Pattern Classification Based on Fuzzy Relations

SHINICHI TAMURA, STUDENT MEMBER, IEEE, SEIHAKU HIGUCHI, AND KOKICHI TANAKA, SENIOR MEMBER, IEEE

Abstract—**A method of classifying patterns using fuzzy relations is described. To start with, we give a suitable value of the measure of subjective similarity to each pair of patterns that is taken from the population of patterns to be classified. Then a similitude between any two patterns is calculated by using the composition of a fuzzy relation. The similitude induces an equivalence relation. Consequently, we can classify the present population of the patterns into some classes by the equivalence relation. An experiment of the classification of portraits has been performed to test the method proposed here.**

I. Introduction

SINCE Zadeh published the fuzzy set theory [1]–[6], it has been applied to some fields such as automata, learning, and control [7]–[10]. We introduce a concept of the fuzzy relation [1] to measure the subjective similarity as follows.

In the classification of smells and the classification of pictures, etc., subjective information plays an important role. This subjective information may be represented by the fuzzy relation that corresponds to the subjective similarity. However, since such a primary fuzzy relation is made on the basis of a personal subject, it does not satisfy the axioms of distance. Hence in this paper we construct an n-step fuzzy relation by the composition of the fuzzy relation, and define a similitude as a limit value of the n-step fuzzy relation in order to satisfy the axioms of distance. The similitude defined in such a way induces an equivalence relation. Thus we can classify patterns by the equivalence relation.

II. Fuzzy Relation

Let X be a set of patterns. The fuzzy relation A on X is characterized by $f_A(x,y) \in [0,1]$, for all $x,y \in X$. In this paper, we first consider a one-step fuzzy relation $f_1(x,y)$ satisfying the two conditions

$$f_1(x,x) = 1, \qquad \forall x \in X \tag{1}$$

$$f_1(x,y) = f_1(y,x), \qquad \forall x,y \in X. \tag{2}$$

Condition (1) means that x is perfectly the same with x. Condition (2) means that the fuzzy relation considered here is symmetric. Assume that the value of this one-step fuzzy relation $f_1(x,y)$ is given to each of the pairs of patterns in X. Any $f_1(x,y)$ will do if it satisfies conditions (1) and (2); for example, subjective similarities, normalized correlations, or potential functions, etc., may be conceived.

Now, we define the n-step fuzzy relation $f_n(x,y)$ by

$$f_n(x,y) = \sup_{x_1,x_2,\cdots,x_{n-1}\in X} \min\,[f_1(x,x_1),f_1(x_1,x_2), \cdots, f_1(x_{n-1},y)], \qquad n = 2,3,\cdots.$$

Then

$$\begin{aligned} f_{n+1}(x,y) &= \sup_{x_1,\cdots,x_{n-1},x_n\in X} \min\,[f_1(x,x_1), \cdots, f_1(x_{n-1},x_n),f_1(x_n,y)] \\ &\geq \sup_{x_1,\cdots,x_{n-2},x_{n-1}\in X} \min\,[f_1(x,x_1), \cdots, f_1(x_{n-1},y),f_1(y,y)] \\ &= f_n(x,y). \end{aligned}$$

Therefore, we see

$$0 \leq f_1(x,y) \leq f_2(x,y) \leq \cdots \leq f_n(x,y) \leq f_{n+1}(x,y) \leq \cdots \leq 1 \tag{3}$$

and we have the similitude $f(x,y)$ in $[0,1]$ such that

$$f(x,y) = \lim_{n\to\infty} f_n(x,y).$$

We will show some important properties of $f(x,y)$ in the following.

Definition 1: Let x and y be two elements of X. Then x and y are said to have a stronger relation than λ, written $xR_\lambda y$, iff $f(x,y) \geq \lambda$. Symbolically this is expressed as

$$xR_\lambda y \Leftrightarrow f(x,y) \geq \lambda.$$

Lemma 1: For all $x,y,z \in X$,

$$f(x,z) \geq \min\,[f(x,y),f(y,z)].$$

Proof: See Appendix I.

Theorem 1: R_λ is an equivalence relation on X.

Proof:

1) From (3) and the assumptions, we have

$$1 = f_1(x,x) \leq f(x,x) \leq 1.$$

Then $f(x,x) = 1$, or $xR_\lambda x$, for all $\lambda \in [0,1]$.

2) By the assumption, $f_1(x,y) = f_1(y,x)$. Then $f_n(x,y) = f_n(y,x)$, and we can conclude that $f(x,y) = f(y,x)$. This means that $xR_\lambda y \Leftrightarrow yR_\lambda x$.

Manuscript received March 12, 1970; revised September 1, 1970.
The authors are with the Department of Information and Computer Sciences, Faculty of Engineering Science, Osaka University, Toyonaka, Japan 560.

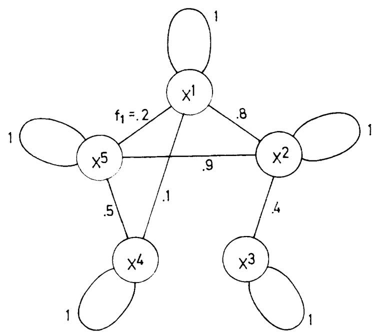

Fig. 1. One-step fuzzy relation $f_1(x,y)$ of Example 1.

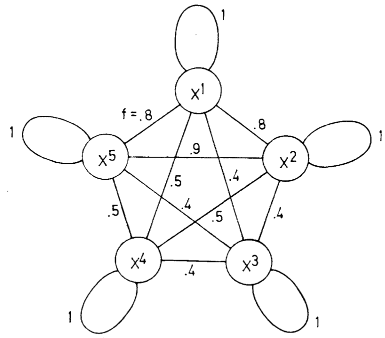

Fig. 2. $f(x,y)$ of Example 1.

3) From Lemma 1, we have

$$f(x,z) \geq \min\,[f(x,y),f(y,z)].$$

Therefore, we can conclude that

$$xR_\lambda y, yR_\lambda z \Rightarrow xR_\lambda z. \qquad \text{Q.E.D.}$$

Thus, by Theorem 1, we see that we can classify the patterns using the partition induced by the equivalence relation R_λ with the appropriate threshold λ.

Example 1. Let $X = \{x^1,x^2,\cdots,x^5\}$ and $f_1(x,y)$ be as follows.

	x^1	x^2	x^3	x^4	x^5
x^1	1				
x^2	0.8	1			
x^3	0	0.4	1		
x^4	0.1	0	0	1	
x^5	0.2	0.9	0	0.5	1

This table is illustrated in Fig. 1. Then we have $f(x,y) = f_3(x,y)$ as follows.

	x^1	x^2	x^3	x^4	x^5
x^1	1				
x^2	0.8	1			
x^3	0.4	0.4	1		
x^4	0.5	0.5	0.4	1	
x^5	0.8	0.9	0.4	0.5	1

This table is illustrated in Fig. 2. We have the partitions (see, e.g., Harrison [11])

$$R_0 = R_{0.3} = \{[x^1,x^2,x^3,x^4,x^5]\}$$
$$R_{0.45} = \{[x^1,x^2,x^4,x^5],[x^3]\}$$
$$R_{0.55} = \{[x^1,x^2,x^5],[x^4],[x^3]\}$$
$$R_{0.85} = \{[x^1],[x^2,x^5],[x^4],[x^3]\}$$
$$R_1 = \{[x^1],[x^2],[x^5],[x^4],[x^3]\}.$$

Thus the patterns are classified by the partition induced by R_λ.

Theorem 2: Let $\lambda \geq \mu$. Then R_λ refines R_μ.

Proof: It is sufficient to show that $xR_\lambda y \Rightarrow xR_\mu y$. Assume that $xR_\lambda y$, then $f(x,y) \geq \lambda \geq \mu$. Therefore, $xR_\mu y$. Q.E.D.

Conversely, if $f_1(x,y)$ is changed, we have the following theorem.

Theorem 3: Assume $f_1(x,y) \leq f_1'(x,y)$, for all $x,y \in X$, let R_λ be an equivalence relation induced by $f_1(x,y)$, and let R_λ' be one by $f_1'(x,y)$. Then R_λ refines R_λ'.

Proof: The proof is obvious.

Theorem 4: If $f(x,y) \neq 1$, for all $x,y \in X$ such that $x \neq y$, then $\rho(x,y) = 1 - f(x,y)$ satisfies the axioms of distance.

Proof:

1) Since by the assumptions $0 \leq f(x,y) < 1$, for $x \neq y$, and $f(x,x) = 1$, we have $\rho(x,y) > 0$, for $x \neq y$, and $\rho(x,x) = 0$.
2) Since $f(x,y) = f(y,x)$, then $\rho(x,y) = \rho(y,x)$.
3) By Lemma 1,

$$f(x,z) \geq \min\,[f(x,y),f(y,z)] \geq f(x,y) + f(y,z) - 1.$$

Then

$$\rho(x,z) \leq \rho(x,y) + \rho(y,z). \qquad \text{Q.E.D.}$$

When $f(x,y) = 1$ holds for some $x \neq y$, the assumption of Theorem 4 is not satisfied. In such a case, the following theorem can be easily demonstrated.

Theorem 5: Let R_1 be a set of the equivalence classes induced by R_1 on X, and $\bar{x}$ and $\bar{y}$ be two elements of R_1. Let x and y be arbitrary elements in $\bar{x}$ and $\bar{y}$, respectively. Then $\rho(\bar{x},\bar{y}) = 1 - f(x,y)$ satisfies the axioms of distance.

Note that if we make changes such as $f \leftrightarrow \rho$, $\geq\, \leftrightarrow\, \leq$, sup $\leftrightarrow$ inf, and max $\leftrightarrow$ min, our approach will be changed into the complementary one (see Mizumoto *et al.* [10]).

When the threshold λ is not changed, we may memorize only whether each $f_1(x,y)$ is greater than λ or not, instead of memorizing the values of $f_1(x,y)$. For such a case, let us consider the transitive closure (see Harrison [11]). The transitive closure of Q_λ, written $\hat{Q}_\lambda$, is defined as

$$\hat{Q}_\lambda = \bigcup_{i=1}^{\infty} Q_\lambda^{\,i} = Q_\lambda \cup (Q_\lambda Q_\lambda) \cup (Q_\lambda Q_\lambda Q_\lambda) \cup \cdots$$

where Q_λ is a relation on X. Let

$$xQ_\lambda y \Leftrightarrow f_1(x,y) \geq \lambda.$$

Then, since $xQ_\lambda x$ and $xQ_\lambda y \Leftrightarrow yQ_\lambda x$, for all $\lambda \in [0,1]$, $\hat{Q}_\lambda$ is the equivalence relation on X. Roughly speaking, the classification by $\hat{Q}_\lambda$ is based on whether there is a path connecting two patterns or not.

Theorem 6: For all λ in $[0,1]$, $\hat{Q}_\lambda$ refines R_λ.

Proof: It is sufficient to show that $x\hat{Q}_\lambda y \Rightarrow xR_\lambda y$. Assume $x\hat{Q}_\lambda y$, then, for some integer n, $xQ_\lambda{}^n y$. Then there exist $x_1,x_2,\cdots,x_{n-1}$ in X such that

$$f_1(x,x_1) \geq \lambda, f_1(x_1,x_2) \geq \lambda,\cdots,f_1(x_{n-1},y) \geq \lambda.$$

That is,

$$f_n(x,y) \geq \min\,[f_1(x,x_1),f_1(x_1,x_2),\cdots,f_1(x_{n-1},y)] \geq \lambda.$$

Then

$$f(x,y) \geq f_n(x,y) \geq \lambda$$

or

$$xR_\lambda y. \qquad \text{Q.E.D.}$$

In Section IV we will show that if X is finite, $\hat{Q}_\lambda$ becomes equal to R_λ. Furthermore, we can easily obtain the same theorems as Theorems 2 and 3 for $\hat{Q}_\lambda$.

III. Abbreviation Form

We show an abbreviation form of $f_n(x,y)$ as a preliminary step to discussing the properties when the pattern set X is finite. Let

$$g_n(x,x_1,x_2,\cdots,x_{n-1},y) = \min\,[f_1(x,x_1),f_1(x_1,x_2),\cdots, f_1(x_{n-1},y)].$$

Then

$$f_n(x,y) = \sup_{x_1,\cdots,x_{n-1}\in X} g_n(x,x_1,\cdots,x_{n-1},y).$$

Generally speaking, if $x_i = x_j$, then

$$g_n(x,x_1,\cdots,x_i,x_{i+1},\cdots,x_j,x_{j+1},\cdots,x_{n-1},y) \leq g_{n-j+i}(x,x_1,\cdots,x_i,x_{j+1},\cdots,x_{n-1},y).$$

This implies that we can remove the loop $(x_i,x_{i+1},\cdots,x_j)$ in the string $(x,x_1,\cdots,x_{n-1},y)$ when we calculate $f_n(x,y)$. Therefore, we have the abbreviation form

$$f_n(x,y) = \max_{k\in K} g_{nk}'(x,y)$$

where

$$g_{nk}'(x,y) = \begin{cases} \sup\limits_{(x_1,\cdots,x_k)\in X_k} g_n(x,x_1,\cdots,x_k,y,y,\cdots,y), & k = 1,2,\cdots,p \\ f_1(x,y), & k = 0 \end{cases}$$

$$X_k = \{(x_1,\cdots,x_k) \mid x_1 \in X - \{x,y\}, x_2 \in X - \{x,x_1,y\}, \cdots, x_k \in X - \{x,x_1,\cdots,x_{k-1},y\}\}$$

$$p = \min\,(n-1, N-2)$$

$$K = \{0,1,\cdots,p\}.$$

In our case, since $f_1(x,x) = 1$, when X is finite, we can easily show that (see Wee *et al.* [9], Mizumoto *et al.* [10])

$$f_{N-1}(x,y) = f_N(x,y) = \cdots = f(x,y)$$

where N is the number of elements in X.

Theorem 7: Let $x \neq y$ and $X' = X - \{x\}$. Then

$$f_n(x,y) \leq \sup_{x_1\in X'} f_j(x,x_1), \qquad j = 1,2,\cdots,n.$$

Proof: See Appendix II.

IV. Finite Sets

In actual cases, we usually deal with only the finite number of the patterns. Let us consider the equivalence relations on such a finite set.

Theorem 8: If the number N of elements in X is finite, then

$$R_\lambda = \hat{Q}_\lambda = Q_\lambda{}^{N-1}.$$

Proof: The latter half of the theorem is obvious (see Harrison [11]). Let us show that R_λ is equal to $\hat{Q}_\lambda$. Since we have shown that $\hat{Q}_\lambda$ refines R_λ in Theorem 6, it is sufficient to show that R_λ refines $\hat{Q}_\lambda$. Assume that $xR_\lambda y$. Then

$$f(x,y) = f_{N-1}(x,y) = \max_{x_1,\cdots,x_{N-2}\in X} \min\,[f_1(x,x_1), f_1(x_1,x_2),\cdots,f_1(x_{N-2},y)] \geq \lambda.$$

This means that there exist $x_1,\cdots,x_{N-2} \in X$ such that

$$f_1(x,x_1) \geq \lambda, f_1(x_1,x_2) \geq \lambda,\cdots,f_1(x_{N-2},y) \geq \lambda$$

so that

$$xQ_\lambda x_1, x_1Q_\lambda x_2,\cdots,x_{N-2}Q_\lambda y.$$

We see $xQ_\lambda{}^{N-1}y$ or $x\hat{Q}_\lambda y$. Q.E.D.

When X is finite, it is sometimes convenient to use the fuzzy matrix representation. We represent the fuzzy matrix as

$$F = \|f_1(x^i,x^j)\|, \qquad i,j = 1,2,\cdots,N.$$

Let us show some fundamental properties of our fuzzy matrix. We denote by a_{ij} the (i,j)th entry of a fuzzy matrix A, where $0 \leq a_{ij} \leq 1$. We define

$$A \prec B \Leftrightarrow a_{ij} \leq b_{ij}$$

$$I = \|m_{ij}\|$$

where

$$m_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$$

$$C = A \circ B \Leftrightarrow c_{ij} = \max_k \min\,(a_{ik},b_{kj})$$

$$A^{m+1} = A^m \circ A$$

$$A^0 = I$$

$$C = \max\,(A,B) \Leftrightarrow c_{ij} = \begin{cases} a_{ij}, & \text{if } a_{ij} \geq b_{ij} \\ b_{ij}, & \text{otherwise.} \end{cases}$$

Then, since all the diagonal elements of F are now equal to unity, we have (see Mizumoto *et al.* [10])

$$I \prec F \prec F^2 \prec \cdots \prec F^{N-1} = F^N = \cdots = F^\infty.$$

The (i,j)th entry of F^k is $f_k(x^i,x^j)$. Hence we can calculate $f(x^i,x^j)$ easily and rather quickly by using $F^k \circ F^k = F^{2k}$.

As in Example 1, we have

$$F = \begin{bmatrix} 1 & 0.8 & 0 & 0.1 & 0.2 \\ 0.8 & 1 & 0.4 & 0 & 0.9 \\ 0 & 0.4 & 1 & 0 & 0 \\ 0.1 & 0 & 0 & 1 & 0.5 \\ 0.2 & 0.9 & 0 & 0.5 & 1 \end{bmatrix}$$

$$F^2 = F \circ F = \begin{bmatrix} 1 & 0.8 & 0.4 & 0.2 & 0.8 \\ 0.8 & 1 & 0.4 & 0.5 & 0.9 \\ 0.4 & 0.4 & 1 & 0 & 0.4 \\ 0.2 & 0.5 & 0 & 1 & 0.5 \\ 0.8 & 0.9 & 0.4 & 0.5 & 1 \end{bmatrix}$$

$$F^3 = \begin{bmatrix} 1 & 0.8 & 0.4 & 0.5 & 0.8 \\ 0.8 & 1 & 0.4 & 0.5 & 0.9 \\ 0.4 & 0.4 & 1 & 0.4 & 0.4 \\ 0.5 & 0.5 & 0.4 & 1 & 0.5 \\ 0.8 & 0.9 & 0.4 & 0.5 & 1 \end{bmatrix} = F^4 = F^5 = \cdots = F^\infty.$$

Thus

$$I \prec F \prec F^2 \prec F^3 = F^4 = \cdots = F^\infty.$$

Next, we show a divided calculation method of a fuzzy matrix. If $C = A \circ B$, then

$$\begin{aligned} c_{ij} &= \max_{1 \le k \le N} \min (a_{ik}, b_{kj}) \\ &= \max \left[\max_{1 \le k \le \beta} \min (a_{ik}, b_{kj}), \max_{\beta+1 \le k \le N} \min (a_{ik}, b_{kj}) \right]. \end{aligned}$$

We can then obtain the following results. If A and B are divided as

$$A = \begin{bmatrix} D & E \\ F & G \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} H & J \\ K & L \end{bmatrix}$$

then

$$\begin{aligned} A \circ B &= \begin{bmatrix} \max (D \circ H, E \circ K) & \max (D \circ J, E \circ L) \\ \max (F \circ H, G \circ K) & \max (F \circ J, G \circ L) \end{bmatrix} \\ &= \max \left(\begin{bmatrix} D \circ H & D \circ J \\ F \circ H & F \circ J \end{bmatrix}, \begin{bmatrix} E \circ K & E \circ L \\ G \circ K & G \circ L \end{bmatrix} \right). \end{aligned}$$

This is the same form as that of the nonfuzzy matrix.

Let us consider the case when a new pattern x^{N+1} comes into X. The pattern set X becomes $X + \{x^{N+1}\}$. We denote a new n-step fuzzy relation on $X + \{x^{N+1}\}$ by $f_n'(x,y)$. Generally speaking, $f_n'(x,y)$ becomes different from $f_n(x,y)$. After some manipulations, we obtain

$$\begin{aligned} f_n'(x^i,x^j) &= \max [f_n(x^i,x^j), \min \{f_{n-1}'(x^i,x^{N+1}), \\ &\quad f_1'(x^{N+1},x^j)\}, \min \{f_{n-2}'(x^i,x^{N+1}), \\ &\quad f_2'(x^{N+1},x^j)\}, \cdots, \min \{f_1'(x^i,x^{N+1}), \\ &\quad f_{n-1}'(x^{N+1},x^j)\}], \qquad x^i,x^j \in X \\ f_n'(x^{N+1},x^{N+1}) &= 1 \\ f_n'(x^{N+1},x^i) &= f_n'(x^i,x^{N+1}) \\ &= \max_{x_1 \in X} \min [f_1'(x^{N+1},x_1), f_{n-1}(x_1,x^i)], \\ &\qquad x^i \in X. \end{aligned}$$

However, in almost all cases it is easier to recompute the F^n than to calculate by these formulas.

V. Experimental Result

Portraits obtained from 60 families were used in our experiment, each of which is composed of between four and seven members. The reason why we chose the portraits is that we had conceived that even if parents do not resemble each other in face, they may be connected through their children, and consequently we could classify the portraits into families. First, we divided the 60 families into 20 groups, each of which was composed of 3 families. Each group was, on the average, composed of 15 members. The portraits of each group were presented to a different student to give the values of the subjective similarity $f_1(x,y)$ by 5 rank representation to all pairs between them. The reason why we used the 5 rank representation instead of continuous value representation is that it had been proved that the human being cannot distinguish into more than 5 ranks in the end. Twenty students joined in this experiment. Two examples of the experiment are shown in Table I, Table II, Fig. 3, and Fig. 4. In Table I, the 5 rank representations are converted to the values in [0,1]. In our case, the number of patterns is not so many that we can classify by inspection without calculating $f_n(x,y)$.

Since the levels of the subjective values are different according to individuals, the threshold was determined in each group as follows. As we bring down the threshold, the number of classes decrease. Hence, under the assumption that the number of classes c to be classified was known to be 3, bringing down the threshold we stopped at the value which divided the patterns into 3 classes (collection of the patterns composed of more than 2 patterns that have a stronger relation than λ with each other) and some non-connected patterns. However, as in the present case, when some $f_1(x,y)$ take the same value, sometimes there is no threshold by which the patterns are divided into just c given classes. In such a case, we made it possible to divide them into just c classes by stopping the threshold at the value where the patterns are divided into less than c classes and separating some connections randomly that have a minimum $f_1(x,y)$ of connections that have the stronger relation than the threshold.

The correctly classified rates, the misclassified rates, and the rejected rates of 20 groups were within the range of 50–94 percent, 0–33 percent, and 0–33 percent, respectively, and we obtained the correctly classified rate 75 percent, the misclassified rate 13 percent, and the rejected rate 12 percent as the averages of the 20 groups. Here, since the classes made in this experiment have no label, we calculated these rates by making a one-to-one correspondence between 3 families and 3 classes, so as to have the largest number of correctly classified patterns.

VI. Conclusion

We have studied pattern classification using subjective information and performed experiments involving classification of portraits. The method of classification proposed here is based on the procedure of finding a path connecting 2

TABLE I
SUBJECTIVE SIMILARITIES OF FIG. 3

Portrait Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1															
2	0	1														
3	0	0	1													
4	0	0	0.4	1												
5	0	0.8	0	0	1											
6	0.5[a]	0	0.2	0.2	0	1										
7	0	0.8	0	0	0.4	0	1									
8	0.4	0.2	0.2	0.5[a]	0	0.8	0	1								
9	0	0.4	0	0.8	0.4	0.2	0.4	0	1							
10	0	0	0.2	0.2	0	0	0.2	0	0.2	1						
11	0	0.5[a]	0.2	0.2	0	0	0.8	0	0.4	0.2	1					
12	0	0	0.2	0.8	0	0	0	0	0.4	0.8	0	1				
13	0.8	0	0.2	0.4	0	0.4	0	0.4	0	0	0	0	1			
14	0	0.8	0	0.2	0.4	0	0.8	0	0.2	0.2	0.6	0	0	1		
15	0	0	0.4	0.8	0	0.2	0	0	0.2	0	0	0.2	0.2	0	1	
16	0.6	0	0	0.2	0.2	0.8	0	0.4	0	0	0	0	0.4	0.2	0	1

[a] This value was converted from 0.6 to 0.5 for division into just three classes.

TABLE II
THE $f(x,y)$ $(= f_6(x,y) = f_7(x,y) = \cdots)$ OF TABLE I

Portrait Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1															
2	0.4	1														
3	0.4	0.4	1													
4	0.5	0.4	0.4	1												
5	0.4	0.8	0.4	0.4	1											
6	0.6	0.4	0.4	0.5	0.4	1										
7	0.4	0.8	0.4	0.4	0.8	0.4	1									
8	0.6	0.4	0.4	0.5	0.4	0.8	0.4	1								
9	0.5	0.4	0.4	0.8	0.4	0.5	0.4	0.5	1							
10	0.5	0.4	0.4	0.8	0.4	0.5	0.4	0.5	0.8	1						
11	0.4	0.8	0.4	0.4	0.8	0.4	0.8	0.4	0.4	0.4	1					
12	0.5	0.4	0.4	0.8	0.4	0.5	0.4	0.5	0.8	0.8	0.4	1				
13	0.8	0.4	0.4	0.5	0.4	0.6	0.4	0.6	0.5	0.5	0.4	0.5	1			
14	0.4	0.8	0.4	0.4	0.8	0.4	0.8	0.4	0.4	0.4	0.8	0.4	0.4	1		
15	0.5	0.4	0.4	0.8	0.4	0.5	0.4	0.5	0.8	0.8	0.4	0.8	0.5	0.4	1	
16	0.6	0.4	0.4	0.5	0.4	0.8	0.4	0.8	0.5	0.5	0.4	0.5	0.6	0.4	0.5	1

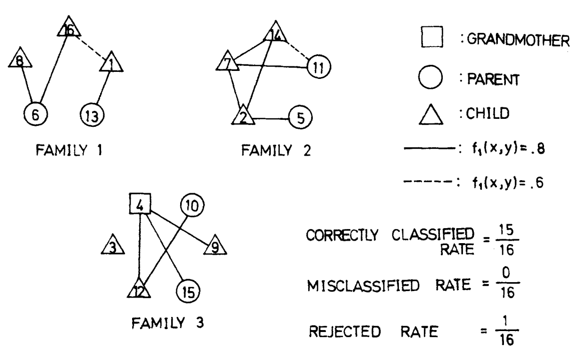

Fig. 3. Portrait classification of Table I.

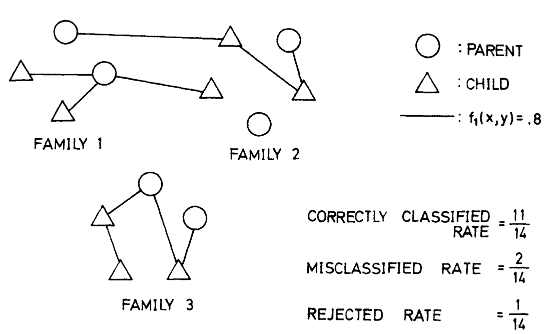

Fig. 4. Example of portrait classification.

patterns. Therefore, this method may be combined with nonsupervised learning and may also be applicable to information retrieval [12] and path detection.

Appendix I

Proof of Lemma 1

$$\begin{aligned} f_{m+n}(x,z) &= \sup_{x_1,\cdots,x_{m+n-1}} \min [f_1(x,x_1),\cdots,f_1(x_{m+n-1},z)] \\ &\geq \sup_{x_1,\cdots,x_{m-1}} \sup_{x_{m+1},\cdots,x_{m+n-1}} \min [f_1(x,x_1),\cdots, f_1(x_{m-1},y),f_1(y,x_{m+1}),\cdots,f_1(x_{m+n-1},z)] \\ &= \sup_{x_1,\cdots,x_{m-1}} \sup_{x_{m+1},\cdots,x_{m+n-1}} \min [\min \{f_1(x,x_1),\cdots,f_1(x_{m-1},y)\}, \min \{f_1(y,x_{m+1}),\cdots,f_1(x_{m+n-1},z)\}] \\ &= \min [\sup_{x_1,\cdots,x_{m-1}} \min \{f_1(x,x_1),\cdots,f_1(x_{m-1},y)\}, \sup_{x_{m+1},\cdots,x_{m+n-1}} \min \{f_1(y,x_{m+1}),\cdots,f_1(x_{m+n-1},z)\}] \\ &= \min [f_m(x,y), f_n(y,z)]. \end{aligned}$$

We obtain the following inequality as $m \to \infty$ and $n \to \infty$:

$$f(x,z) \geq \min [f(x,y), f(y,z)].$$

Appendix II

Proof of Theorem 7

$$\begin{aligned} f_n(x,y) &= \sup_{x_j,\cdots,x_{n-1}\in X'} \sup_{x_1,\cdots,x_{j-1}\in X'} \min [\min \{f_1(x,x_1),\cdots,f_1(x_{j-1},x_j)\}, \min \{f_1(x_j,x_{j+1}),\cdots,f_1(x_{n-1},y)\}] \\ &= \sup_{x_j,\cdots,x_{n-1}\in X'} \min \left[\sup_{x_1,\cdots,x_{j-1}\in X'} \min \{f_1(x,x_1),\cdots,f_1(x_{j-1},x_j)\}, \min \{f_1(x_j,x_{j+1}),\cdots,f_1(x_{n-1},y)\} \right] \\ &= \sup_{x_j,\cdots,x_{n-1}\in X'} \min [f_j(x,x_j), \min \{f_1(x_j,x_{j+1}),\cdots, f_1(x_{n-1},y)\}] \leq \sup_{x_j,\cdots,x_{n-1}\in X'} f_j(x,x_j) \\ &= \sup_{x_j\in X'} f_j(x,x_j). \end{aligned}$$

Acknowledgment

The authors wish to thank M. Mizumoto, J. Toyoda, and other members of Prof. Tanaka's laboratory for their helpful suggestions and discussions.

References

[1] L. A. Zadeh, "Fuzzy sets," *Inform. Contr.*, vol. 8, pp. 338–353, June 1965.
[2] ——, "Fuzzy sets and systems," *Proc. 1965 Symp. on Syst. Theory* (Brooklyn, N.Y.), pp. 29–39.
[3] R. Bellman, R. Kalaba, and L. Zadeh, "Abstraction and pattern classification," *J. Math. Anal. Appl.*, vol. 13, pp. 1–7, January 1966.
[4] L. A. Zadeh, "Fuzzy algorithms," *Inform. Contr.*, vol. 12, pp. 99–102, February 1968.
[5] ——, "Probability measures of fuzzy events," *J. Math. Anal. Appl.*, vol. 23, pp. 421–427, August 1968.
[6] ——, "Toward a theory of fuzzy systems," Electron. Res. Lab., College of Engineering, University of California, Berkeley, Rep. ERL-69-2, June 1969.
[7] J. A. Goguen, "*L*-fuzzy sets," *J. Math. Anal. Appl.*, vol. 18, pp. 145–174, April 1967.
[8] E. S. Santos and W. G. Wee, "General formulation of sequential machines," *Inform. Contr.*, vol. 12, pp. 5–10, January 1968.
[9] W. G. Wee and K. S. Fu, "A formulation of fuzzy automata and its application as a model of learning systems," *IEEE Trans. Syst. Sci. Cybern.*, vol. SSC-5, pp. 215–223, July 1969.
[10] M. Mizumoto, J. Toyoda, and K. Tanaka, "Some considerations on fuzzy automata," *J. Comput. Syst. Sci.*, vol. 3, pp. 409–422, November 1969.
[11] M. A. Harrison, *Introduction to Switching and Automata Theory.* New York: McGraw-Hill, 1965, pp. 9–12, 296.
[12] G. Salton, "Associative document retrieval techniques using bibliographic information," *J. Ass. Comput. Mach.*, vol. 10, pp. 440–457, 1963.