition with the truth-qualifier τ may be stated as follows:

$$p^* \triangleq X \text{ is } F \text{ is } \tau, \qquad (3.105)$$

where τ is a fuzzy subset of interval [0, 1] which expresses the degree of truth. For example, true may be expressed as in Figure 3.3.

---

[8] A more detailed discussion about qualifiers may be found in [1], [7], and [8].

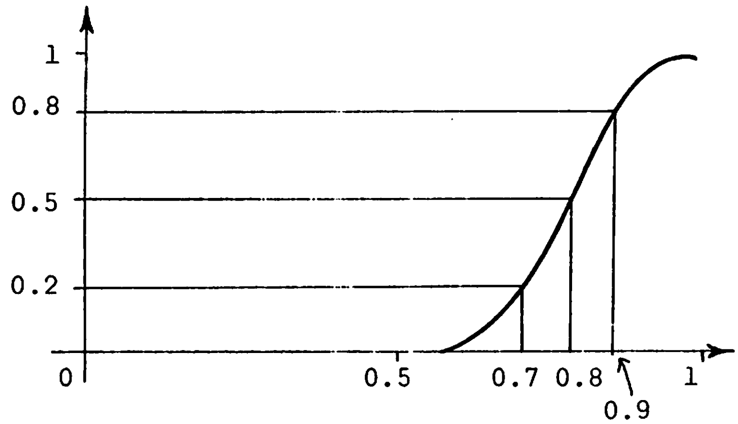

Figure 3.3: Fuzzy set; true.

A proposition with a truth-qualifier τ, p*, can be transformed into a proposition without τ, q, by using the extension principle.[9/] In this case, the possibility distribution derived from the proposition q becomes the composition of the binary relation in the space [0, 1] × U, μ_F^{-1}, and the unary relation (or fuzzy subset) on [0, 1], τ, i.e., if

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

then

$$p^* \triangleq X \text{ is } F \text{ is } \tau \rightarrow \Pi_{A(X)} = F^+, \qquad (3.106)$$

where $F^+ = \mu_F^{-1} \circ \tau$, in which μ_F^{-1} is the inverse of the membership function of F and ∘ is the operation of composition.

---

[9]Let g be a mapping from U to V. Thus,

$$v = g(u),$$

where u and v are generic elements of U and V, respectively.

Let F be a fuzzy subset of U expressed as

$$F = \int_U \mu_F(u)/u.$$

By the extension principle, the image of F under g is given by

$$g(F) = \int_U \mu_F(u)/g(u).$$

Cf. Bellman, R. E., and L. A. Zadeh [1].

For example, let the proposition with a truth-qualifier be p* $\triangleq$ It is very true that X is a small integer, the fuzzy subset of small integer be (3.23) and the truth-qualifier true, which is a fuzzy set too, be defined by

true = 0.2/0.7 + 0.5/0.8 + 0.8/0.9 + 1/1.

From (3.13),

very true = 0.04/0.7 + 0.25/0.8 + 0.6/0.9 + 1/1.

On the other hand, we can interpret $\mu^{-1}_{\text{small integer}}$ as a binary relation in the space of [0, 1] × Integer, and express it as follows:

$\mu^{-1}_{\text{small integer}}$

[0,1] \ Integer	1	2	3	4	5	6
0.2	0	0	0	0	0	1
0.5	0	0	0	0	1	0
0.8	0	0	0	1	0	0
1	1	1	1	0	0	0

Table 3.5: $\mu^{-1}_{\text{small integer}}$

Then, using Kaufmann's expression,[10] we can calculate $\mu^{-1}_{\text{small integer}} \circ$ very true as follows:

[10] Cf. Kaufmann, A. [4].

0.7	0.8	0.1	1
0.04	0.25	0.6	1

∘

$\mu^{-1}_{s.i.}$	1	2	3	4	5	6
0.2	0	0	0	0	0	1
0.5	0	0	0	0	1	0
0.8	0	0	0	1	0	0
1	1	1	1	0	0	0

=

1	2	3
$(0.25 \wedge 0) \vee (1 \wedge 1)$	$(0.25 \wedge 0) \vee (1 \wedge 1)$	$(0.25 \wedge 0) \vee (1 \wedge 1)$

4	5	6
$(0.25 \wedge 1) \vee (1 \wedge 0)$	$(0.25 \wedge 0) \vee (1 \wedge 0)$	$(0.25 \wedge 0) \vee (1 \wedge 0)$

=

1	2	3	4	5	6
1	1	1	0.25	0	0

. (3.107)

Roughly speaking, this result implies that the proposition with the truth-qualifier, "It is very true that X is a small integer" is transformed into the proposition without the qualifier, "X is a very, very, and very small integer."

Now, if two propositions with the truth-qualifiers are

$$p^* \triangleq X \text{ is } F \text{ is } \tau$$

and

$$q^* \triangleq X \text{ is } F \text{ is } \tau^*,$$

and, if also,

$$\tau^* \subset \tau, \qquad (3.108)$$

then, from (3.99) and (3.106),

$$\mu_F^{-1} \circ \tau^* \subset \mu_F^{-1} \circ \tau. \qquad (3.109)$$

Thus we can obtain the following statement:

Proposition 3.10. In two propositions with truth-qualifiers, τ and τ^*, if $\tau^* \subset \tau$, then the proposition with τ^* is at least as informative as the other with τ; and the magnitude of discriminability of the former with τ^* is at least as large as the magnitude of discriminability of the latter with τ, provided that the possibility distributions derived from the propositions are convex.

The implication of the proposition is self-evident. For example, in the case of discrimination between Roger and Jack, it is clearer to discriminate between them by the propositions, "It is very true that Roger is big" and "It is very true that Jack is not big" than by the following, "It is more or less true that Roger is big" and "It is true that Jack is not big."

CONCLUDING REMARKS

In decision-making, discrimination of fuzzy elements is as important as dealing with randomness. The importance of the former will be understood when one realizes that identification, e.g., pattern-recognition, diagnosis, etc., can be interpreted as a special kind of discrimination, in which we discriminate, not between two fuzzy objects, but between a fuzzy object and a conceptualized pattern or image.

But the study of discrimination of fuzzy elements has not yet been fully explored owing mainly to the fact that it has been less than thirteen years since L. A. Zadeh initiated the serious study of fuzziness in 1965.

As a step toward approaching that problem, this paper presented a measure for discrimination of fuzzy elements, which was called discriminative effect of information, through the concept of degree of separation. Then, several basic ways of information-processing are discussed from the point of view of the discriminative effect of information.

The writer is well aware that this is a very small step into the vast areas which are as yet unknown. I would like to express my heart-felt gratitude to Professor L. A. Zadeh, without whose encouragement for the development of this idea, this paper could not have been done.

ACKNOWLEDGMENT

Research sponsored by the U.S. Army Research Office Grant DAHC04-75-G-0056 and the National Science Foundation Grant MCS77-07568.

Yuji Enta is presently a visiting scholar from the Department of Business Administration, Hosei University, 2-17-1, Fujimi Chiyoda-ku, Tokyo, Japan.

REFERENCES

1. Bellman, R. E., and L. A. Zadeh, "Local and fuzzy logics," Memorandum No. ERL-M584, University of California, Berkeley, 1976.
2. De Luca, A., and S. Termini, "A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory," Inf. Control 20 (1972), pp. 301-312.
3. Enta, Y., "Fuzzy choice models," Keiei Shirin 13(4) (1977), pp. 79-91 (in Japanese).
4. Kaufmann, A., Introduction to the theory of fuzzy subsets (Vol. 1, Fundamental theoretical elements), Academic Press, New York (1975). (Translated by D. L. Swanson)
5. Lakoff, G., "Hedges: A study in meaning criteria and the logic of fuzzy concepts," J. of Phil. Logic 2 (1973), pp. 458-508.
6. Zadeh, L. A., "Fuzzy sets," Inf. and Control 8 (1965), pp. 338-353.
7. Zadeh, L. A., "A fuzzy-set-theoretic interpretation of linguistic hedges," J. of Cybernetics 2 (1972), pp. 4-34.
8. Zadeh, L. A., "PRUF - A meaning representation language for natural languages," Memorandum No. ERL-M77/61, University of California, Berkeley, 1977.
9. Zadeh, L. A., "Fuzzy sets as a basis for a theory of possibility," Fuzzy Sets and Systems 1 (1978), pp. 3-28.