Chapter 2
Pattern Classification Based on Conventional Interpretation of MFI

Abstract Our aim is to design a pattern classifier using fuzzy relational calculus (FRC) which was initially proposed by Pedrycz (Pattern Recognition **23** (1/2), 121–146, 1990). In the course of doing so, we first consider a particular interpretation of the multidimensional fuzzy implication (MFI) to represent our knowledge about the training data set. Subsequently, we introduce the notion of a fuzzy pattern vector to represent a population of training patterns in the pattern space and to denote the antecedent part of the said particular interpretation of the MFI. We introduce a new approach to the computation of the derivative of the fuzzy max-function and min-function using the concept of a generalized function. During the construction of the classifier based on FRC, we use fuzzy linguistic statements (or fuzzy membership function to represent the linguistic statement) to represent the values of features (e.g., feature F_1 is small and F_2 is big) for a population of patterns. Note that the construction of the classifier essentially depends on the estimate of a fuzzy relation $\Re$ between the input (fuzzy set) and output (fuzzy set) of the classifier. Once the classifier is constructed, the nonfuzzy features of a pattern can be classified. At the time of classification of the nonfuzzy features of the test patterns, we use the concept of fuzzy masking to fuzzify the nonfuzzy feature values of the test patterns. The performance of the proposed scheme is tested on synthetic data. Finally, we use the proposed scheme for the vowel classification problem of an Indian language.

2.1 Introduction

In real world pattern classification problems, fuzziness is connected with diverse facets of cognitive activity of human being. The sources of fuzziness are related to labels expressed in pattern space, as well as, labels of classes taken into

K. S. Ray, *Soft Computing Approach to Pattern Classification and Object Recognition*, DOI: 10.1007/978-1-4614-5348-2_2,

account in classification procedures. Although a lot of scientific developments have already been made in the area of pattern classification, existing techniques of pattern classification remain inferior to the human classification processes which perform extremely complex tasks. Hence, we attempt to develop a plausible tool using fuzzy relational calculus (FRC) for modeling and mimicking the cognitive process of human reasoning for pattern classification. The FRC approach to pattern classification can take care uncertainties in feature values of patterns under different conditions like measurement error, noise, etc. Though there are several existing approaches to designing a classifier using the concept of fuzzy set/fuzzy logic (Mori 1983; Mori and Laface 1980; Hirota 1988; Seif and Aguilar-Martin 1980; Dubois and Jaulent 1985; Hirota et al. 1987; Huntsberger et al. 1985; Kickert and Koppleaar 1986; Siy and Chen 1974; Shimura 1975; Huntsberger et al. 1986; Lee 1972; Zadeh et al. 1975; Bortolan and Degani 1983; Bortolan et al. 1988; Degani and Bortolan 1987a, b; Pedrycz 1985a, b, c; Watanabe 1985; Kumar 1977; Saitta and Tarasso 1981; Woodbury and Clive 1974; Bezdek and Pal 1992; Simpson 1992), we have selected the concept proposed by Pedrycz (1990) and suitably modified it to incorporate our new concept of computation of the derivative of the fuzzy max—function and min—function. To represent the knowledge about the training data set, we consider the conventional interpretation of multidimensional fuzzy implication (MFI) (Sugeno and Takagi 1983; Tsukamoto 1979). We introduce a novel notion of fuzzy pattern vector as stated in Appendix-A to represent a population of patterns (a set of patterns) in the pattern space. It represents the antecedent part of the said particular interpretation of the MFI to meaningfully carry out the task of pattern classification using FRC. During the construction of the classifier based on FRC, we use fuzzy linguistic statements (or fuzzy membership function to represent the linguistic statement) to represent the values of features (e.g., feature F_1 is small and F_2 is big) for a population of patterns (a set of patterns) represented by the above said notion of Fuzzy pattern vector. Note that the construction of the classifier essentially depends on the estimation of a fuzzy relation $\Re$ between the antecedent part and consequent part of the rules. As, for a given problem of pattern classification and object recognition, there is no specific guideline to select a particular logical operator, e.g., Mamani's min operator, Zadeh's arithmetic rule etc. (Mizumoto 1985; Zadeh 1970) to translate a fuzzy implication to a fuzzy relation we estimate it (the relation $\Re$), based on different learning scheme using soft computing tools. Thus, in our entire treatment for classification and recognition we replace logic by learning. The estimated $\Re$ is the core of the classifier (recognizer). Once the classifier is constructed, the nonfuzzy features of a pattern can be classified. At the time of classification of the nonfuzzy features of the test patterns, we use the concept of fuzzy masking to fuzzify the nonfuzzy feature values of the test patterns. The performance of the proposed scheme is tested on synthetic data. Finally, we use the proposed scheme for the vowel classification problem of an Indian language.

2.2 Statement of the Problem

For the present problem, let us consider the conventional interpretation of a MFI [see Appendix-A, Eq. A.1 & A.2] as stated below;

$$\text{a) if } x \text{ is A and } y \text{ is } B \text{ then } z \text{ is } C$$

or

$$\text{b) if } x \text{ is A then } y \text{ is } B \text{ then } z \text{ is } C. \tag{2.1}$$

The notion of a fuzzy pattern vector (see Appendix-A) represents the antecedent clauses of (a) of (2.1) and locates a population of patterns P in the quantized pattern space. Assume that the quantized pattern space consists of "c" universes U_1, U_2, …, U_c in the form $U = U_1 \times U_2 \times \cdots \times U_c$, where each U_i represents the universe on the ith feature axis F_i, $i = 1, 2, \ldots, c$.

Assume that D is a fuzzy relation [formed by the antecedent clauses of (a) of (2.1)], which is a fuzzy set in quantized product space U, namely μ_D: $U \to [0, 1]$. Also, assume that there exists a set C_{class} of finite number of classes $c_1, c_2, \ldots, c_n$, i.e., $C_{class} = \{c_1, c_2, \ldots, c_n\}$, by which the finite range of the pattern space is covered. The consequent clause of a) of (2.1) is a fuzzy set $C = \sum_{j=1}^{n} \mu_c(c_j)/c_j$, where $\mu_c(C_j)$ denotes the degree belongingness of the population of patterns P to the class c_j, for $j = 1, 2, \ldots, n$ (see Example A.1 of Appendix-A). Therefore, by considering the conventional interpretation of a MFI, the fuzzy set D formed by the antecedent clauses of (a) of (2.1) is associated with the fuzzy set C which represents the consequent clause of (a) of (2.1). Hence, there exists a relation between D and C. More precisely, D and C are related via a certain relation $\Re$ (i.e.,$D\Re C$), which is presently unknown and has to be estimated, based on the training data set, for the design of the classifier. Now, for the testing of the classifier, we specify how C is derived from given D and estimated $\Re$. We may consider the fuzzy relational equation, namely, a direct equation

$$C = Do\Re \tag{2.2}$$

where $0 \equiv$ max–t composition operator, where t is a T—norm operator.

Equation (2.2) can be rewritten, in terms of the membership function, in the following form:

$$\mu_C(C_j) = \bigvee_{u \in U} [\mu_D(u) t \mu_R(u, c_j)] \quad \text{for } j = 1, 2, \ldots, n. \tag{2.3}$$

This explicit form of (2.3) is needed for actual design study of the classifier.

Let us assume that the training set consists of ordered pairs

$$(P_1, C_2), (P_2, C_2), \ldots, (P_\text{k}, C_\text{k})$$

and the classifier relation is supposed to specify a system of equations

$$C_l = D_l o \Re_l, \; l = 1, 2, 3, \ldots, k \tag{2.4}$$

then the fuzzy relation which satisfies (2.4) is given by

$$\widehat{\Re} = \bigcap\nolimits_{l=1}^{k} \Re_l. \tag{2.5}$$

But the above mentioned system of equations in (2.4) may not have a solution (Pedrycz 1990). Hence, in this chapter we look for an approximate solution of the system of fuzzy relation equations in (2.4).

2.3 Existing Method to Solve Fuzzy Relation Equation

The numerical solution of fuzzy relational equation has been proposed by several researchers (Wang 1993; Pedrycz 1983, 1985a, b, 1988, 1991, 1995; Ikoma et al. 1993; Hellendoorn 1992; Dinola et al. 1991; Chakraborty 1985; Ovchinnikov and Riera 1992; Gottwald 1994; Wangming 1986). In this section, we briefly review the method proposed by Pedrycz (1983). We focus our attention on max—composition operator of fuzzy relational equations, which are defined on finite spaces

$$C = Do\Re \tag{2.6}$$

where $0 \equiv \max{-t}$ composition operator, D and C are the fuzzy sets defined on the universe of discourses $U = \{u_1, u_2, \ldots, u_m\}$ and $C_{class} = \{c_1, c_2, \ldots, c_n\}$, respectively, and $\Re$ is the fuzzy relation on $U \times C_{class}$. Let $r_{ij} = (u_i,\ c_j/u_i\varepsilon U,\ c_j\varepsilon C_{class})$, $i = 1{,}2, \ldots, m$, $j = 1{,}2, \ldots, n$; then, the fuzzy sets D and C and fuzzy relation $\Re$ are as follows:

$$\begin{aligned} D &= [\mu_D(u_i)]_{1\times m} \in F(U) \\ C &= [\mu_C(c_j)]_{1\times n} \in F(C_{class}) \\ \Re &= [\mu_\Re(r_{ij})]_{m\times n} \in F(U \times C_{class}). \end{aligned} \tag{2.7}$$

If the universe U of the quantized pattern space consists of 'c' features, say F_i, $i = 1{,}2, \ldots, c$, the D is a fuzzy set defined on the quantized product spaces of U_1, $U_2, \ldots, U_c$, that is $U = U_1 \times U_2 \times \cdots \times U_c$, where $U_i = \{u_1^i, u_2^i, \ldots, u_{m_i}^i\}$ is the universe of the ith feature axis F_i with card $(U_i) = m_i$. Let D^i be the fuzzy set on U_i, i.e., $D^i = [\mu_{D^i}(u_j^i)]_{1\times m_i} \in F(U_i)$ for $i = 1, 2, \ldots, c$; then, card$(U) = m = \prod_{i=1}^{c} m_i$ and u_i is the c tuple each of type $u_i = (u_{i_1}^1, u_{i_2}^2, \ldots, u_{i_c}^c / u_{i_p}^p \in U_p, p = 1, 2, \ldots, c)$ and corresponding membership value belonging to D is determined by (2.8) shown below;

$$\begin{aligned} &(1)\ \mu_D(u_i) = \bigwedge\nolimits_{p=1}^{c} \{\mu_{D^p}(u_{ip}^p)\} = \mu_{D^1}(u_{i_1}^1) \wedge \mu_{D^2}(u_{i_2}^2) \ldots \wedge \mu_{D^c}(u_{i_c}^c) \\ &(2)\ \mu_D(U_i) = \prod\nolimits_{p=1}^{c} \{\mu_{D^p}(U_{ip}^p)\} = \mu_{D^1}(u_{i_1}^1) \cdot \mu_{D^2}(u_{i_2}^2) \ldots \mu_{D^c}(u_{i_c}^c) \end{aligned} \tag{2.8}$$

where $i = \sum_{\mathrm{p}=1}^{\mathrm{c}-1} (\Pi_{k>p}^{c} m_k)(i_p - 1) + i_c$, for each $i_p = 1, 2, \ldots, m_p$, $p = 1, 2, \ldots, c$.

Equation (2.6) can be put in the following form:

$$\mu_c(c_j) = \bigvee_{i=1}^{m} \{\mu_D(u_i) t \mu_{\Re}(r_{ij}\}, \text{ for, } j = 1, 2, \ldots, n \tag{2.9}$$

where t is the t-norm operator.

Thus, from (2.8) and (2.9), where t of (2.9) is one of the operators in {prod, min}, we get following four types of problems:

Type I: by using (1) of (2.8) and $t \equiv$ prod of (2.9);

Type II: by using (2) of (2.8) and $t \equiv$ prod of (2.9);

Type III: by using (1) of (2.8) and $t \equiv$ min of. (2.9);

Type IV: by using (2) of (2.8) and $t \equiv$ min of (2.9).

Let E be the sum of the square of the error over $p = 1, 2, \ldots, n$ and is defined by

$$E = \sum\nolimits_{p=1}^{n} \left\{\mu_C(c_p) - \mu_{\tilde{C}}(c_p)\right\}^2 \tag{2.10}$$

where C is the calculated fuzzy set using (2.9), and $\tilde{C}$ is the desired fuzzy set.

Now, the basic problem is to estimate $\Re = [\mu_{\Re}(r_{ij})]_{m\times n}$ via some given D and C which minimize E defined in (2.10) and satisfying $\{\mu_{\Re}(r_{ij}) \wedge (1 - \mu_{\Re}(r_{ij}))\} \geq 0,\ \forall i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

A general method to solve an optimization problem, defined above, is to solve a set of equations, which form the necessary conditions for a minimum of the square of the error defined in (2.10). Thus, we have $[(\partial E)/(\partial \mu_{\Re}(r_{ij}))]_{m\times n} = [0]_{m\times n}$. Now, we discuss the applicability of Newton's method and its simplification.

The Newton's iterative scheme for finding the solution of $\Re = [\mu_{\Re}(r_{ij})]_{m\times n}$ is

$$\mu_{\Re}(r_{ij})^{(s+1)} = \mu_{\Re}(r_{ij})^{(s)} - \alpha_s \cdot \frac{\partial E}{\partial \mu_{\Re}(r_{ij})} \mid \Re = \Re^{(s)} \tag{2.11}$$

where $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n \cdot \alpha_s$ is the convergent factor and also is an nonincreasing gain factor depending on the number of iteration. It can be described as $\alpha_s = 1/(2.0 + s^k) \cdot \geq 0$ is chosen empirically in order to achieve good convergent properties and avoid significant oscillations in the iteration procedure (Pedrycz 1983).

Now

$$\frac{\partial E}{\partial \mu_{\Re}(r_{ij})} = 2\{\mu_C(C_j) - \mu_{\tilde{C}}(C_j)\} P_{ij} \tag{2.12}$$

where

$$P_{ij} = \frac{\partial \mu_C(C_j)}{\partial \mu_{\Re}(r_{ij})}, \tag{2.13}$$

i.e.,

$$\begin{aligned} P_{ij} &= \frac{\partial}{\partial \mu_{\Re}(r_{ij})}\left[\bigvee\nolimits_{p=1}^{m}\{\mu_D(u_p)t\mu_{\Re}(r_{pj})\}\right] \\ &= \frac{\partial}{\partial \mu_{\Re}(r_{ij})}\left[\bigvee\nolimits_{p\neq i}^{m}\{\mu_D(u_p)t\mu_{\Re}(r_{pj})\}\right\} \times \vee\{\mu_D(u_i)t\mu_{\Re}(r_{ij})\}] \end{aligned} \tag{2.14}$$

for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

If we consider t-norm operator as "prod," then the (2.9) is written as

$$\mu_C(c_j) = \bigvee\nolimits_{i=1}^{m}\{\mu_D(u_i)\cdot\mu_{\Re}(r_{ij})\}, \quad \text{for, } j = 1, 2, \ldots, n \tag{2.15}$$

and in this case P_{ij} in (2.14) is determined as (2.16), for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

$$P_{ij} = \begin{cases} \mu_D(u_i), & \textit{if } \bigvee\nolimits_{p\neq i}^{m}\{\mu_D(u_p)\cdot\mu_{\Re}(r_{pj})\} \leq \mu_D(u_i)\cdot\mu_{\Re}(r_{ij}) \\ 0, & \textit{otherwise.} \end{cases} \tag{2.16}$$

Again, if we consider t-norm operator as "min", then (2.9) is written as

$$\mu_C(c_j) = \bigvee\nolimits_{i=1}^{m}\{\mu_D(u_i)\wedge\mu_{\Re}(r_{ij})\}, \quad \text{for, } j = 1, 2, \ldots, n \tag{2.17}$$

and in this case, P_{ij} is determined as

$$P_{ij} = \begin{cases} 1, & \textit{if } \bigvee\limits_{p\neq i}^{m}\{\mu_D(u_p)\wedge\mu_{\Re}(r_{pj})\} \leq \mu_D(u_i)\wedge\mu_{\Re}(r_{ij}) \\ & \textit{and } \mu_D(u_i) \geq \mu_{\Re}(r_{ij}) \\ 0, & \textit{otherwise} \end{cases} \tag{2.18}$$

for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

Here, the derivative of the max—function and min-function in the (2.14), (2.16), and (2.18), respectively are as follows:

$$\frac{\partial}{\partial w}(w \vee a) = \begin{cases} 1, & \textit{if } w > a \\ 0, & \textit{if } w < a \end{cases} \tag{2.19}$$

where $a = \bigvee\nolimits_{p\neq i}^{m}\{\mu_D(u_p)t\mu_{\Re}(r_{pj})\}$ and $w = \mu_D(u_i)t\mu_{\Re}(r_{ij})$ and

$$\frac{\partial}{\partial z}(z \wedge b) = \begin{cases} 1, & \textit{if } z < b \\ 0, & \textit{if } z > b \end{cases} \tag{2.20}$$

where $b = \mu_D(u_i)$ and $z = \mu_{\Re}(r_{ij})$, which are piecewise differentiable and is undefined at $w = a$ for max-function in (2.19) and $z = b$ for min-function in (2.20). Thus, we get some problems in our numerical computation (Ikoma et al. 1993) which may be overcome by defining the derivatives at $w = a$ and $z = b$, respectively as follows

$$\frac{\partial}{\partial w}(w \vee a) = \begin{cases} 1, & \text{if } w \geq a \\ 0, & \text{if } w < a \end{cases} \tag{2.21}$$

and

$$\frac{\partial}{\partial z}(z \wedge b) = \begin{cases} 1, & \text{if } z \leq b \\ 0, & \text{if } z > b \end{cases}. \tag{2.22}$$

Both formulas for the computation of the derivatives of the max and min functions, as mentioned above, return either 0 or 1 value of the derivatives. Such two-valued results of the derivatives have some inherent difficulties, in connection to the convergence of the solution as mentioned in (Ikoma et al. 1993). To overcome such difficulties there are some propositions in (Ikoma et al. 1993). In the following section, we will provide an alternative approach based on generalized functions (see Appendix-B).

The above method for solving fuzzy relational equations can be extended to simultaneous fuzzy relational equations (Pedrycz 1983) as given below.

The simultaneous fuzzy relational equations for given total k number of data in the training set are as follows:

$$C_l = D_l o \Re, \quad l, = 1, 2, \ldots, k \tag{2.23}$$

and their membership functions are as follows

$$\mu_{C_l}(c_j) = \bigvee\nolimits_{i=1}^{m} \{\mu_{D_l}(u_i) t \mu_{\Re}(r_{ij})\}, \quad \text{for } j = 1, 2, \ldots, n \tag{2.24}$$

where $l = 1, 2, \ldots, k$, and

$$\begin{aligned} D_l &= [\mu_{D_l}(u_i)]_{1\times m} \in F(U) \\ C_l &= [\mu_{C_l}(c_j)]_{1\times n} \in F(C_{class}) \\ \Re &= [\mu_{\Re}(r_{ij})]_{m\times n} \in F(U \times C_{class}). \end{aligned} \tag{2.25}$$

In this case, the error E is taken by summing over all the data set. Thus, (2.10) is modified as follows:

$$E = \sum_{l=1}^{k} \sum_{p=1}^{n} \{\mu_{C_l}(c_p) - \mu_{\tilde{C}_l}(c_p)\}^2 \tag{2.26}$$

satisfying $\{\mu_{\Re}(r_{ij}) \wedge (1 - \mu_{\Re}(r_{ij}))\} \geq 0$, $i = 1,2,\ldots,m$ and $j = 1,2,\ldots,n$ where C_l is the calculated fuzzy set using (2.24), and $\tilde{C}$ is the desired fuzzy set. The iterative scheme of (2.11) for finding the relation $\Re$ remains the same. Only the expression α_s in (2.11) could be modified as $\alpha_s = 1/(2 \times k + s^k)$, which depends on the number of data k.

2.4 Modified Approach to Solve Fuzzy Relational Equation

We modify the above said approach to solve the fuzzy relational equation (FRE) by incorporating a rigorous treatment on the computation of the derivative max-function and min-function indicated in the (2.21) and (2.22), respectively.

2.4.1 Derivative of Max-Function

Let the maximum value of $h_i,\ i = 1, 2, \ldots, s$ be determined by a function called max—function and defined by

$$h_{\max} = \bigvee\nolimits_{i=1}^{s} h_i. \tag{2.27}$$

Now, our intention is to calculate the derivative of max-function defined as above with respect to one of its variables. Hence, we transfer the said max-function of (2.27) into the following functional form

$$G(h_1,\ h_2,\ h_3,\ \ldots,\ h_s,\ h_{\max}) \equiv \sum\nolimits_{i=1}^{s} \{H(h_{\max} - h_i) - 1\} + G_e = 0 \tag{2.28}$$

where $H(h_{\max} - h_i)$ is the *Heaviside function* defined by

$$H(h_{\max} - h_i) = \begin{cases} 1, & \text{if } h_{\max} > h_i \\ 0, & \text{otherwise} \end{cases} \tag{2.29}$$

and G_e is the number of $h_i's,\ i = 1, 2, \ldots, s$ that are equal to $h_{\max}$. Also, it is a constant and independent of $h_i,\ i = 1, 2, \ldots, s$ and $h_{\max}$.

Now, by using implicit function theorem, we write

$$\frac{\partial h_{\max}}{\partial h_r} = \frac{\partial G}{\partial h_r} \Big/ \frac{\partial G}{\partial h_{\max}} \tag{2.30}$$

where $r \in \{1, 2, \ldots, s\}$.

We calculate the partial derivatives $(\partial G)/(\partial h_r)$ and $(\partial G)/(\partial h_{\max})$, using the derivative of *Heaviside function* in (B6) of Appendix-B, as follows:

$$\frac{\partial G}{\partial h_r} = -\delta(h_{\max} - h_r) \tag{2.31}$$

$$\frac{\partial G}{\partial h_{\max}} = \sum\nolimits_{i=1}^{s} \delta(h_{\max} - h_i) \tag{2.32}$$

where $\delta(*)$ is the *Dirac delta* function.

Using Eqs. (2.31) and (2.32) in (2.30), we get

$$\frac{\partial h_{\max}}{\partial h_r} = \frac{\delta(h_{\max} - h_r)}{\sum_{i=1}^{s} \delta(h_{\max} - h_i)} = \begin{cases} \frac{1}{N_{\max}}, & \text{if } h_{\max} = h_r \\ 0, & \text{otherwise} \end{cases}. \tag{2.33}$$

where $N_{\max} \triangleq$ number of terms h_i, satisfying the condition $h_{\max} = h_i,\ i = 1, 2, \ldots, s$ i.e., $N_{\max} = \sum_{i=1}^{s} \delta(h_{\max} - h_i)$, which never vanishes because at least one of $h_i^{'},\ si = 1, 2, \ldots, s$ must be equal to $h_{\max}$. So $(\partial h_{\max})/(\partial h_r)$ in (2.33) always exists everywhere.

2.4.2 Derivative of Min-Function

Let the minimum value of $h_i,\ i = 1, 2, \ldots, s$ can be determined by a function called min-function and defined by

$$h_{\min} = \bigwedge_{i=1}^{s} h_i. \tag{2.34}$$

Now, our intention is to calculate the derivative of min-function defined as above with respect to one of its variables. Hence, we transfer the said min—function (2.34) in the following functional form:

$$L(h_1, h_2, h_3, \ldots, h_s, h_{\min}) \equiv \sum_{i=1}^{s} \{H(h_i - h_{\min}) - 1\} + L_e = 0 \tag{2.35}$$

where L_e is the number of $h_i^{'}s,\quad i = 1, 2, \ldots, s$ that are equal to h_{min}. Also it is a constant independent of $h_i,\quad i = 1, 2, \ldots, s$ and $h_{\min}$.

Now, by using the implicit function theorem, we write

$$\frac{\partial h_{\min}}{\partial h_r} = -\frac{\partial L}{\partial h_r} \Big/ \frac{\partial L}{\partial h_{\min}} \tag{2.36}$$

where $r \in \{1, 2, \ldots, s\}$.

We calculate the partial derivatives $(\partial L)/(\partial h_r)$ and $(\partial L)/(\partial h_{\min})$, using the derivative of *Heaviside function* in (B6) of Appendix-B, as follows:

$$\frac{\partial L}{\partial h_r} = \delta(h_r - h_{\min}) \tag{2.37}$$

$$\frac{\partial L}{\partial h_{\min}} = -\sum_{i=1}^{s} \delta(h_i - h_{\min}). \tag{2.38}$$

Using (2.37) and (2.38) in (2.36), we get

$$\begin{aligned}\frac{\partial h_{\min}}{\partial h_r} &= \frac{\delta(h_r - h_{\min})}{\sum_{i=1}^{s} \delta(h_i - h_{\min})} \\ &= \begin{cases} \frac{1}{N_{\min}}, & \textit{if } h_{\min} = h_r \\ 0, & \textit{otherwise} \end{cases}\end{aligned} \tag{2.39}$$

where $N_{\min} \triangleq$ number of terms h_i, $i = 1, 2, \ldots, s$, satisfying the condition $h_{\min} = h_i$, i.e., $N_{\min} = \sum_{i=1}^{s} \delta(h_i - h_{\min})$, which never vanishes because at least one of $h_i^{'}s$, $i = 1, 2, \ldots, s$ must be equal to $h_{\min}$. So $(\partial h_{\min})/(\partial h_r)$ in (2.39) always exists everywhere.

Thus, from the above discussion, we understand that both the derivative of max and min functions depend on the derivative of the *Heaviside function* which is discussed, for general readability of the chapter, in the Appendix-B.

2.4.3 Modified Approach to the Computation of Derivative of Fuzzy-Max and Fuzzy-Min Functions

For the implementation of the expression of the derivative of fuzzy max and fuzzy -min functions, we approximate the *Delta function* using a finite pulse shown in Fig. 2.1. The motivation behind the approximation of the *Delta function* by a finite pulse is to incorporate the notion uncertainties built in the given data, which are all attached with fuzzy membership functions, indicating their (data) degree of possibilities to take part in any decision making process. Thus, if we approximate the *Delta function* by a finite pulse with width β, that means we try to take care of the possibilities of all the data that fall within the range of β in our computation of the derivative of a fuzzy-max and fuzzy-min functions. Using these approximations, we formulate the approximate derivative of the max and min functions, respectively, as follows:

$$\frac{\partial h_{\max}}{\partial h_r} \approx \begin{cases} \frac{1}{N_{\max}}, & \textit{if } h_{\max} - h_r \le \beta \\ 0, & \textit{otherwise} \end{cases} \tag{2.40}$$

where $N_{\max} \triangleq$ number of terms h_i, satisfying the condition $h_{\max} - h_i \le \beta$, for $i = 1, 2, \ldots, s$, and the parameter β controls the width of the pulse.

$$\frac{\partial h_{\min}}{\partial h_r} \approx \begin{cases} \frac{1}{N_{\min}}, & \textit{if } h_r - h_{\min} \le \beta \\ 0, & \textit{otherwise} \end{cases} \tag{2.41}$$

where $N_{\min} \triangleq$ number of terms h_i, satisfying the condition $h_i - h_{\min} \le \beta$, for $i = 1, 2, \ldots, s$.

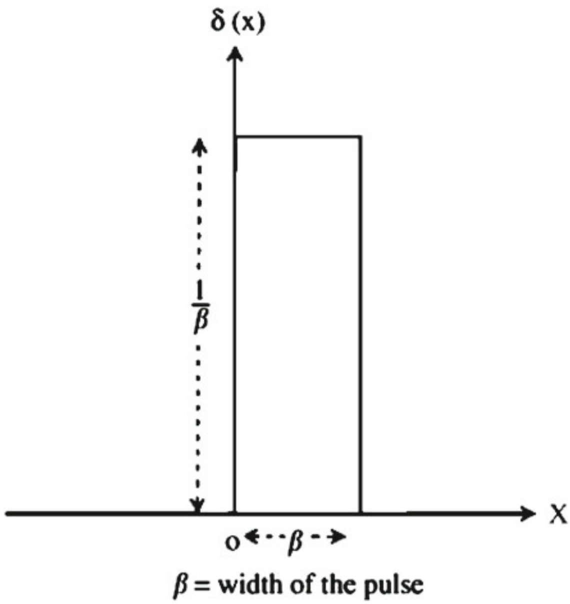

Fig. 2.1 Approximation of *Delta function* δ (x)

Now, the expression in (2.13) can be written as

$$P_{ij} = \frac{\partial \mu_C(c_j)}{\partial\{\mu_D(u_i) t \mu_{\Re}(r_{ij})\}} \cdot \frac{\partial\{\mu_D(u_i) t \mu_{\Re}(r_{ij})\}}{\partial \mu_{\Re}(r_{ij})}. \tag{2.42}$$

Comparing (2.27) with (2.15) we have $h_{\max} = \mu_C(c_j)$ and $h_i = \mu_D(u_i) \cdot \mu_{\Re}(r_{ij}),\ i = 1, 2, \ldots, m$. Using (2.40) in the above (2.42) where $t \equiv$ 'prod' we get the derivative, P_{ij} of (2.15) as

$$P_{ij} = \begin{cases} \frac{\mu_D(u_i)}{N_{\max}}, & \mathit{if}\, \mu_C(c_j) - \{\mu_D(u_i) \cdot \mu_{\Re}(r_{ij})\} \le \beta \\ 0, & \mathit{otherwise} \end{cases} \tag{2.43}$$

where $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. These results are used only for the problems of Types and II of Sect. 2.2.

Comparing (2.34) with (2.17), we have $h_{\min} = \mu_D(u_i) \wedge \mu_{\Re}(r_{ij})$. Now, there is only one variable in $h_{\min}$ as given above so $N_{\min} = 1$ only when $\mu_D(u_i) \ge \mu_{\Re}(r_{ij})$. Therefore, the derivative

$$\frac{\partial h_{\min}}{\partial \mu_{\Re}(r_{ij})} = \begin{cases} 1, & \mathit{if}\, \mu_D(u_i) \ge \mu_{\Re}(r_{ij}) \\ 0, & \mathit{otherwise} \end{cases} \tag{2.44}$$

Using (2.44) in (2.42) where $t \equiv$ 'min', we get the derivative, P_{ij} of (2.17) as

$$P_{ij} = \begin{cases} \frac{1}{N_{\max}} & \mathit{if}\, \mu_C(c_j) - \{\mu_D(u_i) \wedge \mu_{\Re}(r_{ij})\} \le \beta \\ \mathit{and} & \mu_D(u_i) \ge \mu_{\Re}(r_{ij}) \\ 0 & \mathit{otherwise} \end{cases} \tag{2.45}$$

Table 2.1 Fuzzy sets D_l^1, $D_l^2, \tilde{C}_l$, $l = 1, 2.\ldots..8$ for the fuzzy system

1	Membership values of fuzzy set										
	D_l^1			D_l^2				$\tilde{C}_l$			
1	0.00	0.27	0.80	0.97	0.11	0.00	0.00	0.00	0.35	0.80	0.11
2	1.00	0.23	0.00	0.00	0.35	1.00	0.00	0.08	0.66	0.30	0.18
3	0.00	0.05	1.00	0.08	0.70	0.30	0.00	0.00	0.03	0.15	0.70
4	0.38	0.92	0.00	0.00	0.02	0.15	0.95	0.01	0.12	0.38	0.71
5	0.37	0.48	0.00	0.00	0.10	0.30	0.90	0.02	0.20	0.37	0.37
6	0.00	0.90	0.02	0.00	0.10	0.45	0.02	0.00	0.12	0.30	0.36
7	0.20	1.00	0.08	0.00	0.12	0.50	0.01	0.02	0.13	0.33	0.40
8	0.75	0.12	0.00	0.00	0.10	0.15	0.80	0.01	0.25	0.75	0.60

Table 2.2 Fuzzy sets D_l, $l = 1, 2, \ldots\ldots 8$ for type I

1	Membership values of fuzzy set D_l											
1	0.00	0.00	0.00	0.00	0.27	0.11	0.00	0.00	0.80	0.11	0.00	0.00
2	0.00	0.35	1.00	0.00	0.00	0.23	0.23	0.00	0.00	0.00	0.00	0.00
3	0.00	0.00	0.00	0.00	0.05	0.05	0.05	0.00	0.08	0.70	0.30	0.00
4	0.00	0.02	0.15	0.38	0.00	0.02	0.15	0.92	0.00	0.00	0.00	0.00
5	0.00	0.10	0.30	0.37	0.00	0.10	0.30	0.48	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	0.10	0.45	0.02	0.00	0.02	0.02	0.02
7	0.00	0.12	0.20	0.01	0.00	0.12	0.50	0.01	0.00	0.08	0.08	0.01
8	0.00	0.10	0.15	0.75	0.00	0.10	0.12	0.12	0.00	0.00	0.00	0.00

where $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. These results are used only for the problems of Types III and IV of Sect. 2.2.

2.4.4 Algorithm for the Estimation of $\Re$

This algorithm gives the step-by-step calculation of $\Re$ using the modified computational approach.

Step (1) Start with an initial trial values of $\Re^{(0)} = [\mu_{\Re}(r_{ij})]_{m \times n}$ such that $\{\mu_{\Re}(r_{ij}) \wedge (1 - \mu_{\Re}(r_{ij}))\} \geq 0$.

Step (2) Set the width of the pulse β, convergent factor κ, the error threshold ε, and maximum number of iterations $s_{\max}$. Set the initial iteration number $s = 0$.

Step (3) Set new iteration number $s = s + 1$.

Step (4) Using the given fuzzy data D_l and $\tilde{C}_l$, evaluate C_l by (2.23) and E by (2.26).

Step (5) Evaluate $(\partial E)/(\partial \mu_{\Re}(r_{ij}))$, in (2.12) using (2.43) (for the problems of Types III and IV) for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$ and α_s.

Step (6) Update the values of $\mu_{\Re}(r_{ij})$ using the Newton's iterative scheme [see (2.11)], $\mu_{\Re}(r_{ij})^{(s+1)} = \mu_{\Re}(r_{ij})^{(s)} - \alpha \cdot \frac{\partial E}{\partial \mu_{\Re}(r_{ij})} |\Re = \Re^{(s)}$,where $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

Table 2.3 Fuzzy sets D_l, $l = 1, 2, \ldots\ldots 8$ for type II

1	Membership values of fuzzy set D_l											
1	0.00	0.00	0.00	0.00	0.26	0.03	0.00	0.00	0.77	0.09	0.00	0.00
2	0.00	0.00	0.35	1.00	0.00	0.08	0.23	0.00	0.00	0.00	0.00	0.00
3	0.00	0.00	0.00	0.00	0.00	0.03	0.01	0.00	0.70	0.70	0.30	0.00
4	0.00	0.01	0.06	0.36	0.00	0.02	0.14	0.87	0.00	0.00	0.00	0.00
5	0.00	0.04	0.11	0.33	0.00	0.05	0.14	0.43	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	0.09	0.40	0.02	0.00	0.00	0.01	0.00
7	0.00	0.02	0.10	0.01	0.00	0.12	0.50	0.01	0.01	0.01	0.04	0.00
8	0.00	0.07	0.11	0.60	0.00	0.01	0.02	0.10	0.00	0.00	0.00	0.00

Step (7) Now test whether $\{\mu_{\Re}(r_{ij}) \wedge (1 - \mu_{\Re}(r_{ij}))\} \geq 0$ or not, for all $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. If not, then construct a set of index pairs $NF = \{(i,j)/\{\mu_{\Re}(r_{ij}) \wedge (1 - \mu_{\Re}(r_{ij}))\} < 0,\ i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n\}$. Set $\{\mu_{\Re}(r_{ij}) \wedge (1 - \mu_{\Re}(r_{ij}))\} = 0,\ \forall (i,j) \in NF$.

Step (8) Repeat from Step 3 until $E < \in$ and/or $s = s_{\max}$.

2.4.5 Illustration of the Modified Approach to the Estimation of $\Re$

We illustrate the modified method based on the data set (see Table 2.1) given by Pedrycz (1983). Here, $U = U_1 \times U_2$, card $(U_1) = m_1 = 3$, and card $(U_2) = m_2 = 4$. Therefore, card $(U) = m_1 \times m_2 = 12$. Now the membership values of $D_l \in F(U)$, given by the formula $\mu_{D_l}(u_i) = (\mu_{D_l^1}(u_{i_1}^1) \wedge \mu_{D_l^2}(u_{i_2}^2))$, where $i = 4(i_1 - 1) + i_2, i_1 = 1, 2, 3$ and $i_2 = 1, 2, 3, 4$, are shown in Table 2.2, and those of D_l obtained by the formula $\mu_{D_l}(u_i) = (\mu_{D_l^1}(u_{i_1}^1) \cdot \mu_{D_l^2}(u_{i_2}^2))$, where $i = 4(i_1 - 1) + i_2$, $i_1 = 1, 2, 3$, and $i_2 = 1, 2, 3, 4$ are shown in Table 2.3. We start with an initial trial values $\mu_{\Re}(r_{ij}) = 0,\ i = 1, 2, \ldots 12$ and $j = 1, 2, 3, 4$. The value $\kappa = 10^{-4}$ is chosen to ensure good convergence properties. The width of the pulse is $\beta = 0.05$. The error threshold is $\varepsilon = 10^{-3}$, and the maximum number of iterations is s_{max} 500. The values of the error, calculated every 25 steps of iterations, are displayed in Fig. 2.2. The solutions of $\Re s$, of the problems of Types I and II of Sect. 2.2, are shown in Table 2.4.

2.5 Design of the Classifier Based on Fuzzy Relational Calculus

In classifier design (see Fig. 2.3), two phases exist, namely, the learning phase (training phase), where we estimate the fuzzy relation $\Re$ based on the algorithm of Sect. 2.4.4, and the testing phase (classification phase), where we test the performance of the classifier using (2.3) which involves the expression $\Re$.

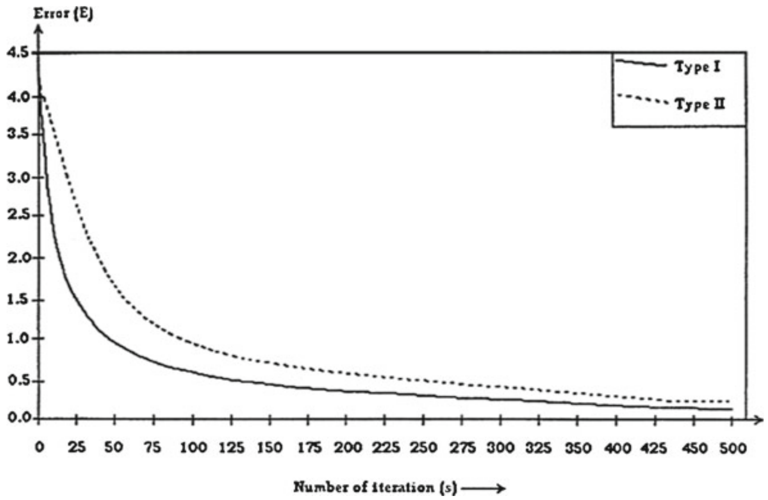

Fig. 2.2 Squared error (E) against each 25 iteration (s)

Table 2.4 Solutions of relation $\Re$ of problems type I and type II

0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.02	0.03	0.03	0.03	0.02	0.04	0.03	0.02
0.08	0.67	0.30	0.18	0.08	0.67	0.30	0.14
0.01	0.33	1.00	0.80	0.02	0.43	1.00	1.00
0.00	0.02	0.04	0.01	0.00	0.03	0.03	0.01
0.01	0.05	0.05	0.04	0.01	0.02	0.03	0.03
0.01	0.21	0.67	0.80	0.02	0.26	0.70	0.83
0.01	0.09	0.19	0.78	0.02	0.12	0.43	0.82
0.00	0.44	1.00	0.07	0.00	0.45	1.00	0.14
0.00	0.02	0.04	1.00	0.00	0.01	0.03	1.00
0.00	0.01	0.01	0.05	0.00	0.00	0.01	0.05
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Type I				**Type II**		

At the beginning of the training phase, we discretize (quantize) the individual feature axis and the entire pattern space in the following way.

Determine the lower and upper bounds of the data of ith feature value. Let f_j^i be the j th data of the ith feature F_i, and let d_i be the length of segmentation along ith feature axis.

Minimum of the data f_j^i, $j = 1, 2, \ldots$ of the ith feature is $f_{\min}^i = \min_j(f_j^i)$. Let $r_{\min}^i$ be the remainder when $f_{\min}^i$ is divided by d_i Therefore, the lower bound of the ith feature axis is

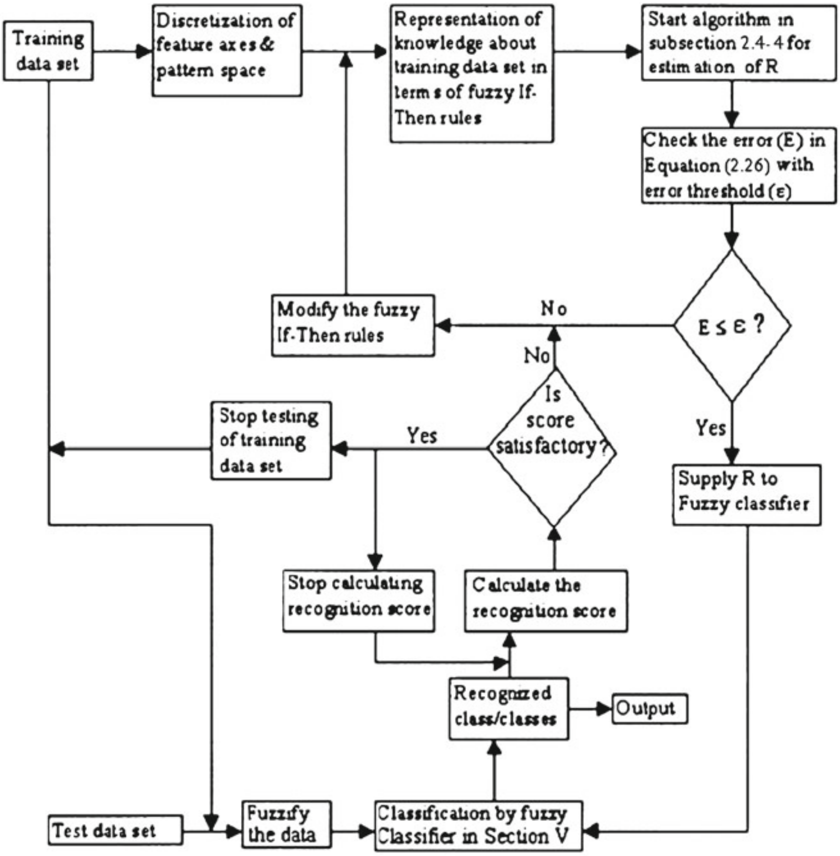

Fig. 2.3 Classifier based on fuzzy relational calculus

$$LB_i = \begin{cases} f^i_{\min}, & \textit{if } r^i_{\min} = 0 \\ f^i_{\min} - r^i_{\min}, & \textit{otherwise}. \end{cases} \tag{2.46}$$

This LB_i is taken as the ith coordinate of the origin.

Again, maximum of the data $f^i_j, j = 1, 2, \ldots$ of the ith feature is $f^i_{\max} = \max_j(f^i_j)$. Let $r^i_{\max}$ be the remainder when $f^i_{\max}$ is divided by d_i. Therefore, the upper bound of the ith feature axis is

$$UB_i = \begin{cases} f^i_{\max}, & \textit{if } r^i_{\max} = 0 \\ f^i_{\max} + (d_i - r^i_{\max}), & \textit{otherwise}. \end{cases} \tag{2.47}$$

Table 2.5 Fuzzy sets in $F\ (U_i)$

(*)	u_1^i	u_2^i	…	…	…	$u_{m_i}^i$
D_1^i	$\mu D_1^i (u_1^i)$	$\mu D_1^i\ (u_2^i)$	…	…	…	$\mu D_1^i\ \left(u_{m_i}^i\right)$
D_2^i	$\mu D_2^i\ (u_1^i)$	$\mu D_2^i\ (u_2^i)$	…	…	…	$\mu D_2^i\ \left(u_{m_i}^i\right)$
⋮	⋮	⋮	⋮			⋮
$D_{k_i}^i$	$\mu D_{k_i}^i\ (u_1^i)$	$\mu D_{k_i}^i\ (u_2^i)$	…	…	…	$\mu D_{k_i}^i\ \left(u_{m_i}^i\right)$

Let U_i be the universe of discourse on the ith feature axis F_i; then, U_i has $m_i = (UB_i - LB_i)/d_i$ generic elements and these are $u_j^i,\ j = 1, 2,\ \ldots\ldots,\ m_i$ which we define as follows

$$u_j^i = \begin{cases} [LB_i + (j-1)\cdot d_i, LB_i + j\cdot d_i)] \\ for\, j = 1, 2, \ldots, (m_i - 1) \\ [LB_i + (m_i - 1)\cdot d_i, UB_i] \\ for\, j = m_i. \end{cases} \tag{2.48}$$

Let the universe on the ith feature axis $U_i = \{u_1^i, u_2^i, \ldots, u_{m_i}^i\}$.. Let the Cartesian product space of the universe $U_i,\ i - 1, 2,\ \ldots,\ c$ be U, i.e., $U = U_1 \times U_2 \times \cdots \times U_c$ having elements each of type $u_i = (u_{i_1}^1, u_{i_2}^2, \ldots, u_{i_c}^c / u_{i_p}^p \varepsilon U_p, p = 1, 2, \ldots, c)$, where $i\ =\ \sum_{\mathrm{p}=1}^{\mathrm{c}-1} (\Pi_{l>p}^c m_l)(i_p - 1)\ + i_c$ for each $i_p = 1, 2,\ \ldots,\ m_p$, $p = 1, 2,\ \ldots,\ c$.

Now, we define k_i fuzzy sets on U_i, say, $D_j^i,\ j = 1, 2,\ \ldots,\ k_i$ which are in shown in Table 2.5. So there are $k = \Pi_{j=1}^c k_j$ fuzzy **If–Then** rules as follows:

R_l: If F_1 is $D_{j_1}^1$ and F_2 is $D_{j_2}^2$ and… F_c is $D_{j_c}^c$, then C is $C_l \in F(C_{class})$, where $l\ = \sum_{\mathrm{p}=1}^{\mathrm{c}-1}\ (\Pi_{q>p}^c k_q)(j_p - 1) + j_c$ for each j_p = 1,2,…, k_p, p = 1,2,…, c, and C_{class} is the universe of discourse constructed by all the classes in the pattern space, i.e., $C_{class} = \{c_1, c_2, \ldots, c_n\}$.

If D_l is the fuzzy set which is a fuzzy pattern vector (see Definition 1.1) formed by the antecedent clauses of the rule R_l, i.e., $D_l \in F(U)$, then the membership value of the belongingness of u_i in D_l is determined by (2.8). According to the fuzzy implication method, we write $\Re$: $F(U_1) \times F(U_2) \times \cdots \times F(U_c) \rightarrow F(C_{class})$.

The membership value of the class $c_p \in C_{class}$ when $D_{j_1}^1$ is on U_1, $D_{j_2}^2$ is on U_2, etc., is taken in the following way:

$$\mu_{C_l}(C_p) = \bigvee\nolimits_{u_i \in c_p \cap FZ(D_{j_1}^1, D_{j_2}^2, \ldots, D_{j_c}^c)} (\mu_{D_l}(u_i)) \tag{2.49}$$

Table 2.6 Fuzzy sets in $F\ (C_{class})$ for o∈'min', 'prod'}

(o)	D_1^2	D_1^2	...	...	...	$D_{k_2}^2$
D_1^1	C_{11}	C_{12}	...	...	...	C_{1k_2}
D_2^1	C_{21}	C_{22}	...	...	...	C_{2k_2}
⋮	⋮	⋮	⋮			⋮
$D_{k_1}^1$	$C_{k_1 1}$	$C_{k_1 2}$	...	...	...	$C_{k_1 k_2}$

where $FZ(D_{j_1}^1, D_{j_2}^2, \ldots, D_{j_c}^c)$ is the zone which represents the tip of the fuzzy pattern vector (see Fig. A.1 in Appendix-A) and is constructed by the fuzzy sets $D_{j_1}^1,\ D_{j_2}^2,\ \ldots,\ D_{j_c}^c$ of the rule R_l, where $j_p = 1, 2, \ldots, k_p$ for each $p = 1, 2, \ldots, c$.

For two-dimensional (2-D) pattern space, we may construct the rules in the compact form as shown in Table 2.6.

$R_{i_1 i_2}$: If F_1 is $D_{i_1}^1$ and F_2 is $D_{i_2}^2$, then C is $C_{i_1 i_2} \varepsilon F(C_{class})$ and the membership value of each class $c_p, p\varepsilon\{1, 2, \ldots, n\}$ of the fuzzy set $C_{i_1 i_2}$, where $i_l = 1, 2, \ldots, k_l$, $l = 1, 2$ will be determined by (2.49). Based on the generated fuzzy rules as stated above, we estimate the fuzzy relation $\Re$ at the end of training phase using the algorithm of the Sect. 2.4.4. In the course of estimating $\Re$, if the error given by (2.26) does not reach the desired threshold, even after a sufficient number of iterations, we may have to modify the initial fuzzy **if–then** rules to represent our knowledge about the training data set. On the other hand, after reaching the error threshold, we cross-verify the quality of the estimated $\Re$ by checking the classification score of the training data set (based on the fuzzy **if–then** rules which were initially generated for estimating $\Re$). If the classification score of the training data set (which are fuzzified by fuzzy masking at the time of testing) does not reach the satisfactory threshold (say 80 % recognition score is set as threshold), we may have to modify the initial fuzzy **if–then** rules to represent our knowledge about the training data set. After satisfactory estimation of $\Re$, we switch over to the testing phase (classification phase), where we consider the classification of data which does not belong to the training data set.

At the testing phase (classification phase), we use (2.3), as stated in Sect. 2.1. The features of the selected patterns are fuzzified using the concept of the fuzzy masking. The classification results obtained from (2.3) produces a fuzzy set $C = \sum_{i=1}^{n} \mu C(c_i)/c_i$, which represents the degree of occurrence of each test pattern at different classes in the quantized pattern space. We, thus, get a fuzzy classification of a test pattern. To calculate the recognition score from the above result, we have to go through a certain decision process. In the first stage of our decision process, we increase the level of confidence by prescribing a α-cut of the fuzzy set C, i.e.,

$$C_\alpha = \{c_i/\mu_C(c_i) \geq \alpha; c_i \in C_{class}\}.$$

If $C_\alpha = 0$ (empty set), then the given test pattern is not recognized by the present classifier. Otherwise;

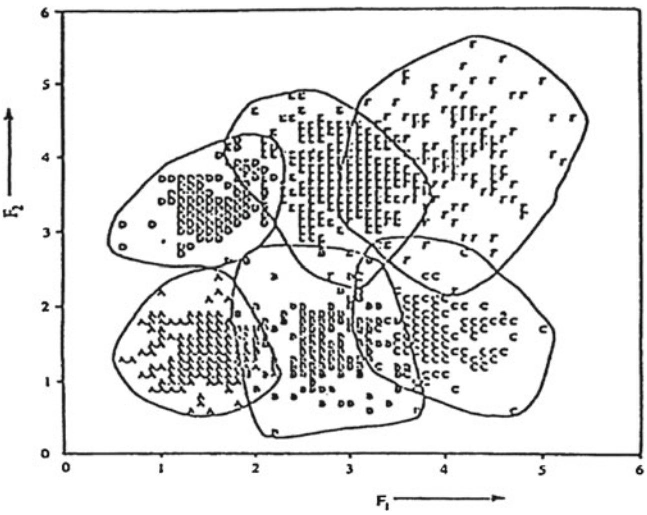

Fig. 2.4 First synthetic data

$$hgt(C) = \bigvee_{c_i \in C_\alpha} \mu_C(c_i).$$

Now, we get the set of recognized classes as

$$Class_{recognize} = \{c_i / hgt(C) - \mu_C(c_i) \leq \theta, c_i \in C_\alpha\}$$

where θ is a small threshold prescribed by the designer to capture the relative change in membership values among the elements of the recognized classes $Class_{recognize}$.

1. In case $Class_{recognize}$ is a singleton set, then the given test pattern is recognized uniquely.,
2. Otherwise, multiple classifications of the given test pattern occur.

The notion multiple classification is very natural in the case of test patterns occurring at overlapped classes. Such choice of multiple classifications sometimes stands as a kind of grace, to take care of all uncertainties (e.g., uncertainties in the representation knowledge about training patterns, uncertainties in the process of fuzzification, through fuzzy masking, of the test patterns etc.) in our classification process.

2.6 Effectiveness of the Proposed Method

To test the effectiveness of our design, as stated in Sect. 2.4, we consider the classification of two synthetic data as shown in Figs. 2.4 and 2.5. At the time of writing fuzzy **If–Then** rules for the classifier, we may consider complete cover of

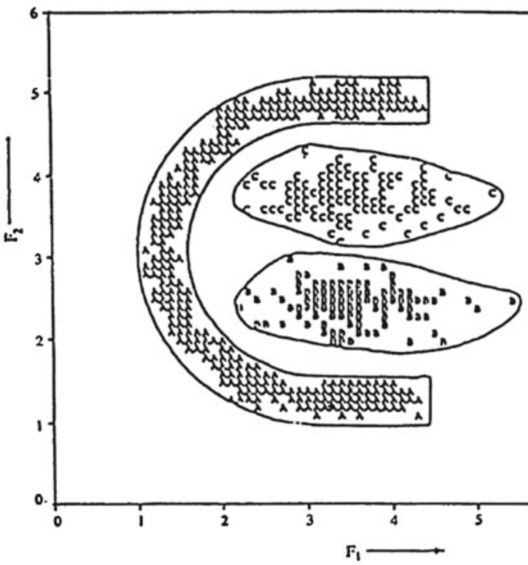

Fig. 2.5 Second synthetic data

the pattern space (see Appendix-A), but as the consideration of complete cover of the pattern space does not bring any significant change in classification score, for practical purposes, without loss of generality, we consider partial cover of the pattern space.

2.6.1 Classification of First Synthetic Data

For the data shown in Fig. 2.4, we choose length of segmentations $d_1 = 0.5 = d_2$. Therefore, we get $LB_1 = 0 = LB_2$ by (2.46) and $UB_1 = 6 = UB_2$ by (2.47). Thus, $m_1 = (UB_1 - LB_1)/d_1 = 12$ and $m_2 = (UB_2 - LB_2)/d_2 = 12$.

1. *For the Problem of Type I of* Sect. 2.2 We define $k_1 = 4$ fuzzy sets on U_1 and $k_2 = 3$ fuzzy sets on U_2 which are shown in Tables 2.7 and 2.8 respectively and $k = k_1 \times k_2 = 12$ fuzzy **If–Then** rules and their consequent parts are shown in Table 2.9.

   Now we start with initial trial values of $\mu_{\Re}(r_{ij}) = 0$, $i = 1,2,\ldots, m$ and $j = 1,2,\ldots, n$, $k = 10^{-3}$, $\beta = 0.05$, $\varepsilon = 10^{-2}$, $S_{\max} = 500$ and terminate the iteration scheme at S_{max}. The classification scores are shown in Table 2.10.

2. *For the Problem of Type II of* Sect. 2.2: We define $k_1 = 4$ fuzzy sets on U_1 and $k_2 = 4$ fuzzy sets on U_2, so we can find $k = k_1 \times k_2 = 16$ fuzzy If–Then rules.

Table 2.7 Fuzzy sets in $F\ (U_1)$ for the first synthetic data for the problem of type I

(*)	u_1^1	u_2^1	u_3^1	u_4^1	u_5^1	u_6^1	u_7^1	u_8^1	u_9^1	u_{10}^1	u_{11}^1	u_{12}^1
D_1^1	0.3	0.7	1.0	0.7	0.3	0.1	0.0	0.0	0.0	0.0	0.0	0.0
D_2^1	0.0	0.0	0.1	0.3	0.7	1.0	0.7	0.3	0.1	0.0	0.0	0.0
D_3^1	0.0	0.0	0.0	0.0	0.1	0.3	0.7	1.0	0.7	0.3	0.1	0.0
D_4^1	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.3	0.7	1.0	0.7	0.3

Table 2.8 Fuzzy sets in $F(U_2)$ for the first synthetic data for the problem of type I

(*)	u_1^2	u_2^2	u_3^2	u_4^2	u_5^2	u_6^2	u_7^2	u_8^2	u_9^2	u_{10}^2	u_{11}^2	u_{12}^2
D_1^2	0.3	0.7	1.0	0.7	0.3	0.1	0.0	0.0	0.0	0.0	0.0	0.0
D_2^2	0.0	0.0	0.0	0.1	0.3	0.7	1.0	0.7	0.3	0.1	0.0	0.0
D_3^2	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.3	0.7	1.0	0.7	0.3

Table 2.9 Fuzzy sets in $F\ (C_{class})$ for the first synthetic data for the problem of type I

Rule	Antecedent part	Consequent part					
l	F_1 *and* F_2	c_1	c_2	c_3	c_4	c_5	c_6
R_1	D_1^1 *and* D_1^2	1.00	0.70	0.00	0.10	0.10	0.00
R_2	D_1^1 *and* D_2^2	0.30	0.30	0.00	1.00	0.70	0.10
R_3	D_1^1 *and* D_3^2	0.00	0.00	0.00	0.70	0.70	0.10
R_4	D_2^1 *and* D_1^2	0.70	1.00	0.70	0.10	0.30	0.30
R_5	D_2^1 *and* D_2^2	0.30	0.70	0.70	0.70	1.00	1.00
R_6	D_2^1 *and* D_3^2	0.00	0.00	0.00	0.70	1.00	0.70
R_7	D_3^1 *and* D_1^2	0.10	1.00	1.00	0.10	0.30	0.30
R_8	D_3^1 *and* D_2^2	0.10	0.70	0.70	0.10	1.00	1.00
R_9	D_3^1 *and* D_3^2	0.00	0.00	0.00	0.10	0.70	1.00
R_{10}	D_4^1 *and* D_1^2	0.00	0.30	1.00	0.00	0.10	0.30
R_{11}	D_4^1 *and* D_2^2	0.00	0.10	0.70	0.00	0.30	1.00
R_{12}	D_4^1 *and* D_3^2	0.00	0.00	0.00	0.00	0.30	1.00

Table 2.10 Classification scores of first synthetic data for the problem of type I

From	To						Number of data	Recognition score (%)
	A	*B*	*C*	*D*	*E*	*F*		
A	109	2	0	0	0	0	116	93.97
B	7	105	16	0	1	1	117	89.74
C	0	71	71	0	1	1	79	89.87
D	0	0	0	77	10	9	89	86.52
E	0	28	0	1	135	98	151	89.40
F	0	2	0	0	64	111	115	96.52
Total	109	105	71	77	135	111		91.15

Table 2.11 Classification scores of first synthetic data for the problem of type II

From	To						Number of data	Recognition score (%)
	A	*B*	*C*	*D*	*E*	*F*		
A	114	6	0	5	4	4	116	93.28
B	10	110	18	2	3	3	117	94.02
C	0	73	73	0	1	1	79	92.41
D	0	0	0	79	12	11	89	88.76
E	0	28	0	3	137	117	151	90.73
F	0	3	1	0	65	113	115	98.26
Total	114	110	73	79	137	113		93.85

Now we start with initial trial values of $\mu_{\Re}(r_{ij}) = 0$, $i = 1,2,\ldots, m$ and $j = 1,2,\ldots, n$, $\kappa i = 10^{-3}$, $\beta = 0.05$, $\varepsilon = 10^{-2}$, $S_{\max} = 500$ and terminate the iteration scheme at S_{max}. The classification scores are shown in Table 2.11.

2.6.2 *Classification of Second Synthetic Data*

For the data shown in Fig. 2.5, we choose length of segmentations $d_1 = 0.5 = d_2$. Therefore, we get, $LB_1 = 0 = LB_2$ by (2.46) and $UB_1 = 6 = UB_2$ by (2.47). Thus, $m_1 = (UB_1 - LB_1)/d_1 = 12$ and $m_2 = (UB_2 - LB_2)/d_2 = 12$. By (2.48), we get

$$U_1 = \{u_1^1, u_2^2, \ldots, u_{12}^1\}$$
$$U_2 = \{u_1^2, u_2^2, \ldots, u_{12}^2\}.$$

For both of the problems of Types I and II of Sect. 2.2, we define $\kappa_1 = 3\,fuzzy$ sets on U_1 and $\kappa_2 = 4$ fuzzy sets on U_2 so we can fmd $\kappa = \kappa_1 \times \kappa_2 = 12$ fuzzy **If–Then** rules.

Now, for both the problems, we start with initial trial values of $\mu_{\Re}(r_{ij}) = 0$, $i = 1,2,\ldots, m$ and $j = 1,2,\ldots, n$, $\kappa = 10^{-3}$, $\beta = 0.05$, $\varepsilon = 10^{-2}$, $S_{max} = 500$ and terminate the iteration scheme at S_{max}. The classification scores are shown in Table 2.12.

2.7 Applications

After achieving satisfactory results on a synthetic set of data, we apply the proposed design for the vowel classification problem of an Indian language, namely Telugu.

Table 2.12 Classification scores of second synthetic data

(a) Problem of Type I

From	To			Number of data	Recognition score (%)
	A	*B*	*C*		
A	287	13	6	287	100.00
B	8	96	0	97	98.97
C	12	0	99	100	99.00
Total	287	96	99		99.59

(b) Problem of Type II

From	To			Number of data	Recognition score (%)
	A	*B*	*C*		
A	283	43	38	287	98.61
B	2	96	0	97	98.97
C	3	0	99	100	99.00
Total	283	96	99		98.76

2.7.1 Experimental Results

A number of discrete phonetically balanced speech samples for the vowels of Telugu language in CNC (consonant–vowel nucleus-consonant) form are selected. A CNC combination is taken because the form of consonants connected to a vowel is responsible for influencing the role and quality of vowels. These speech units are recorded by five informants on an AKAI type recorder. The spectrographic analysis has been done on Kay Sona-graph Model 7029-A, which is a very standard audio frequency spectrum analyzer that produces a permanent record of the spectra of any complex waveform in the range of 5 Hz to 16 kHz. For the present study of vowels, the spectrographic representation of frequency versus time has been done for 800 Telugu words uttered by three male informants in the age group of 25–30 year chosen from 15 educated persons. The total bandwidth of the system is 80 Hz to 8 kHz with a resolution of 300 Hz. The experiment deals with the formant frequencies at the steady state of the Telugu vowels and their variations on different consonantal context and for different speakers. The average positions (see Table 2.13) of different Telugu vowels with respect to cardinal vowels and their distribution in F_1–F_2 frequency planes are considered (see Fig. 2.6). The present investigation has been carried out with the Telugu vowels (listed in Table 2.13) both short and long. It is well known that the first three formant frequencies carry most of the information regarding the vowel quality. But for all practical purposes of vowel classification, we can use the first two formant frequencies, i.e., F_1 and F_2.

In the following, we discuss the classification results for Telugu vowels.

For the data shown in Fig. 2.6, we choose length of segmentations $d_1 = 50$ and $d_2 = 100$. Therefore, we get, $LB_1 = 200$ and $LB_2 = 600$ by (2.46) and $UB_1 = 850$ and $UB_2 = 2600$ by (2.47). Thus, $m_1 = (UB_1 - LB_1)/d_1 = 13$ and $m_2 = (UB_2 - LB_2)/d_2 = 20$. By (2.48), we get

Table 2.13 Average formant frequencies of Telugu vowel

Phonetic symbol	Average format frequencies (Hz)		
	F_1	F_2	F_3
/∂/	606	1,473	2,420
/a:/	710	1,240	2,400
/i⊢/	365	2,116	2,757
/i:/	325	2,260	2,836
/u⊣/	370	1,066	2,500
/u:/	348	923	2,543
/e⊢/	517	1,796	2,633
/e:/	470	1,883	2,633
/o⊣/	476	1,133	2,630
/o:/	486	1,000	2,540
/ae/	575	1,744	2,700

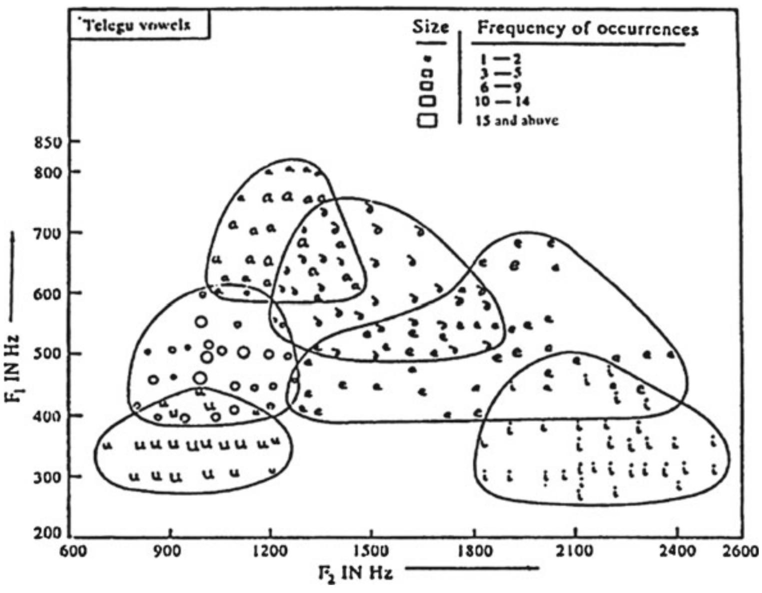

Fig. 2.6 Telugu vowels

$$U_1 = \{u_1^1, u_2^1, \ldots, u_{13}^1\}$$
$$U_2 = \{u_1^2, u_2^2, \ldots, u_{20}^2\}.$$

For both the problems of Types I and II of Sect. 2.2, we define $\kappa_1 = 5$ fuzzy sets on U_1 and $\kappa_2 = 7$ fuzzy sets on U_2, so we can fmd $\kappa = \kappa_1 \times \kappa_2 = 35$ fuzzy **If–Then** rules.

Table 2.14 Classification scores of Telugu vowel for the problem of type I

From	To						Number of data	Recognition score (%)
	a	e	i	o	u	δ		
a	82	4	0	29	0	74	83	98.80
e	7	193	52	21	7	82	200	96.50
i	0	48	119	0	0	0	133	89.47
o	15	26	0	114	72	30	116	98.28
u	0	10	0	48	107	2	112	95.54
δ	41	32	0	19	1	65	66	98.48
Total	82	193	119	114	107	65		95.77

Table 2.15 Classification scores of Telugu Vowel for the problem of type II

From	To						Number of data	Recognition score (%)
	a	e	i	o	u	∂		
a	82	4	0	29	0	74	83	98.80
e	7	196	52	21	7	83	200	98.00
i	0	50	126	0	0	0	133	94.74
o	16	26	0	114	72	30	116	98.28
u	0	10	0	48	108	2	112	96.43
∂	40	32	0	19	1	65	66	98.48
Total	82	196	126	114	108	65		97.32

Now, for both the problems, we start with initial trial values of $\mu_{\Re}(r_{ij}) = 0, 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, $\kappa, = 10^{-3}$, $\beta =, 0.05$, $\varepsilon = 10^{-2}$, $S_{\max} = 1000$ and terminate the iteration scheme at S_{max}. The classification scores are shown in Tables 2.14 and 2.15.

2.8 Comparative Study

In Table 2.16 we have compared the performance (in terms of recognition score) of the present classifier with those of some existing ones. The results shown in Table 2.16 indicate that the performance of the present design of the classifier is comparable with those of some existing ones.

2.9 Conclusion

In this chapter, we consider a particular interpretation [i.e., (a) of (2.1)] of MFI and introduce a notion of fuzzy pattern vector which represents the antecedent part of the interpretation (a) of (2.1). The advantage of considering such notion is two-fold. First, we can describe a population of training patterns by linguistic features.

Table 2.16 Comparative study

Different types of classifier	Recognition score (%)		
	First synthetic data (Fig. 2.4)	First synthetic data (Fig. 2.5)	First synthetic data (Fig. 2.6)
Bayesian	78.0	80.9	80.3
Support vector machine (SVM)	86.13	87.2	85
Conventional multilayer perception (MLP)	88.15	78.0	90.0
Present method with max–min operator	91.15	99.59	95.77
Present method with max-product operator	93.85	98.76	97.32

Second, the notion of fuzzy pattern vector helps us formulate the consequent part of (a) of (2.1) (see Example A.1 of Appendix-A). We develop a new approach to the computation of the derivative of the fuzzy max/min function. A detail design of pattern classifier based on FRC is developed and very promising results are obtained. We compute the performance of the present classifier with those of some existing classifiers and get satisfactory response. A neural net (Pao 1989) version of the present design to estimate the fuzzy relation $\Re$ (for classification problem) would be the scope for future work. In the present design study we have only considered the problems of Types I and II of Sect. 2.2. Similar results are also obtainable for the problems of Types III and IV.

References

J. C. Bezdek, S. K. Pal (eds.), *Fuzzy Models for Pattern Recognition: Methods that Search for Structures in Data* (IEEE, New York, 1992)

G. Bortolan, R. Degani, Ranking of fuzzy alternatives in electrocardiography, in *Fuzzy Information, Knowledge Representation and Decision, Analysis*, ed. by E. Sanchez, M. M. Gupta (Pergamon, Oxford, 1983), pp. 397–402

G. Bortolan, R. Degani, K. Hirota, W. Pedrycz Classification of ECG Signals-utilization of fuzzy pattern matching, in *Proceedings International Workshop Fuzzy System* (Applic, Iizuka, 1988)

M. K. Chakraborty, Some aspects of [0, 1] fuzzy relations and a few suggestions toward its use. Approx. Reas. Expert Syst. pp. 139–157 (1985)

R. Degani, G. Bortolan, Computerized electrocardiogram diagnosis: Fuzzy approach, in *Encyclopedia of Systems and Control*. (Pergamon, Oxford, 1987)

R. Degani, G. Bortolan, Fuzzy numbers in computerized electrocardiography. Fuzzy Sets Syst. **24**, 345–362 (1987)

A. Dinola, W. Pedrycz, S. Sessa, E. Sanchez, Fuzzy relation equations theory as a basis of fuzzy modeling: An overview. Fuzzy Sets Syst. **40**, 415–429 (1991)

D. Dubois, M. C. Jaulent, Techniques for extracting fuzzy regions, in *First IFSA Congress*, vol. II (Mallorca, Spain, 1985), 1–6 July

S. Gottwald, Approximately solving fuzzy relation equations: Some mathematical results and some heuristic proposals. Fuzzy Sets Syst. **66**, 175–193 (1994)
H. Hellendoorn, The generalized modus ponens considered as a fuzzy relation. Fuzzy Sets Syst. **48**, 29–48 (1992)
K. Hirota, Fuzzy robot vision and fuzzy controlled robot, in *NATO ASI CIM*, ed. by I. B. Turksen (Springer, Berlin, 1988)
K. Hirota, K. Iwami, W. Pedrycz, FCM-AD (fuzzy cluster means with additional ata) and its application to aerial images, in *Proc. II IFSA Congress*, vol. II (Tokyo, Japan, 1987), pp. 729–732
T. L. Huntsberger, C. L. Jacobs, R. L. Canon, Iterative fuzzy image segmentation. Pattern Recognit. **18**, 131–138 (1985)
T. L. Huntsberger, C.H. Rangarajan, S.N. Jayaramurthy, Representation of uncertainty in computer vision using fuzzy sets. IEEE Trans. Comput. **C–2**, 145–156 (1986)
N. Ikoma, W. Pedrycz, K. Hirota, Estimation of fuzzy relational matrix by using probabilistic descent method. Fuzzy Sets Syst. **57**, 335–349 (1993)
W. J. Kickert, H. Koppleaar, Application of fuzzy set theory to syntactic pattern recognition of handwritten capitals. IEEE Trans. Syst. Man Cybern. **SMC–6**, 148–151 (1986)
A. Kumar, A real-time system for pattern recognition of human sleep stages by fuzzy systems analysis. Pattern Recognit. **9**, 43–46 (1977)
E. T. Lee, Proximity measure for the classification of geometric figures. J. Cybern. **2**, 43–59 (1972)
M. Mizumoto, Extended fuzzy reasoning, in *Approximate Reasoning in Expert Systems*, ed. by M. M. Gupta, A. Kandel, W. Bandler, J. B. Kiszka (North-Holland, Amsterdam, 1985), pp. 71–85
R. Di Mori, *Computerized Models of Speech Using Fuzzy Algorithms* (Plenum, New York, 1983)
R. Di Mori, P. Laface, Use of fuzzy algorithms for phonetic and phonetic labeling of continuous speech. IEEE Trans. Pattern Anal. Machine Intell. **2**, 136–148 (1980)
S. V. Ovchinnikov, T. Riera, On fuzzy classifications. Fuzzy Sets Syst. **49**, 119–132 (1992)@@@@
Y. H. Pao, *Adaptive Pattern Recognition and Neural Networks* (Addison Wesley Publishing Company, Boston, 1989)
W. Pedrycz, Numerical and applications aspects of fuzzy relational equation. Fuzzy Sets Syst. **11**, 1–18 (1983)
W. Pedrycz, Applications of fuzzy relational equations for methods of reasoning in presence of fuzzy data. Fuzzy Sets Syst. **16**, 163–174 (1985a)
W. Pedrycz, On generalized fuzzy relational equations and their applications. J. Math. Anal. Appl. **107**, 520–536 (1985b)
W. Pedrycz ECG Signal classification with the aid of linguistic classifier, in *Proceedings XIV International Conference Medicine Biomedical Engineering*, (Spain, 1985c) pp. 11–16 Aug
W. Pedrycz, Approximate solution of Fuzzy relational equations. Fuzzy Sets Syst. **28**, 183–201 (1988)
W. Pedrycz, Fuzzy sets in pattern recognition methodology and methods. Pattern Recogn. **23**(1/2), 121–146 (1990)
W. Pedrycz, Processing of relational structures: Fuzzy relational equations. Fuzzy Sets Syst. **40**, 77–106 (1991)
W. Pedrycz, Genetic algorithms for learning in fuzzy relational structures. Fuzzy Sets Syst. **69**(1), 37–52 (1995)
L. Saitta, P. Tarasso, Fuzzy characteristics of coronary disease. Fuzzy Sets Syst. **5**, 245–258 (1981)
A. Seif, J. Aguilar-Martin, Multi-group classification using fuzzy correlation. Fuzzy Sets Syst. **3**, 109–122 (1980)
M. Shimura, Applications of fuzzy set theory to pattern recognition. J. JAACE **19**, 243–248 (1975)

P. K. Simpson, Fuzzy min-max neural network-Part 1: Classification. IEEE Trans. Neural Networks **13**, 776–786 (1992)
P. Siy, C. S. Chen, Fuzzy logic for handwritten numerical character recognition. IEEE Trans. Syst. Man Cybern **SMC–4**, 570–575 (1974)
M. Sugeno, T. Takagi, Multidimensional fuzzy reasoning. Fuzzy Sets Syst. **9**, 313–325 (1983)
Y. Tsukamoto, An approach to fuzzy reasoning method, in *Advance in Fuzzy Set Theory and Applications*, ed. by M. M. Gupta, R. K. Ragade, R. R. Yager (North-Holland, Amsterdam, 1979), pp. 137–149
H. F. Wang, Numerical analysis on fuzzy relation equations with various operators. Fuzzy Sets Syst. **53**, 155–166 (1993)
W. U. Wangming, Fuzzy reasoning and fuzzy relational equations. Fuzzy Sets Syst. **20**, 67–78 (1986)
S. Watanabe, *Pattern Recognition: Human and Mechanical* (Wiley, New York, 1985)
M. A. Woodbury, J. Clive, Clinical pure types as fuzzy partition. J. Cybern. **3**, 111–121 (1974)
L. A. Zadeh, Theory of Approximate Reasoning, in *Machine Intelligence*, ed. by D. Michie, L. I. Mikulich, J. E. Hayes (Ellis Horwood, Chichester, 1970), pp. 149–194
L. A. Zadeh, K. S. Fu, K. Tanaka, M. Shimura (eds.), *Fuzzy Sets and Their Applications to Cognitive and Decision Processes* (Academic, New York, 1975)

Soft Computing Approach to Pattern Classification and
Object Recognition
A Unified Concept
Ray, K.S.
2012, XII, 173 p. 46 illus., Hardcover
ISBN: 978-1-4614-5347-5