Similarity Relations and Fuzzy Orderings

L. A. ZADEH
Department of Electrical Engineering and Computer Sciences and Electronics Research Laboratory, University of California, Berkeley, California

ABSTRACT

The notion of "similarity" as defined in this paper is essentially a generalization of the notion of equivalence. In the same vein, a fuzzy ordering is a generalization of the concept of ordering. For example, the relation $x \succ y$ ($x$ is much larger than $y$) is a fuzzy linear ordering in the set of real numbers.

More concretely, a similarity relation, $S$, is a fuzzy relation which is reflexive, symmetric, and transitive. Thus, let $x, y$ be elements of a set $X$ and $\mu_S(x, y)$ denote the grade of membership of the ordered pair $(x, y)$ in $S$. Then $S$ is a similarity relation in $X$ if and only if, for all $x, y, z$ in $X$, $\mu_S(x, x) = 1$ (reflexivity), $\mu_S(x, y) = \mu_S(y, x)$ (symmetry), and $\mu_S(x, z) \geq \vee (\mu_S(x, y) \land \mu_S(y, z))$ (transitivity), where $\vee$ and $\land$ denote max and min, respectively.

A fuzzy ordering is a fuzzy relation which is transitive. In particular, a fuzzy partial ordering, $P$, is a fuzzy ordering which is reflexive and antisymmetric, that is, $(\mu_S(x, y) > 0$ and $x \neq y) \Rightarrow \mu_S(x, y) = 0$. A fuzzy linear ordering is a fuzzy partial ordering in which $x \neq y \Rightarrow \mu_S(x, y) > 0$ or $\mu_S(y, x) > 0$. A fuzzy preordering is a fuzzy ordering which is reflexive. A fuzzy weak ordering is a fuzzy preordering in which $x \neq y \Rightarrow \mu_S(x, y) > 0$ or $\mu_S(y, x) > 0$.

Various properties of similarity relations and fuzzy orderings are investigated and, as an illustration, an extended version of Szpilrajn's theorem is proved.

1. INTRODUCTION

The concepts of equivalence, similarity, partial ordering, and linear ordering play basic roles in many fields of pure and applied science. The classical theory of relations has much to say about equivalence relations and various types of orderings [1]. The notion of a distance, $d(x, y)$, between objects $x$ and $y$ has long been used in many contexts as a measure of similarity or dissimilarity between elements of a set. Numerical taxonomy [2], factor

†This work was supported in part by a grant from the National Science Foundation, NSF GK-10656X, to the Electronics Research Laboratory, University of California, Berkeley, California.

Information Sciences 3 (1971), 177–200

analysis [3], pattern classification [4-7], and analysis of proximities [8-10] provide a number of concepts and techniques for categorization and clustering. Preference orderings have been the object of extensive study in econometrics and other fields [11, 12]. Thus, in sum, there exists a wide variety of techniques for dealing with problems involving equivalence, similarity, clustering, preference patterns, etc. Furthermore, many of these techniques are quite effective in dealing with the particular classes of problems which motivated their development.

The present paper is not intended to add still another technique to the vast armamentarium which is already available. Rather, its purpose is to introduce a unifying point of view based on the theory of fuzzy sets [13] and, more particularly, fuzzy relations. This is accomplished by extending the notions of equivalence relation and ordering to fuzzy sets, thereby making it possible to adapt the well-developed theory of relations to situations in which the classes involved do not have sharply defined boundaries. Thus, the main contribution of our approach consists in providing a unified conceptual framework for the study of fuzzy equivalence relations and fuzzy orderings, thereby facilitating the derivation of known results in various applied areas and, possibly, stimulating the discovery of new ones.

In what follows, our attention will be focused primarily on defining some of the basic notions within this conceptual framework and exploring some of their elementary implications. Although our approach might be of use in areas such as cluster analysis, pattern recognition, decision processes, taxonomy, artificial intelligence, linguistics, information retrieval, system modeling, and approximation, we shall make no attempt in the present paper to discuss its possible applications in these or related problem areas.

2. NOTATION, TERMINOLOGY, AND PRELIMINARY DEFINITIONS

In [13], a fuzzy (binary) relation $R$ was defined as a fuzzy collection of ordered pairs. Thus, if $X = \{x\}$ and $Y = \{y\}$ are collections of objects denoted generically by $x$ and $y$, then a fuzzy relation from $X$ to $Y$ or, equivalently, a fuzzy relation in $X \cup Y$, is a fuzzy subset of $X \times Y$ characterized by a membership (characteristic) function $\mu_R$ which associates with each pair $(x, y)$ its "grade of membership," $\mu_R(x, y)$, in $R$. We shall assume for simplicity that the range of $\mu_R$ is the interval $[0, 1]$ and will refer to the number $\mu_R(x, y)$ as the strength of the relation between $x$ and $y$.

\footnote{In an independent work which came to this writer's attention [14], S. Tamura, S. Higuchi, and K. Tanaka have applied fuzzy relations to pattern classification, obtaining some of the results described in Section 3.}
In the following definitions, the symbols \( \lor \) and \( \land \) stand for \( \max \) and \( \min \), respectively.

The **domain** of a fuzzy relation \( R \) is denoted by \( \text{dom} \, R \) and is a fuzzy set defined by

\[
\mu_{\text{dom} \, R}(x) = \lor_y \mu_R(x, y), \quad x \in X,
\]

where the supremum, \( \lor \), is taken over all \( y \) in \( Y \). Similarly, the **range** of \( R \) is denoted by \( \text{ran} \, R \) and is defined by

\[
\mu_{\text{ran} \, R}(y) = \lor_x \mu_R(x, y), \quad x \in X, \, y \in Y.
\]

The **height** of \( R \) is denoted by \( h(R) \) and is defined by

\[
h(R) = \lor_x \lor_y \mu_R(x, y).
\]

A fuzzy relation is **subnormal** if \( h(R) < 1 \) and **normal** if \( h(R) = 1 \).

The **support** of \( R \) is denoted by \( \text{S}(R) \) and is defined to be the non-fuzzy subset of \( X \times Y \) over which \( \mu_R(x, y) > 0 \).

The **union** of \( R \) and \( Q \) is denoted by \( R \cup Q \) (rather than \( R + Q \)) and is defined by \( \mu_{R \cup Q} = \lor_x \lor_y \mu_R(x, y) \mu_Q(x, y) \), that is

\[
\mu_{R \cup Q}(x, y) = \max(\mu_R(x, y), \mu_Q(x, y)), \quad x \in X, \, y \in Y.
\]

Consistent with this notation, if \( \{R_\alpha\} \) is a family of fuzzy (or non-fuzzy) sets, we shall write \( \bigcup_\alpha R_\alpha \) to denote the union \( \bigcup_\alpha R_\alpha \).

The **intersection** of \( R \) and \( Q \) is denoted by \( R \cap Q \) and is defined by \( \mu_{R \cap Q} = \mu_R \land \mu_Q \).

The **product** of \( R \) and \( Q \) is denoted by \( RQ \) and is defined by \( \mu_{RQ} = \mu_R \mu_Q \).

Note that, if \( R, Q, \) and \( T \) are any fuzzy relations from \( X \) to \( Y \), then

\[
R(Q + T) = RQ + RT.
\]

The **complement** of \( R \) is denoted by \( R' \) and is defined by \( \mu_{R'} = 1 - \mu_R \).

If \( R \subset X \times Y \) and \( Q \subset Y \times Z \), then the composition, or, more specifically, the **max-min composition**, of \( R \) and \( Q \) is denoted by \( R \circ Q \) and is defined by

\[
\mu_{R \circ Q}(x, z) = \lor_y (\mu_R(x, y) \land \mu_Q(y, z)), \quad x \in X, \, z \in Z.
\]

The \( n \)-fold composition \( R \circ R \ldots \circ R \) is denoted by \( R^n \).

From the above definitions of the composition, union, and containment it follows at once that, for any fuzzy relations \( R \subset X \times Y \), \( Q \subset Y \times Z \), and \( S \subset Z \times W \), we have

\[
R \circ (Q \circ S) = (R \circ Q) \circ S,
\]

and
\[ R \circ (Q + T) = R \circ Q + R \circ T, \quad (7) \]

Note. On occasion it may be desirable to employ an operation \( * \) other than \( \wedge \) in the definition of the composition of fuzzy relations. Then (5) becomes
\[ R \circ Q = \bigvee_{y} (\mu_{R}(x, y) \ast \mu_{Q}(y, z)), \quad (9) \]
with \( R \ast Q \) called the max-star composition of \( R \) and \( Q \).

In order that (6), (7), and (8) remain valid when \( \wedge \) is replaced by \( * \), it is sufficient that \( * \) be associative and monotone non-decreasing in each of its arguments, which assures the distributivity of \( * \) over \(+\). A simple example of an operation satisfying these conditions and having the interval \([0, 1]\) as its range is the product. In this case, the definition of the composition assumes the form
\[ \mu_{R \ast Q}(x, z) = \bigvee_{y} (\mu_{R}(x, y) \cdot \mu_{Q}(y, z)), \quad (10) \]
where we use the symbol \( \cdot \) in place of \( \wedge \) to differentiate between the max-min and max-product compositions. In what follows, in order to avoid a confusing multiplicity of definitions, we shall be using (5) for the most part as our definition of the composition, with the understanding that, in all but a few cases, an assertion which is established with (5) as the definition of the composition holds true also for (10) and, more generally, (9) (provided (6), (7), and (8) are satisfied).

Note also that, when \( X \) and \( Y \) are finite sets, \( \mu_{R} \) may be represented by a relation matrix whose \((x, y)\)th element is \( \mu_{R}(x, y) \). In this case, the defining equation (5) implies that the relation matrix for the composition of \( R \) and \( Q \) is given by the max-min product\(^{2}\) of the relation matrices for \( R \) and \( Q \).

Level Sets and the Resolution Identity

For \( \alpha \) in \([0, 1]\), an \( \alpha \)-level-set of a fuzzy relation \( R \) is denoted by \( R_{\alpha} \) and is a non-fuzzy set in \( X \times Y \) defined by
\[ R_{\alpha} = \{(x, y) | \mu_{R}(x, y) \geq \alpha \}. \quad (11) \]
Thus, the \( R_{\alpha} \) form a nested sequence of non-fuzzy relations, with
\[ \alpha_{1} \geq \alpha_{2} \Rightarrow R_{\alpha_{1}} \subseteq R_{\alpha_{2}}. \quad (12) \]
\(^{2}\) An exhaustive discussion of operations having properties of this type can be found in [15].

\(^{3}\) In the max-min (or quasi-Boolean) product of matrices with real-valued elements, \( \wedge \) and \( \vee \) play the roles of product and addition, respectively [16, 17].
An immediate and yet important consequence of the definition of a level set is stated in the following proposition:

**Proposition 1.** Any fuzzy relation from $X$ to $Y$ admits of the resolution

$$R = \sum_\alpha \alpha R_\alpha, \quad 0 < \alpha < 1,$$

(13)

where $\Sigma$ stands for the union (see (4)) and $\alpha R_\alpha$ denotes a subnormal non-fuzzy set defined by

$$\mu_{\alpha R_\alpha}(x, y) = \alpha \mu_{R_\alpha}(x, y), \quad (x, y) \in X \times Y.$$  

(14)

or equivalently

$$\mu_{\alpha R_\alpha}(x, y) = \alpha, \quad \text{for} \ (x, y) \in R_\alpha,$$

$$= 0, \quad \text{elsewhere}.$$  

Proof. Let $\mu_{R_\alpha}(x, y)$ denote the membership function of the non-fuzzy set $R_\alpha$ in $X \times Y$ defined by (11). Then (11) implies that

$$\mu_{R_\alpha}(x, y) = 1, \quad \text{for} \ \mu_R(x, y) \geq \alpha,$$

$$= 0, \quad \text{for} \ \mu_R(x, y) < \alpha,$$

and consequently the membership function of $\Sigma_\alpha \alpha R_\alpha$ may be written as

$$\mu_{\Sigma_\alpha \alpha R_\alpha}(x, y) = \bigvee_\alpha \alpha \mu_{R_\alpha}(x, y)$$

$$= \bigvee_{\alpha \leq \mu_R(x, y)} \alpha$$

$$= \mu_R(x, y),$$

which in turn implies (13).

Note. It is understood that in (13) to each $R_\alpha$ corresponds a unique $\alpha$. If this is not the case, e.g., $\alpha_1 \neq \alpha_2$ and $R_{\alpha_1} = R_{\alpha_2}$, then the two terms are combined by forming their union, yielding $(\alpha_1 \vee \alpha_2) R_\alpha$. In this way, a summation of the form (13) may be converted into one in which to each $R_\alpha$ corresponds a unique $\alpha$. Furthermore, if $X$ and $Y$ are finite sets and the distinct entries in the relation matrix of $R$ are denoted by $\alpha_k$, $k = 1, 2, \ldots, K$, where $K$ is a finite number, then (13) assumes the form

$$R = \sum_k \alpha_k R_{\alpha_k}, \quad 1 \leq k \leq K.$$  

(16)

As a simple illustration of (13), assume $X = Y = \{x_1, x_2, x_3\}$, with the relation matrix $\mu_R$ given by

$$\mu_R = \begin{bmatrix}
1 & 0.8 & 0 \\
0.6 & 1 & 0.9 \\
0.8 & 0 & 1
\end{bmatrix}.$$
In this case, the resolution of $R$ reads

$$R = 0.6((x_1, x_1), (x_1, x_2), (x_2, x_2), (x_2, x_3), (x_3, x_1), (x_3, x_3)) + 0.8((x_1, x_1), (x_1, x_2), (x_2, x_2), (x_2, x_3), (x_3, x_1), (x_3, x_3)) + 0.9((x_1, x_1), (x_2, x_2), (x_2, x_3), (x_3, x_1)) + 1((x_1, x_1), (x_2, x_2), (x_3, x_3)). \quad (17)$$

In what follows, we assume that $X = Y$. Furthermore, we shall assume for simplicity that $X$ is a finite set, $X = \{x_1, x_2, \ldots, x_n\}$.

3. SIMILARITY RELATIONS

The concept of a similarity relation is essentially a generalization of the concept of an equivalence relation. More specifically:

**Definition.** A similarity relation, $S$, in $X$ is a fuzzy relation in $X$ which is

(a) reflexive, i.e.,

$$\mu_S(x, x) = 1, \quad \text{for all } x \in \text{dom } S, \quad (18)$$

(b) symmetric i.e.,

$$\mu_S(x, y) = \mu_S(y, x), \quad \text{for all } x, y \in \text{dom } S, \quad (19)$$

and (c) transitive, i.e.,

$$S \circ S \subseteq S, \quad (20)$$

or, more explicitly,

$$\mu_S(x, z) \geq \bigvee_y (\mu_S(x, y) \wedge \mu_S(y, z)).$$

**Note.** If $*$ is employed in place of $\circ$ in the definition of the composition, the corresponding definition of transitivity becomes

$$S \circ S \subseteq S \ast S$$

or, more explicitly

$$\mu_S(x, z) \geq \bigvee_y (\mu_S(x, y) * \mu_S(y, z)).$$

When there is a need to distinguish between the transitivity defined by (20) and the more general form defined by (21), we shall refer to them as *max-min* and *max-star* transitivity, respectively.

An example of the relation matrix of a similarity relation $S$ is shown in Figure 1. It is readily verified that $S = S \circ S$ and also that $S = S \ast S$. 
Transitivity

There are several aspects of the transitivity of a similarity relation which are in need of discussion. First, note that, in consequence of (18), we have

\[ S \supset S^2 \supset S^k, \quad k = 3, 4, \ldots \]  

and hence

\[ S \supset S^2 \Leftrightarrow S = S', \]  

where

\[ S = S + S^2 + S^3 + \ldots \]  

is the transitive closure of \( S \). Thus, as in the case of equivalence relations, the condition that \( S \) be transitive is equivalent to

\[ S = S = S + S^2 + S^3 + \ldots \]  

An immediate consequence of (25) is that the transitive closure of any fuzzy relation is transitive. Note also that for any \( S \)

\[ S = S^2 \Rightarrow S = S' \]

and, if \( S \) is reflexive, then

\[ S = S^2 \Leftrightarrow S = S'. \]

The significance of (25) is made clearer by the following observation. Let \( x_1, \ldots, x_k \) be \( k \) points in \( X \) such that \( \mu(x_{i1}, x_{i2}), \ldots, \mu(x_{i_k-1}, x_{i_k}) \) are all \( > 0 \). Then the sequence \( C = (x_{i1}, \ldots, x_{i_k}) \) will be said to be a chain from \( x_{i1} \) to \( x_{i_k} \), with the strength of this chain defined as the strength of its weakest link, that is,

\[ \text{strength of } (x_{i1}, \ldots, x_{i_k}) = \mu(x_{i1}, x_{i2}) \land \ldots \land \mu(x_{i_{k-1}}, x_{i_k}). \]

From the definition of the composition (equation (5)), it follows that the \((i,j)\)th element of \( S^l, \ l = 1, 2, 3, \ldots, \) is the strength of the strongest chain of length \( l \) from \( x_i \) to \( x_j \). Thus, the transitivity condition (25) may be stated in words as: for all \( x_i, x_j \) in \( X' \),

\[ \text{strength of } S \text{ between } x_i \text{ and } x_j = \text{strength of the strongest chain from } x_i \text{ to } x_j. \]
Second, if \( X \) has \( n \) elements, then any chain \( C \) of length \( k \gg n + 1 \) from \( x_1 \) to \( x_k \) must necessarily have cycles, that is, one or more elements of \( X \) must occur more than once in the chain \( C = (x_1, \ldots, x_k) \). If these cycles are removed, the resulting chain, \( C' \), of length \( \leq n \), will have at least the same strength as \( C \), by virtue of (26). Consequently, for any elements \( x_i, x_j \) in \( X \) we can assert that

\[
\text{strength of the strongest chain from } x_i \text{ to } x_j = \text{strength of the strongest chain of length } \leq n \text{ from } x_i \text{ to } x_j. \tag{28}
\]

Since the \((i,j)\)th element of \( S \) is the strength of the strongest chain from \( x_i \) to \( x_n \), (28) implies the following proposition [18], which is well known for Boolean matrices [16]:

**Proposition 2.** If \( S \) is a fuzzy relation characterized by a relation matrix of order \( n \), then

\[
S = S + S^2 + S^3 + \ldots = S + S^2 + \ldots + S^n. \tag{29}
\]

**Note.** Observe that (29) remains valid when in the definition of the composition and the strength of a chain \( \wedge \) is replaced by the product, i.e., \( S^k \), \( k = 2, 3, \ldots \) is replaced by the \( k \)-fold composition \( S \cdot S \cdot \ldots \cdot S \), with \( \cdot \) defined by (10), and (26) is replaced by

\[
\text{strength of } (x_{i_1}, \ldots, x_{i_k}) = \mu(x_{i_1}, x_{i_2}) \mu(x_{i_2}, x_{i_3}) \ldots \mu(x_{i_{k-1}}, x_{i_k}). \tag{30}
\]

Since \( ab \leq a \wedge b \) for \( a, b \in [0, 1] \), it follows that

\[
S \triangleright S \circ S \implies S \triangleright S \circ S, \tag{31}
\]

that is, max-min transitivity implies max-product transitivity. This observation is useful in situations in which the strength of a chain is more naturally expressed by (30) than by (26).

A case in point is provided by the criticisms [19–21] leveled at the assumption of transitivity in the case of weak ordering. Thus, suppose that \( X \) is a finite interval \([a, b]\) and that we wish to define a non-fuzzy preference ordering on \( X \) in terms of two relations \( > \) and \( \approx \) such that

(a) for every \( x, y \) in \( X \), exactly one of \( x > y, y > x, \) or \( x \approx y \) is true,
(b) \( \approx \) is an equivalence relation,
(c) \( > \) is transitive.

In many cases, it would be reasonable to assume that

\[
x \approx y \iff |x - y| < \varepsilon > 0,
\]

where \( \varepsilon \) is a small number (in relation to \( b - a \)) representing an "indifference"
interval. But then, by transitivity of $\approx$, $x \approx y$ for all $x, y$ in $X$, which is inconsistent with our intuitive expectation that when the difference between $x$ and $y$ is sufficiently large, either $x > y$ or $y > x$ must hold.

This difficulty is not resolved by making $\approx$ a similarity relation in $X$ so long as we employ the max-min transitivity in the definition of $\approx$. For, if we make the reasonable assumption that $\mu_\approx(x, y)$ is continuous at $x = y$, then (20) implies that $\mu_\approx(x, y) = 1$ for all $x, y$ in $X$.

The difficulty may be resolved by making $\approx$ a similarity relation and employing the max-product transitivity in its definition. As an illustration, suppose that

$$\mu_\approx(x, y) = e^{-\beta|x-y|}, \quad x, y \in X,$$

where $\beta$ is any positive number. In this case, $\approx$ may be interpreted as "is not much different from."

Let $x, y, z \in [a, b]$, with $x < z$. Then, substituting (32) in

$$\mu_\approx(z, x) = \vee_y \mu_\approx(x, y) \mu_\approx(y, z),$$

we have

$$\mu_\approx(z, x) = \vee_y e^{-\beta|x-y|} e^{-\beta|z-y|}$$

$$= \vee_{y \in [x, z]} e^{-\beta|x-y|} e^{-\beta|x-z|}$$

$$= e^{-\beta|x-z|}$$

$$= \mu_\approx(x, z),$$

which establishes that $\approx^2 = \approx$ and hence that (32) defines a similarity relation which is continuous at $x = y$ and yet is not constant over $X$.

Finally, it should be noted that the transitivity condition (20) implies and is implied by the ultrameric inequality [22] for distance functions. Specifically, let the complement of a similarity relation $S$ be a dissimilarity relation $D$, with

$$\mu_D(x, y) = 1 - \mu_S(x, y), \quad x, y \in X.$$  

(34)

If $\mu_D(x, y)$ is interpreted as a distance function, $d(x, y)$, then (20) yields

$$1 - d(x, z) \geq \vee_y ((1 - d(x, y)) \wedge (1 - d(y, z))),$$

and since

$$(1 - d(x, y)) \wedge (1 - d(y, z)) = 1 - (d(x, y) \vee (d(y, z)))$$

(35)

we can conclude that, for all $x, y, z$ in $X$,

$$d(x, z) < d(x, y) \vee d(y, z), \quad y \in X,$$

(36)

*Information Sciences 3 (1971), 177-200*
which is the ultrameric inequality satisfied by \(d(x, y)\). Clearly, (36) implies the triangle inequality

\[d(x, z) \leq d(x, y) + d(y, z).\]

Thus, (20) implies (36) and (37), and is implied by (36).

Returning to our discussion of similarity relations, we note that one of their basic properties is an immediate consequence of the resolution identity (13) for fuzzy relations. Specifically,

**Proposition 3.** Let

\[S = \sum_{\alpha} \alpha S_\alpha, \quad 0 < \alpha < 1,\]

be the resolution of a similarity relation in \(X\). Then each \(S_\alpha\) in (38) is an equivalence relation in \(X\). Conversely, if the \(S_\alpha\), \(0 < \alpha < 1\), are a nested sequence of distinct equivalence relations in \(X\), with \(\alpha_1 > \alpha_2 \Rightarrow S_{\alpha_2} \subseteq S_{\alpha_1}\), \(S_1\) non-empty and \(\text{dom} S_\alpha = \text{dom} S_1\), then, for any choice of \(\alpha\)'s in \((0, 1]\) which includes \(\alpha = 1\), \(S\) is a similarity relation in \(X\).

**Proof.** First, since \(\mu_\alpha(x, x) = 1\) for all \(x\) in the domain of \(S\), it follows that \((x, x) \in S_\alpha\) for all \(\alpha\) in \((0, 1]\) and hence that \(S_\alpha\) is reflexive for all \(\alpha\) in \((0, 1]\).

Second, for each \(\alpha\) in \((0, 1]\), let \((x, y) \in S_\alpha\), which implies that \(\mu_\alpha(x, y) > \alpha\) and hence, by symmetry of \(S\), that \(\mu_\alpha(y, x) > \alpha\). Consequently, \((y, x) \in S_\alpha\) and thus \(S_\alpha\) is symmetric for each \(\alpha\) in \((0, 1]\).

Third, for each \(\alpha\) in \((0, 1]\), suppose that \((x_1, x_2) \in S_\alpha\) and \((x_2, x_3) \in S_\alpha\). Then \(\mu_\alpha(x_1, x_2) > \alpha\) and \(\mu_\alpha(x_2, x_3) > \alpha\) and hence, by the transitivity of \(S\), \(\mu_\alpha(x_1, x_3) > \alpha\). This implies that \((x_1, x_3) \in S_\alpha\) and hence that \(S_\alpha\) is transitive for each \(\alpha\) in \((0, 1]\).

First, since \(S_1\) is non-empty, \((x, x) \in S_1\) and hence \(\mu_\alpha(x, x) = 1\) for all \(x\) in the domain of \(S_1\).

Second, expressed in terms of the membership functions of \(S\) and \(S_\alpha\), (38) reads

\[\mu_\alpha(x, y) = \bigvee \alpha \mu_\alpha(x, y), \quad x, y \in \text{dom} S.\]

It is obvious from this expression for \(\mu_\alpha(x, y)\) that the symmetry of \(S_\alpha\) for each \(\alpha\) in \((0, 1]\) implies the symmetry of \(S\).

Third, let \(x_1, x_2, x_3\) be some arbitrarily chosen elements of \(X\). Suppose that

\[\mu_\alpha(x_1, x_2) = \alpha \quad \text{and} \quad \mu_\alpha(x_2, x_3) = \beta.\]

Then, \((x_1, x_2) \in S_{\alpha, \beta}\) and \((x_2, x_3) \in S_{\alpha, \beta}\), and consequently \((x_1, x_3) \in S_{\alpha, \beta}\) by the transitivity of \(S_{\alpha, \beta}\).

From this it follows that, for all \(x_1, x_2, x_3\) in \(X\), we have

\[\mu_\alpha(x_1, x_3) > \alpha \land \beta\]
and hence

\[ \mu_\alpha(x_1, x_3) \geq \bigvee_{x_2} (\mu_\alpha(x_1, x_2) \land \mu_\alpha(x_2, x_3)) , \]

which establishes the transitivity of \( S \).

**Partition Tree**

Let \( \pi_\alpha \) denote the partition induced on \( X \) by \( S_\alpha \), \( 0 < \alpha < 1 \). Clearly, \( \pi_\alpha \) is a refinement of \( \pi_\beta \) if \( \alpha \geq \beta \). For, by the definition of \( \pi_\alpha \), two elements of \( X \), say \( x \) and \( y \), are in the same block of \( \pi_\alpha \) iff \( \mu_\alpha(x, y) \geq \alpha \). This implies that \( \mu_\alpha(x, y) \geq \alpha \) and hence that \( x \) and \( y \) are in the same block of \( \pi_\alpha \).

![Partition Tree](image)

**Figure 2.** Partition tree for the similarity relation defined in Figure 1.

A nested sequence of partitions \( \pi_{\alpha_1}, \pi_{\alpha_2}, \ldots, \pi_{\alpha_n} \) may be represented diagrammatically in the form of a *partition tree*, as shown in Figure 2. It should be noted that the concept of a partition tree plays the same role with respect to a similarity relation as the concept of a quotient does with respect to an equivalence relation.

The partition tree of a similarity relation \( S \) is related to the relation matrix of \( S \) by the rule: \( x_i \) and \( x_j \) belong to the same block of \( \pi_\alpha \) iff \( \mu_\alpha(x_i, x_j) \geq \alpha \).

---

4 The notion of a partition tree and its properties are closely related to the concept of the hierarchic clustering scheme described in [22].
This rule implies that, given a partition tree of $S$, one can readily determine $\mu_S(x_i, x_j)$ by observing that

$$\mu_S(x_i, x_j) = \text{largest value of } \alpha \text{ for which } x_i \text{ and } x_j \text{ are in the same block of } \pi_\alpha.$$  \hspace{1cm} (39)

An alternative to the diagrammatic representation of a partition tree is provided by a slightly modified form of the phrase-marker notation which is commonly used in linguistics [23]. Specifically, if we allow recursion and use the notation $\alpha(A,B)$ to represent a partition $\pi_\alpha$ whose blocks are $A$ and $B$, then the partition tree shown in Figure 2 may be expressed in the form of a string:

$$0.2(0.6(0.8(1(x_1, x_3)), 0.8(1(x_4), 1(x_6))), 0.6(0.8(1(x_2), 1(x_5)))).$$ \hspace{1cm} (40)

This string signifies that the highest partition, $\pi_1$, comprises the blocks $(x_1, x_3), (x_4), (x_6), (x_2), \text{ and } (x_5)$. The next partition, $\pi_{0.8}$, comprises the blocks $((x_1, x_3)), ((x_4), (x_6)), \text{ and } ((x_2), (x_5))$. And so on. Needless to say, the profusion of parentheses in the phrase-marker representation of a partition tree makes it difficult to visualize the structure of a similarity relation from an inspection of (40).

**Similarity Classes**

Similarity classes play the same role with respect to a similarity relation as do equivalence classes in the case of an equivalence relation. Specifically, let $S$ be a similarity relation in $X = \{x_1, \ldots, x_n\}$ characterized by a membership function $\mu_S(x_i, x_j)$. With each $x_i \in X$, we associate a *similarity class* denoted by $S[x_i]$ or simply $[x_i]$. This class is a fuzzy set in $X$ which is characterized by the membership function

$$\mu_{S[x_i]}(x_j) = \mu_S(x_i, x_j).$$ \hspace{1cm} (41)

Thus, $S[x_i]$ is identical with $S$ conditioned on $x_i$, that is, with $x_i$ held constant in the membership function of $S$.

To illustrate, the similarity classes associated with $x_1$ and $x_2$ in the case of the similarity relation defined in Figure 1 are

$$S[x_1] = \{(x_1, 1), (x_2, 0.2), (x_3, 1), (x_4, 0.6), (x_5, 0.2), (x_6, 0.6)\}$$

$$S[x_2] = \{(x_1, 0.2), (x_2, 1), (x_3, 0.2), (x_4, 0.2), (x_5, 0.8), (x_6, 0.2)\}.$$  

By conditioning both sides of the resolution (38) on $x_i$ we obtain at once the following proposition:

**Proposition 4.** The similarity class of $x_i, x_i \in X$, admits of the resolution

$$S[x_i] = \sum_\alpha \alpha S_\alpha[x_i],$$ \hspace{1cm} (42)
where $S_\alpha[x_1]$ denotes the block of $S_\alpha$ which contains $x_1$, and $\alpha S_\alpha[x_1]$ is a subnormal non-fuzzy set whose membership function is equal to $\alpha$ on $S_\alpha[x_1]$ and vanishes elsewhere.

For example, in the case of $S[x_1]$, with $S$ defined in Figure 1, we have
\[ S[x_1] = 0.2(x_1, x_2, x_3, x_4, x_5, x_6) + 0.6(x_1, x_3, x_4, x_6) + 1\{x_1, x_3\} \]
and similarly
\[ S[x_2] = 0.2(x_1, x_2, x_3, x_4, x_5, x_6) + 0.8(x_2, x_5) + 1(x_2). \]

The similarity classes of a similarity relation are not, in general, disjoint—as they are in the case of an equivalence relation. Thus, the counterpart of disjointness is a more general property which is asserted in the following proposition:

**Proposition 5.** Let $S[x_i]$ and $S[x_j]$ be arbitrary similarity classes of $S$. Then, the height (see (3)) of the intersection of $S[x_i]$ and $S[x_j]$ is bounded from above by $\mu_S(x_i, x_j)$, that is,
\[ h(S[x_i] \cap S[x_j]) \leq \mu_S(x_i, x_j). \]  

**Proof.** By definition of $h$ we have
\[ h(S[x_i] \cap S[x_j]) = \vee_{x_k} (\mu_S(x_i, x_k) \land \mu_S(x_j, x_k)), \]
which in view of the symmetry of $S$ may be rewritten as
\[ h(S[x_i] \cap S[x_j]) = \vee_{x_k} (\mu_S(x_i, x_k) \land \mu_S(x_k, x_j)). \]

Now the right-hand member of (44) is identical with the grade of membership of $(x_i, x_j)$ in the composition of $S$ with $S$. Thus
\[ h(S[x_i] \cap S[x_j]) = \mu_{S \circ S}(x_i, x_j), \]
which, in virtue of the transitivity of $S$, implies that
\[ h(S[x_i] \cap S[x_j]) \leq \mu_S(x_i, x_j). \]  

Note that, if $S$ is reflexive, then $S^2 = S$ and (45) is satisfied with the equality sign. Thus, for the example of Proposition 4, we have
\[ h(S[x_1] \cap S[x_2]) = 0.2 = \mu_S(x_1, x_2), \]
since $S$ is reflexive.

The following corollary follows at once from Proposition 5:

**Corollary 6.** The height of the intersection of all similarity classes of $X$ is bounded by the infimum of $\mu_S(x_i, x_j)$ over $X$. Thus
\[ h(S[x_1] \cap \ldots \cap S[x_n]) \leq \land_{x_i, x_j} \mu_S(x_i, x_j). \]  

We turn next to the consideration of fuzzy ordering relations.

*Information Sciences* 3 (1971), 177–200
4. FUZZY ORDERINGS

A fuzzy ordering is a fuzzy transitive relation. In what follows we shall define several basic types of fuzzy orderings and dwell briefly upon some of their properties.

A fuzzy relation $P$ in $X$ is a fuzzy partial ordering iff it is reflexive, transitive, and antisymmetric. By antisymmetry of $P$ is meant that

$$\mu_P(x, y) > 0 \quad \text{and} \quad \mu_P(y, x) > 0 \Rightarrow x = y, \quad x, y \in X. \quad (47)$$

(On occasion, we may use the notation $x \leq y$ to signify that $\mu_p(x, y) > 0$.)

$$\mu_P = \begin{bmatrix}
1 & 0.8 & 0.2 & 0.6 & 0.6 & 0.4 \\
0 & 1 & 0 & 0 & 0.6 & 0 \\
0 & 0 & 1 & 0 & 0.5 & 0 \\
0 & 0 & 0 & 1 & 0.6 & 0.4 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}$$

FIGURE 3. Relation matrix for a fuzzy partial ordering.

An example of a relation matrix for a fuzzy partial ordering is shown in Figure 3. The corresponding fuzzy Hasse diagram for this ordering is shown in Figure 4. In this diagram, the number associated with the arc joining $x_i$ to $x_j$ is $\mu_P(x_i, x_j)$, with the understanding that $x_j$ is a cover for $x_i$, that is, there is no $x_k$ in $X$ such that $\mu_P(x_i, x_k) > 0$ and $\mu_P(x_k, x_j) > 0$. Note that the numbers associated with the arcs define the relation matrix by virtue of the transitivity identity $P = P^2$.

5 Alternatively, a fuzzy ordering may be viewed as a metrized ordering in which the metric satisfies the ultrametric inequality.
As in the case of a similarity relation, a fuzzy partial ordering may be resolved into non-fuzzy partial orderings. This basic property of fuzzy partial orderings is expressed by

**Proposition 7.** Let

\[ P = \sum_\alpha \alpha P_\alpha, \quad 0 < \alpha < 1, \]

be the resolution of a fuzzy partial ordering in \( X \). Then each \( P_\alpha \) in (48) is a partial ordering in \( X \). Conversely, if the \( P_\alpha \), \( 0 < \alpha < 1 \), are a nested sequence of distinct partial orderings in \( X \), with \( \alpha_1 > \alpha_2 \Rightarrow P_{\alpha_1} \subset P_{\alpha_2} \), \( P_1 \) non-empty, and \( \text{dom} P_\alpha = \text{dom} P_1 \), then, for any choice of \( \alpha \)'s in \( (0,1) \) which includes \( \alpha = 1 \), \( P \) is a fuzzy partial ordering in \( X \).

Proof. Reflexivity and transitivity are established as in Proposition 3. As for antisymmetry, suppose that \((x,y) \in P_\alpha \) and \((y,x) \in P_\alpha \). Then \( \mu_\alpha(x,y) = \alpha \), \( \mu_\alpha(y,x) = \alpha \) and hence by antisymmetry of \( P_\alpha \), \( x = y \). Conversely, suppose that \( \mu_\alpha(x,y) = \alpha > 0 \) and \( \mu_\beta(y,x) = \beta > 0 \). Let \( \gamma = \alpha \land \beta \). Then \((x,y) \in P_\gamma \) and \((y,x) \in P_\gamma \), and from the antisymmetry of \( P_\gamma \) it follows that \( x = y \).

In many applications of the concept of a fuzzy partial ordering, the condition of reflexivity is not a natural one to impose. If we allow \( \mu_\alpha(x,x), x \in X \), to take any value in \((0,1]\), the ordering will be referred to as irreflexive.

To illustrate the point, assume that \( X \) is an interval \([a,b]\), and \( \mu_\alpha(x,y) = f(y-x) \), with \( f(y-x) = 0 \) for \( y < x \) and \( f(0) = 1 \). Then, as was noted in Section 2 ((31) et seq.), if \( f(x) \) is right-continuous at \( x = 0 \), the max-min transitivity of \( \mu_\alpha \) requires that \( f(x) = 1 \) for \( x > 0 \). However, if we drop the requirement of reflexivity, then it is sufficient that \( f \) be monotone non-decreasing in order to satisfy the condition of transitivity. For, assume that \( f \) is monotone non-decreasing and \( x < y < z, x, y, z, \in [a,b] \). Then

\[ \mu_\alpha(x,z) = f(z-x) = f((z-y) + (y-x)), \]

and, since

\[ f((z-y) + (y-x)) \geq f(z-y), \]

\[ f((z-y) + (y-x)) \geq f(y-x), \]

we have

\[ f((z-y) + (y-x)) \geq f(z-y) \land f(y-x), \]

and therefore

\[ \mu_\alpha(x,z) \geq \bigvee_y (\mu_\alpha(x,y) \land \mu_\alpha(y,z)), \]

which establishes the transitivity of \( P \).

*Information Sciences 3 (1971), 177–200*
It should be noted that the condition is not necessary. For example, it is easy to verify that for any
\[
\frac{1}{b-a} \leq \beta \leq \frac{2}{b-a},
\]
the function
\[
f(x) = \beta x, \quad 0 \leq x \leq 1/\beta,
\]
\[
e = 2 - \beta x, \quad 1/\beta \leq x \leq b-a,
\]
corresponds to a transitive fuzzy partial ordering if \(\beta(b-a) \leq \frac{b}{a}\).

With each \(x \in X\), we associate two fuzzy sets: the dominating class, denoted by \(P_{\geq} [x_i]\) and defined by
\[
\mu_{P_{\geq} [x_i]}(x_j) = \mu_P(x_j, x_i), \quad x_j \in X,
\]
and the dominated class, denoted by \(P_{\leq} [x_i]\) and defined by
\[
\mu_{P_{\leq} [x_i]}(x_j) = \mu_P(x_j, x_i), \quad x_j \in X.
\]

In terms of these classes, \(x_i\) is undominated iff
\[
\mu_P(x_i, x_j) = 0, \quad \text{for all } x_j \neq x_i,
\]
and \(x_i\) is undominating iff
\[
\mu_P(x_j, x_i) = 0, \quad \text{for all } x_j \neq x_i.
\]

It is evident that, if \(P\) is any fuzzy partial ordering in \(X = \{x_1, \ldots, x_n\}\), the sets of undominated and undominating elements of \(X\) are non-empty.

Another related concept is that of a fuzzy upper-bound for a non-fuzzy subset of \(X\). Specifically, let \(A\) be a non-fuzzy subset of \(X\). Then the upper-bound for \(A\) is a fuzzy set denoted by \(U(A)\) and defined by
\[
U(A) - \bigcap_{x_i \in A} P_{\geq} [x_i].
\]

For a non-fuzzy partial ordering, this reduces to the conventional definition of an upper-bound. Note that, if the least element of \(U(A)\) is defined as an \(x_i\) (if it exists) such that
\[
\mu_{U(A)}(x_i) > 0 \quad \text{and} \quad \mu_P(x_i, x_j) > 0 \quad \text{for all } x_j \text{ in the support of } U(A),
\]
then the least upper-bound of \(A\) is the least element of \(U(A)\) and is unique by virtue of the antisymmetry of \(P\).

In a similar vein, one can readily generalize to fuzzy orderings many of the well-known concepts relating to other types of non-fuzzy orderings. Some of these are briefly stated in the sequel.
Preordering

A fuzzy preordering $R$ is a fuzzy relation in $X$ which is reflexive and transitive. As in the case of a fuzzy partial ordering, $R$ admits of the resolution

$$R = \sum_\alpha \alpha R\alpha, \quad 0 < \alpha \leq 1,$$

(56)

where the $\alpha$-level-sets $R\alpha$ are non-fuzzy preorderings.

For each $\alpha$, the non-fuzzy preordering $R\alpha$ induces an equivalence relation, $E\alpha$, in $X$ and a partial ordering, $P\alpha$, on the quotient $X/E\alpha$. Specifically,

$$(x_i, x_j) \in E\alpha \iff \mu_{R\alpha}(x_i, x_j) = \mu_{R\alpha}(x_j, x_i) = 1$$

(57)

and

$$([x_i], [x_j]) \in P\alpha \iff \mu_{R\alpha}(x_i, x_j) = 1 \quad \text{and} \quad \mu_{R\alpha}(x_j, x_i) = 0,$$

(58)

where $[x_i]$ and $[x_j]$ are the equivalence classes of $x_i$ and $x_j$, respectively.

As an illustration, consider the fuzzy preordering characterized by the relation matrix shown in Figure 5. The corresponding relation matrices for $R_{0.2}$, $R_{0.6}$, $R_{0.8}$, $R_{0.9}$, and $R_1$ read as in Figure 6.

\[
\mu_R = \begin{bmatrix}
1 & 0.8 & 1 & 0.8 & 0.8 & 0.8 \\
0.2 & 1 & 0.2 & 0.2 & 0.8 & 0.2 \\
1 & 0.8 & 1 & 0.8 & 0.8 & 0.8 \\
0.6 & 0.9 & 0.6 & 1 & 0.9 & 1 \\
0.2 & 0.8 & 0.2 & 0.2 & 1 & 0.2 \\
0.6 & 0.9 & 0.6 & 0.9 & 0.9 & 1
\end{bmatrix}
\]

Figure 5. Relation matrix of a fuzzy preordering.

The preordering in question may be represented in diagrammatic form as shown in Figure 7. In this figure, the broken lines in each level (identified by $R\alpha$) represent the arcs (edges) of the Hasse diagram of the partial ordering $P\alpha$, rotated clockwise by 90°. The nodes of this diagram are the equivalence classes of the equivalence relation, $E\alpha$, induced by $R\alpha$. Thus, the diagram as a whole is the partition tree of the similarity relation

$$S = 0.2E_{0.2} + 0.6E_{0.6} + 0.8E_{0.8} + 0.9E_{0.9} + 1E_1,$$

with the blocks in each level of the tree forming the elements of a partial ordering $P\alpha$ which is represented by a rotated Hasse diagram.

Linear Ordering

A fuzzy linear ordering $L$ is a fuzzy antisymmetric ordering in $X$ in which for every $x \neq y$ in $X$ either $\mu_L(x, y) > 0$ or $\mu_L(y, x) > 0$. A fuzzy linear ordering

*Information Sciences* 3 (1971), 177–200
$R_{0.2}$
\[
\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 \\
\end{bmatrix}
\]

$R_{0.6}$
\[
\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 \\
\end{bmatrix}
\]

$R_{0.8}$
\[
\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 \\
\end{bmatrix}
\]

$R_{0.9}$
\[
\begin{bmatrix}
1 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 & 1 & 1 \\
\end{bmatrix}
\]

$R_1$
\[
\begin{bmatrix}
1 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
\end{bmatrix}
\]

Figure 6. Relation matrices for the level sets of the preordering defined in Figure 5.

Figure 7. Structure of the preordering defined in Figure 5.
admits of the resolution
\[ L = \sum_\alpha \alpha L_\alpha, \quad 0 < \alpha \leq 1, \tag{59} \]
which is a special case of (48) and in which the \( L_\alpha \) are non-fuzzy linear orderings.

A simple example of an irreflexive fuzzy linear ordering is the relation \( y \succ x \) in \( X = (-\infty, \infty) \). If we define \( \mu_L(x, y) \) by
\[ \mu_L(x, y) = \begin{cases} \frac{1}{1 + (y - x)^2}, & \text{for } y - x > 0, \\ 0, & \text{for } y - x < 0, \end{cases} \]
then \( L \) is transitive (in virtue of (49)), antisymmetric, and \( \mu_L(x, y) > 0 \) or \( \mu_L(y, x) > 0 \) for every \( x \neq y \) in \( (-\infty, \infty) \). Hence \( L \) is a fuzzy linear ordering.

**Weak Ordering**

If we remove the condition of antisymmetry, then a fuzzy linear ordering becomes a *weak ordering*. Equivalently, a weak ordering, \( W \), may be regarded as a special case of a preordering in which for every \( x \neq y \) in \( X \) either \( \mu_W(x, y) > 0 \) or \( \mu_W(y, x) > 0 \).

**Szpilrajn’s Theorem**

A useful example of a well-known result which can readily be extended to fuzzy orderings is provided by the Szpilrajn theorem [24], which may be stated as follows: Let \( P \) be a partial ordering in \( X \). Then, there exists a linear ordering \( L \) in a set \( Y \), of the same cardinality as \( X \), and a one-to-one mapping \( \sigma \) from \( X \) onto \( Y \) (called the Szpilrajn mapping) such that for all \( x, y \) in \( X \)
\[ (x, y) \in P \Rightarrow (\sigma(x), \sigma(y)) \in L. \]

In its extended form, the statement of the theorem becomes:

**Theorem 8.** Let \( P \) be a fuzzy partial ordering in \( X \). Then, there exist a fuzzy linear ordering \( L \) in a set \( Y \), of the same cardinality as \( X \), and a one-to-one mapping \( \sigma \) from \( X \) onto \( Y \) such that
\[ \mu_P(x, y) > 0 \Rightarrow \mu_L(\sigma(x), \sigma(y)) = \mu_P(x, y), \quad x, y \in X. \tag{60} \]

**Proof.** The theorem can readily be established by the following construction for \( L \) and \( \sigma \): Assume that a fuzzy partial ordering \( P \) in \( X = \{x_1, \ldots, x_n\} \) is characterized by its relation matrix, which for simplicity will also be referred to as \( P \). In what follows, the relation matrix shown in Figure 3 and the Hasse diagram corresponding to it (Fig. 4) will be used to illustrate the construction for \( L \) and \( \sigma \).

*Information Sciences 3 (1971), 177–200*
First, we shall show that the antisymmetry and transitivity of $P$ make it possible to relabel the elements of $X$ in such a way that the corresponding relabeled relation matrix $P$ is upper-triangular.

To this end, let $C_0$ denote the set of undominating elements of $X$ (i.e., $x_i \in C_0 \iff$ column corresponding to $x_i$ contains a single positive element (unity) lying on the main diagonal). The transitivity of $P$ implies that $C_0$ is non-empty. For the relation matrix of Figure 3, $C_0 = \{x_1\}$.

Referring to the Hasse diagram of $P$ (Fig. 4), it will be convenient to associate with each $x_j$ in $X$ a positive integer $\rho(x_j; C_0)$ representing the level of $x_j$ above $C_0$. By definition

$$\rho(x_j; C_0) = \max_{x_i \in C_0} d(x_i, x_j),$$

where $d(x_i, x_j)$ is the length of the longest upward path between $x_i$ and $x_j$ in the Hasse diagram. For example, in Figure 4, $C_0 = \{x_1\}$ and $d(x_1, x_3) = 1$, $d(x_1, x_4) = 1$, $d(x_1, x_5) = 1$, $d(x_1, x_6) = 2$, $d(x_1, x_7) = 2$.

Now, let $C_m$, $m = 0, 1, \ldots, M$, denote the subset of $X$ consisting of those elements whose level is $m$, that is,

$$C_m = \{x_j | \rho(x_j; C_0) = m\},$$

with the understanding that, if $x_j$ is not reachable (via an upward path) from some element in $C_0$, then $x_j \in C_0$. For the example of Figure 4, we have $C_0 = \{x_1\}$, $C_1 = \{x_2, x_3, x_4\}$, $C_2 = \{x_5, x_6, x_7\}$, $C_3 = \emptyset$ (empty set). In words,

$$x_j \in C_m \iff \begin{align*}
& (i) \text{ there exists an element of } C_0 \text{ from which } x_j \text{ is reachable} \\
& \quad \text{via a path of length } m, \text{ and} \\
& (ii) \text{ there does not exist an element of } C_0 \text{ from which } x_j \text{ is} \\
& \quad \text{reachable via a path of length } > m.
\end{align*}$$

From (61) and (62) it follows that $C_0, \ldots, C_M$ have the following properties:

(a) Every $x_j$ in $X$ belongs to some $C_m$, $m = 0, \ldots, M$.

(b) $C_0, C_1, \ldots, C_M$ are disjoint. Thus, (a) and (b) imply that the collection $\{C_0, \ldots, C_M\}$ is a partition of $X$. 

Reason. Single-valuedness of \( \rho(x_j; C_0) \) implies that \( x_j \in C_k \) and \( x_j \in C_l \) cannot both be true if \( k \neq l \). Hence the disjointness of \( C_0, \ldots, C_M \).

\( x_i, x_j \in C_m \Rightarrow \mu_p(x_i, x_j) = \mu_p(x_j, x_i) = 0. \) (65)

Reason. Assume \( x_i, x_j \in C_m \) and \( \mu_p(x_i, x_j) > 0 \). Then
\[
\rho(x_j; C_0) > \rho(x_i; C_0),
\]
which contradicts the assumption that \( x_i, x_j \in C_m \). Similarly, \( \mu_p(x_j, x_i) > 0 \) contradicts \( x_i, x_j \in C_m \).

(d) \( x_j \in C_i \) and \( k < l \Rightarrow x_j \) is reachable from some \( x_i \) in \( C_k \).

Reason. If \( x_j \in C_i \), then there exists a path \( T \) of length \( l \) via which \( x_j \) is reachable from some \( x_i \) in \( C_0 \), and there does not exist a longer path via which \( x_j \) is reachable from any element of \( C_0 \). Now let \( x_i \) be the \( k \)th node of \( T \) (counting in the direction of \( C_i \)), with \( k < l \). Then \( x_i \in C_k \), since there exists a path of length \( k \) from \( x_i \) to \( x_j \) and there does not exist a longer path via which \( x_i \) is reachable from any element of \( C_0 \). (For, if such a path existed, then \( x_j \) would be reachable via a path longer than \( l \) from some element of \( C_0 \).) Thus \( x_j \) is reachable from some \( x_i \) in \( C_k \).

An immediate consequence of (d) is that the \( C_m \) may be defined recursively by
\[
C_{m+1} = \{ x_j | \rho(x_j; C_m) = 1 \}, \quad m = 0, 1, \ldots, M,
\] (66)
with the understanding that \( C_M \neq \emptyset \) and \( C_{M+1} = \emptyset \). More explicitly,
\[
x_j \in C_{m+1} \iff \mu_p(x_i, x_j) > 0 \text{ for some } x_i \in C_m \text{ and there does not exist an } x_i \in C_m \text{ and an } x_s \in X \text{ distinct from } x_i \text{ such that } \mu_p(x_i, x_s) > 0 \text{ and } \mu_p(x_s, x_j) > 0.
\] (e) \( x_i \in C_k \) and \( x_j \in C_i \) and \( k < l \Rightarrow \mu_p(x_j, x_i) = 0. \) (67)

Reason. Suppose \( \mu_p(x_j, x_i) > 0 \). By (d), \( x_j \) is reachable from some element of \( C_m \), say \( x_s \). If \( x_s = x_l \), then \( \mu_p(x_l, x_j) > 0 \), which contradicts the antisymmetry of \( P \). If \( x_s \neq x_l \), then by transitivity of \( P \), \( \mu_p(x_l, x_i) > 0 \), which contradicts (c) since \( x_j, x_l \in C_k \).

(f) \( x_i \in C_k \) and \( x_j \in C_i \) and \( \mu_p(x_i, x_j) > 0 \Rightarrow l > k. \) (68)

Reason. By negation of (e) and (c).

The partition \( \{ C_0, \ldots, C_M \} \), which can be constructed from the relation matrix \( P \) or by inspection of the Hasse diagram of \( P \), can be put to use in various ways. In particular, it can be employed to obtain the Hasse diagram of \( P \) from its relation matrix in cases in which this is difficult to do by inspection. Another application, which motivated our discussion of \( \{ C_0, \ldots, C_M \} \), relates to the

Information Sciences 3 (1971), 177–200
possibility of relabeling the elements of $X$ in such a way as to result in an upper-triangular relation matrix. By employing the properties of $(C_0, \ldots, C_M)$ stated above, this can readily be accomplished as follows:

Let $n_m$ denote the number of elements in $C_m$, $m = 0, \ldots, M$. Let the elements of $C_0$ be relabeled, in some arbitrary order, as $y_1, \ldots, y_{n_0}$, then the elements of $C_1$ be relabeled as $y_{n_0+1}, \ldots, y_{n_0+n_1}$, then the elements of $C_2$ be relabeled as $y_{n_0+n_1+1}, \ldots, y_{n_0+n_1+n_2}$, and so on, until all the elements of $X$ are relabeled in this manner. If the new label for $x_i$ is $y_j$, we write

$$Y_j = \sigma(x_i),$$

(69)

where $\sigma$ is a one-to-one mapping from $X = \{x_1, \ldots, x_n\}$ to $Y = \{y_1, \ldots, y_m\}$. Furthermore, we order the $y_j$ linearly by $y_j > y_i \iff i > j$, $i, j = 1, \ldots, n$.

The above relabeling transforms the relation matrix $P$ into the relation matrix $P_r$ defined by

$$P_r(x_i, x_j) = \mu_p(x_i, x_j), \quad x_i, x_j \in X.$$  

(70)

To verify that $P_r$ is upper-triangular, it is sufficient to note that, if $P_r(x_i, x_j) > 0$, for $x_i \neq x_j$, then by (f) $\sigma(x_j) > \sigma(x_i)$.

It is now a simple matter to construct a linear ordering $L$ in $Y$ which satisfies (60). Specifically, for $x_i \neq x_j$, let

$$\mu_L(\sigma(x_i), \sigma(x_j)) = \mu_p(x_i, x_j),$$

(71)

$$= 0, \quad \text{if } \mu_p(x_i, x_j) = 0 \text{ and } \mu_p(x_j, x_i) > 0,$$

$$= \varepsilon, \quad \text{if } \mu_p(x_i, x_j) = \mu_p(x_j, x_i) = 0 \text{ and } \sigma(x_j) > \sigma(x_i),$$

$$= 0, \quad \text{if } \mu_p(x_i, x_j) = \mu_p(x_j, x_i) = 0 \text{ and } \sigma(x_j) < \sigma(x_i),$$

where $\varepsilon$ is any positive constant which is smaller than or equal to the smallest positive entry in the relation matrix $P$.

Note. It is helpful to observe that this construction of $L$ may be visualized as a projection of the Hasse diagram of $P$ on a slightly inclined vertical line $Y$ (Figure 8). The purpose of the inclination is to avoid the possibility that two or more nodes of the Hasse diagram may be taken by the projection into the same point of $Y$.

All that remains to be demonstrated at this stage is that $L$, as defined by (71), is transitive. This is insured by our choice of $\varepsilon$, for, so long as $\varepsilon$ is smaller than or equal to the smallest entry in $P$, the transitivity of $P$ implies the transitivity of $L$, as is demonstrated by the following lemma:

**Lemma 9.** Let $P_r$ be an upper-triangular matrix such that $P_r = P_r^2$. Let $Q$ denote an upper-triangular matrix all of whose elements are equal to $\varepsilon$, where
0 ≤ ε ≤ smallest positive entry in \( P_r \). Then
\[ P_r \lor Q = (P_r \lor Q)^2. \]  
(72)

In other words, if \( P_r \) and \( Q \) are transitive, so is \( P_r \lor Q \).

**Proof.** We can rewrite (72) as
\[ (P_r \lor Q)^2 = P_r^2 \lor P_r \circ Q \lor Q \circ P_r \lor Q^2. \]  
(73)

Now \( P_r^2 = P_r \) and, since \( Q \) is upper-triangular, \( Q = Q^2 \). Furthermore, \( P_r \circ Q = Q \circ P_r = Q \). Hence (72).

To apply this lemma, we note that \( L \), as defined by (71), may be expressed as
\[ L = P_r \lor Q, \]  
(74)

where \( P_r \) and \( Q \) satisfy the conditions of the lemma. Consequently, \( L \) is transitive and thus is a linear ordering satisfying (60). This completes the proof of our extension of Szpirajn's theorem.

**CONCLUDING REMARK**

As the foregoing analysis demonstrates, it is a relatively simple matter to extend some of the well-known results in the theory of relations to fuzzy sets. It appears that such extension may be of use in various applied areas, particularly those in which fuzziness and/or randomness play a significant role in the analysis or control of system behavior.

*Information Sciences* 3 (1971), 177–200
ACKNOWLEDGMENT

The author is indebted to P. Varaiya, E. T. Lee, and J. Yang for helpful criticisms.

REFERENCES

9. J. B. Kruskal, Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis, Psychometrika 29 (1964), 1-27.
10. J. H. Ward, Jr., Hierarchical grouping to optimize an objective function, Amer. Statist. Assoc. 58 (1963), 236-244.

Received August 17, 1970