ELSEVIER International Journal of Approximate Reasoning 20 (1999) 21–45

INTERNATIONAL JOURNAL OF APPROXIMATE REASONING

A proposal on reasoning methods in fuzzy rule-based classification systems [1]

Oscar Cordón [a,2], María José del Jesus [b,3], Francisco Herrera [a,*]

[a] *Department of Computer Science and Artificial Intelligence, University of Granada, ETS Ingenieria Informatica, Avenida Andalucia 38, 18071 Granada, Spain,*

[b] *Department of Computer Science, University of Jaén, 23071 Jaén, Spain*

Received 1 December 1997; accepted 1 August 1998

Abstract

Fuzzy Rule-Based Systems have been succesfully applied to pattern classification problems. In this type of classification systems, the classical Fuzzy Reasoning Method (FRM) classifies a new example with the consequent of the rule with the greatest degree of association. By using this reasoning method, we lose the information provided by the other rules with different linguistic labels which also represent this value in the pattern attribute, although probably to a lesser degree. The aim of this paper is to present new FRMs which allow us to improve the system performance, maintaining its interpretability. The common aspect of the proposals is the participation, in the classification of the new pattern, of the rules that have been fired by such pattern. We formally describe the behaviour of a general reasoning method, analyze six proposals for this general model, and present a method to learn the parameters of these FRMs by means of Genetic Algorithms, adapting the inference mechanism to the set of rules. Finally, to show the increase of the system generalization capability provided by the proposed FRMs, we point out some results obtained by their integration in a fuzzy rule generation process.

Keywords: Fuzzy rule based classification systems; Fuzzy reasoning; Genetic algorithms

* Corresponding author. Tel.: +34 58 24 4019; fax: +34 58 24 3317; e-mail: herrera@decsai.ugr.es.

[1] This research has been supported by CICYT TIC96-0778.

[2] E-mail: ocordon@decsai.ugr.es

[3] E-mail: mjjesus@ujaen.es

0888-613X/99/$ – see front matter
PII: S0888-613X(98)10021-X

1. Introduction

Fuzzy Rule-Based Systems (FRBSs) have been succesfully applied to pattern classification problems ([4,5,7]). The interest in using FRBSs arises from the fact that they provide a good platform to deal with noisy, imprecise or incomplete information which is often handled in any human-cognition system. They are an effort to reconcile the empiric precision of traditional engineering techniques and the interpretability of Artificial Intelligence.

In classification problems, the fundamental role of fuzzy rules is to make the opaque classification schemes, as usually used by a human, transparent in a formal and computer-realizable framework ([30]). Therefore, Fuzzy Rule-Based Classification Systems (FRBCSs) may be assigned two classes of Classification Systems: those which are supposed to work autonomously, and those which are intended to be tools in the hands of the user to help him to take decisions. In the former case, the performance level may be the answer, but in the latter, other dimensions such as comprenhensibility, robustness, versatility, modificability and coherence with previous knowledge may be fundamental in order to allow the system to be accepted for use. The second kind of properties are associated to the fuzzy rule structure. The first one, the performance level, also depends on the Fuzzy Reasoning Method (FRM) employed. The FRM has an important role in FRBCSs in order to find the highest performance level.

As is well known in FRBCSs, the classical FRM, maximum matching, classifies a new example with the consequent of the rule with the greatest association degree ([1,7,15,19,21,22,25,26]). Using this inference method, we lose the information provided by the other fuzzy rules with different linguistic labels which also represent the value in the pattern attribute, although probably to a lesser degree. On the other hand, in fuzzy control it is well-known that the best performance is obtained when we use defuzzification methods that operate on the fuzzy subsets obtained by the fuzzy rules fired ([9]). As regards FRBCSs, in Refs. [4,6,7,20] the information provided for all rules belonging to the set of rules is used for a classification problem.

The aim of this paper is to present new FRMs which allow us to improve the system performance, maintaining its interpretability. The common aspect of the proposals is that all rules fired by a pattern participate in the classification of such pattern. We describe formally the behaviour of a general reasoning method, analyzing six specific proposals for this general model. We present a method to learn the parameters of these FRMs by means of Genetic Algorithms, adapting the inference mechanism to the set of rules, so improving the performace of the FRBCSs. Finally, we point out some results obtained by the integration of the FRMs in a fuzzy rule generation process.

This paper is organized as follows. Section 2 briefly reviews FRBCSs, and describes the different structures for the fuzzy rules. Section 3 presents the classical reasoning method, a general model of fuzzy reasoning, the alternative proposals for FRMs, and some experiments carried out, showing the good behavior of the latter reasoning methods. Section 4 describes a genetic learning algorithm for inference parameters, shows some experiments with this proposal, and introduces some results obtained by the integration of the proposed FRMs in a fuzzy rule generation process. Finally, some concluding remarks are presented in Section 5.

2. Fuzzy rule-based classification systems

Pattern classification problems involve assigning a class C_j from a predefined class set $C = \{C_1, \ldots, C_M\}$ to an object, described as a point in a certain feature space $x \in S^N$.

The problem of designing a classifier is to find a mapping

$$D : S^N \to C$$

optimal in the sense of a certain criterion $\delta(D)$ that determines the classifier performance. Usually, the final goal is to design a classifier that assigns class labels with the smallest possible error across the whole feature space. The classifier may be a set of fuzzy rules, a neural network, a decision tree, etc. When the classifier is a set of fuzzy rules, the resulting system is called a Fuzzy Rule-Based Classification System (that we have denoted by FRBCS).

An FRBCS is composed of a Knowledge Base (KB) and an FRM. The KB is made up by the Rule Base (RB) and the Data Base (DB) that describes the semantic of the Fuzzy subsets associated to the linguistic labels in the if-part of the rules. The FRM uses the information from the KB to determine a label class for all admisible patterns. This structure is shown in Fig. 1.

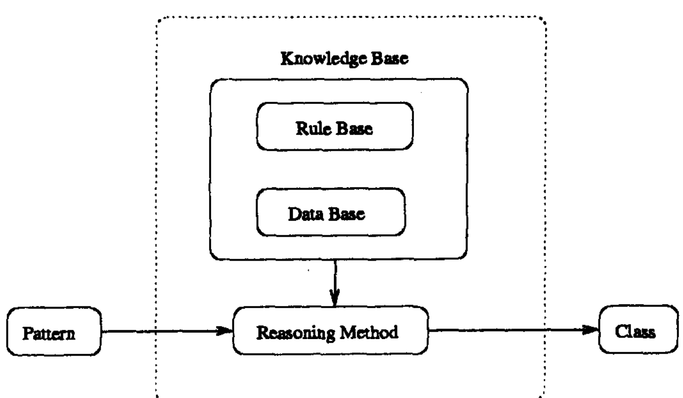

Fig. 1. Fuzzy rule-based classification system.

To build an FRBCS we begin with a set of preclassified examples, from which we must determine:

- the method to find or learn a set of fuzzy rules for the specific classification problem, and
- the fuzzy reasoning method used to classify a new pattern.

Many researches have proposed various methods for generating fuzzy rules from numerical data pairs ([1,6–8,13,15,18,19,22,24,26]). In this paper, we will use an FRBCS constituted by an RB generated by the technique that extends the Wang and Mendel algorithm ([27]) to fuzzy classification rules ([6,7]). In Appendix A, we briefly review this learning method.

For the description of the FRMs, we explain the three types of fuzzy rules that may be used to build the RB in Section 2.1.

2.1. Types of fuzzy rules

We can generate RBs with one of the following three types of rules:

(a) *Fuzzy rules with a class in the consequent* [1,15]. This kind of rules has the following structure:

$$R_k\text{: If } x_1 \text{ is } A_1^k \text{ and } \ldots \text{ and } x_N \text{ is } A_N^k \text{ then } Y \text{ is } C_j,$$

where $x_1, \ldots, x_N$ are the outstanding selected features for the classification problem, $A_1^k, \ldots, A_N^k$ are linguistic labels used to discretize the continuous domain of the variables, and Y is the class C_j to which the pattern belongs.

(b) Fuzzy *rules with a class and a certainty degree in the consequent* [18].

$$R_k\text{: If } x_1 \text{ is } A_1^k \text{ and } \ldots \text{ and } x_N \text{ is } A_N^k \text{ then } Y \text{ is } C_j \text{ with } r^k,$$

where r^k is the certainty degree of the classification in the class C_j for a pattern belonging to the fuzzy subspace delimited by the antecedent. This certainty degree can be determined by the ratio

$$\frac{S_j^k}{S^k},$$

where, considering the matching degree as the compatibility degree between the rule antecedent and the pattern feature values,

- S_j^k is the sum of the matching degrees for the class C_j patterns belonging to the fuzzy region delimited by the antecedent, and
- S^k the sum of the matching degrees for all the patterns belonging to this Fuzzy subspace, regardless its associated class.

(c) *Fuzzy rules with certainty degree for all classes in the consequent* [22,24].

$$R_k\text{: If } x_1 \text{ is } A_1^k \text{ and } \ldots \text{ and } x_N \text{ is } A_N^k \text{ then } (r_1^k, \ldots, r_M^k),$$

where r_j^k is the soundness degree for the rule k to predict the class C_j for a pattern belonging to the fuzzy region represented by the antecedent of the rule.

This degree of certainty can be determined by the same ratio as the type (b) rules.

The last type of rule extends type (a) and type (b) using different values for $(r_1^k, \ldots, r_M^k)$. Considering

$$r_h^k = 1, \quad r_j^k = 0, \quad j \neq h, \quad j = 1, \ldots, M,$$

we have the first case, and with

$$r_h^k = r^k, \quad r_j^k = 0, \quad j \neq h, \quad j = 1, \ldots, M$$

we have the second one.

From this point, to develop the theoretical model of the *reasoning method* in an FRBCS, we will work with type (c) rules. As regards the experiments, we will consider the three types of rules.

3. Fuzzy reasoning methods

As was mentioned earlier, an FRM is an inference procedure that derives conclusions from a set of fuzzy if-then rules and a pattern. The power of fuzzy reasoning is that we can achieve a result even when we do not have an exact match (with degree 1) between a system observation and the antecedents of the rules.

The use of a reasoning method that combines the information of the rules fired with the pattern to be classified can improve the generalization capability of the classification system. We will analyze this idea in this section according to the following structure. First, we describe the classical FRM. After that, we present a general model of reasoning that involves different possibilities as reasoning methods, and we propose six alternative FRMs as some particular new proposals inside the general reasoning model. Finally, in the last section we present the experiments carried out, showing the good behaviour of the alternative reasoning methods.

3.1. Classical fuzzy reasoning method: maximum matching

Suppose that the RB is $R = \{R_1, \ldots, R_L\}$ and there are L_j rules in R that produce class C_j. Clearly in an RB type c), $L_j = L \quad \forall j = 1, \ldots, M$.

For a pattern $E^t = (e_1^t, \ldots, e_N^t)$, *the classical fuzzy reasoning method* considers the rule with the highest combination between the matching degree of the pattern with the if-part and the certainty degree for the classes. It classifies E^t with the class that obtains the highest value. This procedure is described in the following steps.

- Let $R^k(E^t)$ denote *the strength of activation of the if-part of the rule k* (*matching degree*). Usually $R^k(E^t)$ is obtained by applying a t-norm to the degree of satisfaction of the clauses (x_i is A_i^k)

$$R^k(E^t) = T(\mu_{A_1^k}(e_1^t), \ldots, \mu_{A_N^k}(e_N^t)).$$

In the literature, there are some proposals for this conjunction operator; some of them are the following: Ishibuchi et al., uses the t-norm product and minimum ([18]), González et al., uses the minimum ([15]), and Mandal et al. ([22]) and Uebele et al. ([26]) calculate the matching degree by means of the arithmetic mean (which is not a t-norm).

- Let $h(R^k(E^t), r_j^k)$ denote *the degree of association of the pattern with class* C_j *according to the rule* k. This degree is obtained by applying a combination operator between $R^k(E^t)$ and r_j^k. For this operator, in Ref. [4] are used the minimum and product ones, in Ref. [18] the product and in Ref. [22] the arithmetic mean.
- For each class C_j, *the degree of association of the pattern with the class*, Y_j, is calculated

$$Y_j = \max_{k \in L_j} h(R^k(E^t), r_j^k).$$

This degree of association is a soundness degree of the classification of the pattern E^t in class C_j.

- The classification for the pattern E^t is the class C_h such

$$Y_h = \max_{j=1,\ldots,M} Y_j.$$

Graphically, this method could be seen as shown in Fig. 2.

This reasoning method uses only one rule -the winner rule- in the inference process and wastes the information associated to all those rules whose association degree with the input pattern is lower than the association degree of the selected rule. If we only consider the rule with the highest association degree, we would only be considering one fuzzy subset for each value of the attributes, and we would not take into account the information of other rules with other

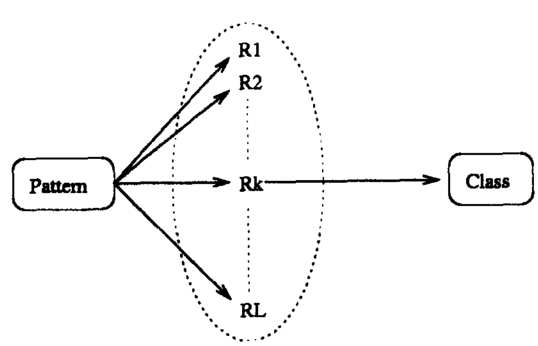

Fig. 2. Fuzzy reasoning method that uses only the winner rule.

fuzzy subsets which also have the value of the attribute on its support but to a lesser degree. Working in this way, *we would be considering an interval discretization in a masked way.*

3.2. General Model of Fuzzy Reasoning

In this section, we present a common and general model of fuzzy reasoning to combine information provided by different rules. This model is an extension of the fuzzy classifier definition presented by Kuncheva in Ref. [21].

Considering a new pattern $E^t = (e_1^t, \ldots, e_N^t)$, the steps of the general reasoning model are the following:

1. *Matching degree.* To calculate the *strength of activation of the if-part for all rules in the RB with the pattern E^t*, using a t-norm ([2,11]).

   $$R^k(E^t) = T(\mu_{A_1^k}(e_1^t), \ldots, \mu_{A_N^k}(e_N^t)), \quad k = 1, \ldots, L.$$

2. *Association degree.* To calculate the *association degree of the pattern E^t with the M classes according to each rule in the RB.*

   $$b_j^k = h(R^k(E^t), r_j^k), \quad j = 1, \ldots, M, \quad k = 1, \ldots, L.$$

3. *Weighting function. To weight* the obtained values, through a function g. A possibility is to increase the higher values of the association degree and penalize the lower ones.

   $$B_j^k = g(b_j^k), \quad j = 1, \ldots, M, \quad k = 1, \ldots, L.$$

4. *Pattern classification soundness degree for all classes.* We use an aggregation function ([2,11]) that combines -for each class- the positive degrees of association calculated in the previous step and produces a system soundness degree for the classification of the pattern in this class.

   $$Y_j = f(B_j^k,\ k = 1, \ldots, L \text{ and } B_j^k > 0), \quad j = 1, \ldots, M.$$

   with f being an aggregation operator verifying $\min \leqslant f \leqslant \max$. It is clear that if we select f as the maximum operator, we have the classical FRM.
5. *Classification.* We apply a decision function F over the soundness degree of the system for the pattern classification for all classes. This function will determine the class label l corresponding to the maximum value.

   $$C_l = F(Y_1, \ldots, Y_M) \quad \text{such as} \quad Y_l = \max_{j=1,\ldots,M} Y_j.$$

The general fuzzy reasoning model is represented graphically in Fig. 3.

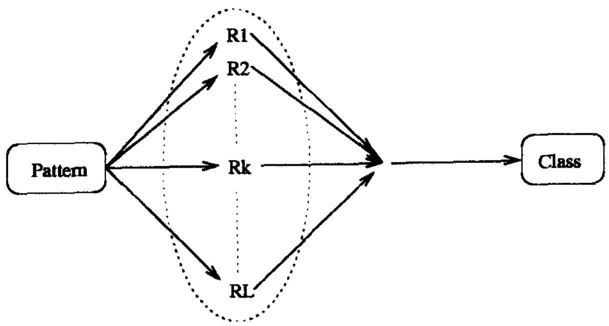

Fig. 3. General fuzzy reasoning model.

3.3. Alternative fuzzy reasoning methods

According to the formulation that we have stated above, there are four steps that we must define to determine a fuzzy reasoning method in an FRBCS. In the following, we present the proposal for the alternative FRMs.

The first two components of the proposals are used classically.

(1) *Pattern matching degree with the if-part of a rule k.* It will be calculated by the min t-norm.

$$R^k(E^t) = \min_{i=1,\dots,N} \mu^k_{A_i}(e^t_i),$$

(2) *Association degree.* To obtain a global response from the rule between the pattern and one class, we define function h such as

$$h(R^k(E^t), r^k_j) = R^k(E^t) r^k_j.$$

The following two steps establish the new fuzzy reasoning proposals according to the general reasoning model.

(3) *Weighting function.* We consider two weighting functions,

$$g_1(x) = x, \quad \forall x \in [0, 1],$$

that leaves the degrees of association without weighting, and

$$g_2(x) = \begin{cases} x^2 & \text{if } x < 0.5, \\ \sqrt{x} & \text{if } x \geqslant 0.5, \end{cases}$$

that favours the degrees of association greater than 0.5 and penalizes the ones under than this value.

(4) *Aggregation function.* Let $(a_1, \dots, a_s)$ be the positive weighted degrees of association for the pattern E^t and the class C_j, according to the rules in the RB, that is

$$(a_1, \dots, a_s) = (B^k_j > 0, k = 1, \dots, L),$$

where s is the number of positive elements for that class.

As we mentioned above, if the aggregation function is the maximum, the FRM is the classical one. We use some alternative aggregation functions that produce an aggregated value between the minimum and the maximum. The functions considered are:

(a) *Normalized addition.*

$$f_1(a_1,\dots,a_s) = \frac{\sum_{i=1}^{s} a_i}{f_{1_{\max}}},$$

where $f_{1_{\max}} = \max_{h=1,\dots,M} \sum_{i=1}^{s_h} a_i$, and $(a_1,\dots,a_{s_h})$ are positive weighted degrees of association for pattern E^t and class C_h, according to the rules in the RB.

This function accumulates the association degree of the pattern with the class C_j for the rules in the RB. Finally this sum is divided by the maximum sum for all classes, to obtain a normalized value. Bardossy et al., studied this FRM as a method for combining fuzzy rule responses, called *additive combination* ([4]). This operator was also presented by Chi et al., in Ref. [7] like a defuzzification method to produce a classification result for an FRBCSs, called *maximum accumulated matching*. Ishibuchi et al., used it in Ref. [20] as well, where the inference result is given by the voting of the fuzzy if-then rules that are compatible with the pattern to be classified and it was called *fuzzy reasoning method based on the maximum vote*.

(b) *Arithmetic mean.*

$$f_2(a_1,\dots,a_s) = \frac{\sum_{i=1}^{s} a_i}{s}.$$

The arithmetic mean is an operator with a compensation degree between the minimum and the maximum, used to synthesize judgement in multicriteria decision processes. The use of this operator is suitable for combining the information given by each local classifier and to obtain an average degree that considers the quality of the rules in the inference process.

(c) *Quasiarithmetic mean.* The quasiarithmetic mean operator is a strictly monotonous and continuous function defined as

$$f_3(a_1,\dots,a_s) = H^{-1}\left[\frac{1}{s}\sum_{i=1}^{s} H(a_i)\right].$$

We will use $H(x) = x^p, \quad p \in \mathrm{R}$ and so, this operator looks like the generalized mean function with equal weight for the unit to all the values to aggregate, and with a compensation degree between the minimum and the maximum, including the arithmetic and geometric means. Exactly,

$$\text{If } p \to -\infty, \quad f_3 \to \min,$$
$$\text{If } p \to +\infty, \quad f_3 \to \max.$$

In Ref. [12] a detailed study may be found about the properties and behaviour of this operator.

(d) *Sowa And-Like.* The Sowa And-Like operator has the following expression:

$$f_4(a_1, \ldots, a_s) = \alpha a_{\min} + (1 - \alpha)\frac{1}{s}\sum_{i=1}^{s} a_i,$$

where $\alpha \in [0, 1]$ and $a_{\min} = \min\{a_1, \ldots, a_s\}$.

This operator presents a behaviour between the arithmetic mean and the minimum.

If $\alpha = 0$, f_4 is the arithmetic mean.
If $\alpha = 1$, f_4 is the minimum.

(e) *Sowa Or-Like.* The Sowa Or-Like operator is defined as

$$f_5(a_1, \ldots, a_s) = \alpha a_{\max} + (1 - \alpha)\frac{1}{s}\sum_{i=1}^{s} a_i,$$

with $\alpha \in [0, 1]$ and $a_{\max} = \max\{a_1, \ldots, a_s\}$.

In this case, and according to the value of α, this operator will give back a value between the following extremes:

If $\alpha = 0$, f_5 is the arithmetic mean.
If $\alpha = 1$, f_5 is the maximum.

(f) *Badd.* The Badd operator expression is:

$$f_6(a_1, \ldots, a_s) = \frac{\sum_{i=1}^{s} a_i^{p+1}}{\sum_{i=1}^{s} a_i^{p}}, \quad p \in \mathrm{R}.$$

This aggregation function gives us an aggregation value between the minimum and maximum according to the value of p.

If $p \to +\infty$, $f_6 \to \max$,
If $p = 0$, f_6 is the arithmetic mean,
If $p \to -\infty$, $f_6 \to \min$.

Again, as we are interested in an aggregation between the arithmetic mean and the maximum, we will consider $p \in \mathrm{R}^+$ in our experiments.

In Ref. [29] information about properties and behaviour is to be found for the last three aggregation functions.

(5) *Classification.* As regards the classification step, we work according to the expression presented in Section 3.2.

3.4. Experiments

We have generated three RBs (one for each type of rules) for three well known sets of samples: IRIS, WINE and PIMA.

The IRIS base of examples is a set of 150 examples of iris flowers with three classes and four attributes. WINE is a data base of wines with thirteen important characteristics, 178 examples, and three classes. PIMA is a set of 768 solved cases of diagnostics of diabetes where eight variables are taken into account and there are two possible classes (having or not having the illness).

Taking into account the characteristics of the example sets, we have considered it interesting to use, as the initial DB, a fuzzy partition constituted by three triangular fuzzy sets in the case of WINE, and five in the case of IRIS and PIMA.

As we mentioned, we have generated the RB by means of the learning method proposed in Ref. [6,7] (extension of Wang and Mendel's algorithm ([27]) to fuzzy classification rules).

The values that have been used for the parameters are:

f_3: Quasiarithmetic mean $p \in \{2, 5, 10, 20, 50\}$,
f_4: Sowa And Like $\alpha \in \{0.3, 0.5\}$,
f_5: Sowa Or Like $\alpha \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$,
f_6: Badd $p \in \{2, 5, 10, 20, 50\}$.

To calculate an error estimation of an FRBCS, we use random resampling ([28]) with five random partitions of the sample base in training and test sets (70% and 30%, respectively).

The outcomes, that are shown in Appendix B, are means of correct classification percentages for training and test sets. In the tables, the row noted by f_0 corresponds to the FRM based on the winner rule, i.e., the classical FRM; f_1 represents the FRM based on aggregation function f_1 (normalized addition), and so on. The column indicated by g_1 corresponds to the FRM with the weighting function g_1 (without weighting), and the column indicated by g_2 corresponds to the FRM with weighting function g_2.

The best test classification percentages are shown in Tables 1–3.

If we analyze the results of the experiments according to the kind of RB, we observe that:

- In the three example sets, when using type (a) rules, the FRM presenting the best results is the one based on the addition function (f_1). This function is the best way to aggregate the information given by the fired rules when they do not provide a degree of accuracy associated to the classification in a class.
- For type (b) and (c) rules, the FRMs based on Badd operator (f_6), Sowa Or-Like (f_5), and Quasi-Arithmetic Mean (f_3) improve the results from the classical FRM.

Table 1
Test classification percentages for IRIS

Type (a) rules				Type (b) rules				Type (c) rules			
f	p	TE	TE	f	p	TE	TE	f	p	TE	TE
		g_1	g_2			g_1	g_2			g_1	g_2
f_0		88.25		f_0		94.32		f_0		94.32	
f_1		92.88	**94.38**	f_3	10	94.32	94.32	f_3	10	94.32	94. 32
f_5	0.5	90.83	92.27	f_5	0.5	94.32	94.32	f_5	0.7	94.32	94. 32
f_3	5	90.83	91.78	f_6	5	94.32	94.32	f_6	10	94.32	94. 32

Table 2
Test classification percentages for WINE

Type (a) rules				Type (b) rules				Type (c) rules			
f	p	TE	TE	f	p	TE	TE	f	p	TE	TE
		g_1	g_2			g_1	g_2			g_1	g_2
f_0		88.36		f_0		91.94		f_0		91.9	
f_1		**92.81**	91.29	f_1		92.05	91.50	f_1		92.73	91.94
f_3	50	81.86	83.99	f_3	20	91.96	91.94	f_3	5	92.76	92.29
f_6	10	82.32	84.42	f_5	0.9	**92.29**	91.94	f_5	0.5	**92.97**	91.97

Table 3
Test classification percentages for PIMA

Type (a) rules				Type (b) rules				Type (c) rules			
f	p	TE	TE	f	p	TE	TE	f	p	TE	TE
		g_1	g_2			g_1	g_2			g_1	g_2
f_0		64.88		f_0		73.23		f_0		74.16	
f_1		**72.11**	70.97	f_3	50	73.33	73.44	f_5	0.3	73.95	74.47
f_5	0.1	59.63	64.56	f_6	5	73.33	**73.53**	f_6	10	74.47	74.58
f_6	50	61.42	67.38	f_6	10	**73.53**	73.43	f_6	20	**74.68**	74.27

As regards the aggregation function we should point out that:

- The Arithmetic Mean (f_2) presents too much compensation between high and small degrees of association of the pattern with a class.
- The FRM based on Sowa And-Like (f_4) is worse than the others due to the fact that its behaviour is to be found between the arithmetic mean and the

minimum.

- The FRMs based on Quasiarithmetic Mean (f_3), Sowa Or-Like (f_5) and Badd (f_6) operators, overcome the classical FRM and the FRM based on addition function (f_1), but there is no an FRM that has the best behaviour for all example sets and types of rules.

The weighting function g_2 improves, in the most cases, the FRBCS prediction, but it does not have an uniform behaviour. This is due to:

1. *The characteristics of the example set considered.* For problems with broad variable domains, regardless of the generation process used, the RB built will not cover some regions in the search space. The test patterns belonging to these areas can have a very small association degree with the rules that, when weighted up, may be cancelled, and there is no classification for the pattern. In these situations the system behaves worse with the weighting function proposed (g_2).
2. *The characteristics of the RB generation process.* Depending on the generation method used, this deficiency can be accentuated. For the generating method used in this paper, the RB is too fitted to the training examples, and the negative effects of the weighting function are shown in WINE and PIMA example bases. This problem may be solved by a search algorithm that determines the best parameter values of the weighting function for a specific problem and RB. A proposal based on this idea is presented in the next section.
3. *The characteristics of the aggregation function.* The parameters of the weighting function also depend on the aggregation function in the FRM. Aggregation functions such as quasiarithmetic mean (f_3) or Badd operator (f_6), implicitly have a weight associated in their definition. A parameter learning method can also adapt the weighting function for the aggregation function.

As the final comments based on these results, we may point out that:

- The alternative reasoning methods show better behaviour than the classical method in the classification of new patterns in all example sets and for the three types of fuzzy rules considered. We may say that *the new reasoning methods improve the performance of the FRBCSs.*
- Furthermore, as is well known, a Classification System may be characterized by two properties: *Abstraction*, that is, the extraction of the information provided by the training samples to build a suitable structure for the system, and *generalization*, that using the information taken out, makes the classification of new patterns possible. In this Section, we have analyzed the behaviour of an FRBCS made up by an FRM and an RB generated regardless of the FRM. The results show that *the FRMs improve the system generalization property without taking part in the abstraction process.*
- The results obtained, even with the FRM based on Sowa And-Like (f_4), demonstrate that *the aggregation of the information provided by the fired rules is important to define an FRM.*

- *There is no set of parameter values for the aggregation functions with the best classification percentage* for all example sets considered.

4. Extending the fuzzy reasoning methods

In the last subsection, the results of the experiments show that there is no FRM with aggregation and weighting function parameters with the best behaviour. These results support the idea of using a search algorithm to determine the best values for the FRM parameters.

We also observed in Section 3 that the FRMs improve the system generalization property without taking part in the abstraction process. The integration of the FRMs in a generation process would allow us to obtain a set of cooperative rules for the classification system that improves the behaviour of the system.

In this Section, we present a Genetic Algorithm (GA) to learn the FRM parameters, obtaining an FRM fitting the specific problem. Then, we describe some results obtained by the integration of the FRMs in an RB generation process.

4.1. Learning parameters

The parameter learning process is an optimization process that we face with GAs. These kinds of algorithms are general search processes that use elements inspired by natural genetics to evolve solutions to problems. It has been proven theoretic and practically, that they are a robust search mechanism in complex spaces ([14,23]).

We will consider two parameter genetic learning processes:

- A genetic process to learn the parameters of the weighting function for the FRMs with the best values in the aggregation function parameters.
- A genetic process to search for the best values for all parameters for the weighting function and aggregation function of the FRM.

The main components of the GAs used are:

1. Genetic representation of the problem solutions. The parameters are real valued and for that we use a Real-Coded Genetic Algorithm ([17]).
2. The *Evaluation function* that gives the fitness of each chromosome is the classification rate for training example sets.
3. The *Selection mechanism* is stochastic universal sampling ([3]) and the elitist selection model ([14]).
4. The *Genetic operators* applied are mechanisms for inducing a suitable exploitation/exploration balance and so avoid premature convergence: Nonuniform mutation ([23]), and Dubois-Dynamic Crossover ([16]).

The experiments carried out are the following. They have been divided into two parts:

Part I. Learning parameters of the weighting function for the FRM based on:
- normalized sum (f_1),
- arithmetic mean (f_2),
- quasiarithmetic mean (f_3), with $p \in \{20, 50\}$,
- Sowa Or-like (f_5), with $\alpha \in \{0.5, 0.7, 0.9\}$, and
- Badd (f_6) with $p \in \{20, 50\}$.

Part II. Learning parameters of the weighting function and the FRM based on the aggregation functions:
- quasiarithmetic mean (f_3),
- Sowa Or-like (f_5), and
- Badd (f_6).

The experiments carried out are shown in Tables 4–6. The row called f_0 presents the results of classical FRM; the row noted by *br* shows the best classification percentages obtained in the previous experiments carried out in Section 3, and the rows referred by I and II show the results from the experiments in Part I and Part II, respectively.

About the results, we may point out that:
- The experiments carried out show an improvement in the test classification results over the classical reasoning method, although not in comparison with the best results obtained in the previous section.

Table 4
Learning parameters for IRIS

	f	p	Type (a) RB		Type (b) RB		Type (c) RB	
			TR	TE	TR	TE	TR	TE
	f_0		90.97	88.25	97.31	94.32	96.96	94.32
br			98.56	94.38	97.31	94.32	96.96	94.32
I	f_1		99.28	**94.87**	98.57	**94.38**	98.02	94.81
	f_2		91.90	90.34	97.65	93.73	98.21	94.32
	f_3	20	91.90	90.34	97.31	94.32	96.96	94.32
	f_3	50	91.90	90.34	97.31	93.89	96.96	93.89
	f_5	0.5	92.80	90.18	97.49	93.83	97.85	93.89
	f_5	0.7	93.18	90.66	97.31	94.32	97.83	94.32
	f_5	0.9	93.18	91.25	97.31	94.32	97.31	93.83
	f_6	20	91.74	88.74	97.31	94.32	96.96	94.32
	f_6	50	90.97	87.82	97.31	93.89	96.96	93.89
II	f_3		92.08	90.34	97.65	93.73	98.22	94.81
	f_5		92.62	89.12	97.84	94.22	98.21	95.23
	f_6		92.50	91.20	97.33	93.43	97.86	**95.78**

Table 5
Learning parameters for WINE

	f	p	Type (a) RB		Type (b) RB		Type (c) RB	
			TR	TE	TR	TE	TR	TE
	f_0		99.53	88.36	98.88	91.94	98.54	91.94
br			97.55	**92.81**	98.88	92.29	98.38	**92.97**
I	f_1		100.00	90.10	98.88	**92.70**	98.71	92.29
	f_2		97.64	82.05	98.84	92.47	98.39	92.29
	f_3	20	97.64	83.26	99.37	90.53	98.54	91.21
	f_3	50	97.64	83.26	99.37	90.86	98.54	91.21
	f_5	0.5	97.80	82.73	99.37	91.96	98.87	92.35
	f_5	0.7	98.11	84.07	99.54	90.97	98.70	91.94
	f_5	0.9	98.44	83.42	99.20	91.61	98.70	91.94
	f_6	20	99.53	87.60	98.88	91.21	98.54	91.21
	f_6	50	99.53	87.00	99.05	90.88	98.71	91.21
II	f_3		97.64	67.75	99.37	89.88	98.71	90.92
	f_5		98.59	82.99	99.37	92.29	98.87	92.35
	f_6		99.53	70.84	99.52	90.51	99.03	89.53

Table 6
Learning parameters for PIMA

	f	p	Type (a) RB		Type (b) RB		Type (c) RB	
			TR	TE	TR	TE	TR	TE
	f_0		89.51	64.88	85.81	73.23	81.94	74.16
br			83.97	**72.11**	85.88	**73.53**	81.91	**74.68**
I	f_1		91.67	70.47	86.02	72.91	82.08	74.16
	f_2		85.15	58.23	86.40	72.19	82.01	74.16
	f_3	20	84.94	57.91	86.12	71.04	82.08	74.06
	f_3	50	84.84	57.39	86.37	72.49	82.05	72.11
	f_5	0.5	87.45	63.34	86.40	72.91	82.08	74.37
	f_5	0.7	87.90	63.25	86.51	73.12	82.26	74.37
	f_5	0.9	89.09	62.95	86.51	73.12	82.43	74.16
	f_6	20	91.46	68.31	86.51	72.81	82.61	74.47
	f_6	50	91.43	65.52	86.61	72.22	82.75	72.22
II	f_3		84.91	54.00	86.26	71.38	82.08	70.16
	f_5		90.59	65.48	86.54	72.91	82.47	74.26
	f_6		91.46	60.75	86.79	73.22	82.78	74.16

- Since we learn the FRM parameters using only the training data, this strategy does not necessarily lead to the improvement of the generalization ability and sometimes an overlearning phenomenon is presented.
- Nevertheless, with the learning algorithm described, we have obtained the best test classification percentages with this generation method for the IRIS example set for all types of rules, and for WINE and type (b) rules. In both cases, we have used the aggregation function f_6 and f_1 respectively, with a weighting function.

The results of the experiments show that *a significant improvement of the test classification percentages has been obtained with respect to classical FRM.*

4.2. Integration of the FRMs in an RB generation process

The integration of the FRMs in an RB learning method might allow us to obtain a set of cooperative rules for the classification system that improves its behaviour. In this section, we briefly show some experients carried out in this line in a previous work.

In Ref. [10], we present a two-stage genetic fuzzy rule learning process that integrates the first two proposals (FRM based on function f_1, and FRM based on function f_2, both without weighting) in the design of the final RB. In the first stage, an iterative generation process generates rules regardless of the reasoning method, and in the second stage, a genetic selection process obtains a cooperative RB integrating the FRM.

In Table 7 we present the best results obtained with the FRMs based on functions f_1 and f_2, without weighting, and type (a) RB, generated by the method described in Appendix A. In Table 8, we describe the outcomes obtained with the integration of the same FRMs with the generating process explained in Ref. [10].

These results are obtained by an FRBCS made up by:

- An RB with the rule structure with the worst behaviour (RB type (a)) and,
- an FRM based on functions f_1 and f_2 that, as we analyzed in previous sections, is not the FRM with the best behaviour.

Table 7
Classification results for an FRBCS with an RB derived from the method in Appendix A

IRIS		WINE		PIMA	
f	TEST	f	TEST	f	TEST
f_0	88.25	f_0	88.36	f_0	64.88
f_1	92.88	f_1	92.81	f_1	72.11
f_2	88.31	f_2	70.01	f_2	57.47

Table 8
Classification results for an FRBCS with an RB obtained integrating the FRMs

IRIS		WINE		PIMA	
f	TEST	f	TEST	f	TEST
f_0	92.85	f_0	89.97	f_0	70.43
f_1	**95.63**	f_1	92.02	f_1	71.96
f_2	94.54	f_2	**93.08**	f_2	**75.30**

Nevertheless, the outcomes show the improvement of the prediction capability of the system.

These results may suggest that an FRBCS with a type (b) or (c) RB, built by a learning process integrating the FRMs based on functions f_3, f_5 or f_6, could obtain very good classification results.

Table 9
IRIS (1)

f	p	Type (a) rules				Type (b) rules			
		TR	TE	TR	TE	TR	TE	TR	TE
		g_1		g_2		g_1		g_2	
f_0		90.97	88.25			97.31	94.32		
f_1		97.29	92.88	98.56	94.38	96.43	93.20	97.13	92.92
f_2		90.80	88.31	89.72	89.71	96.04	92.39	96.23	93.89
f_3	2	91.35	88.80	89.73	91.78	96.23	93.89	96.41	93.89
f_3	5	91.53	90.83	89.56	91.78	96.41	93.89	96.77	94.32
f_3	10	91.18	90.34	89.56	91.78	96.77	94.32	97.31	94.32
f_3	20	91.18	90.34	89.20	91.78	97.31	94.32	97.31	94.32
f_3	50	89.90	90.34	88.81	91.19	97.31	94.32	97.31	94.32
f_4	0.3	88.26	87.82	89.18	89.22	95.11	91.09	96.04	92.88
f_4	0.5	86.82	86.26	88.09	88.25	93.83	89.57	95.86	92.88
f_5	0.1	90.98	88.31	90.64	91.25	96.39	92.39	96.59	93.89
f_5	0.3	92.08	90.34	90.64	92.27	96.77	93.89	96.77	94.32
f_5	0.5	92.08	90.83	90.64	92.27	96.77	94.32	96.77	94.32
f_5	0.7	92.08	90.83	90.64	92.27	97.13	94.32	97.31	94.32
f_5	0.9	92.08	90.83	90.64	92.27	97.31	94.32	97.31	94.32
f_6	2	89.90	90.34	86.45	86.65	96.23	93.89	96.94	94.32
f_6	5	86.45	85.80	86.45	86.65	97.12	94.32	97.31	94.32
f_6	10	85.37	85.80	86.63	87.13	97.31	94.32	97.31	94.32
f_6	20	85.19	85.63	87.00	86.65	97.31	94.32	97.31	94.32
f_6	50	85.91	86.12	90.97	88.25	97.31	94.32	97.31	94.32

Table 10
IRIS (2)

f	p	Type (c) rules			
		TR	TE	TR	TE
		g_1		g_2	
f_0		96.96	94.32		
f_1		96.25	93.20	96.95	93.40
f_2		96.25	92.72	96.95	93.40
f_3	2	96.95	93.50	96.77	94.32
f_3	5	96.96	94.32	96.96	94.32
f_3	10	96.96	94.32	96.96	94.32
f_3	20	96.96	94.32	96.96	94.32
f_3	50	96.96	94.32	96.96	94.32
f_4	0.3	95.35	92.23	97.13	93.40
f_4	0.5	94.41	92.29	97.13	93.40
f_5	0.1	96.93	92.88	96.77	92.92
f_5	0.3	96.77	93.40	96.96	93.89
f_5	0.5	96.77	92.92	96.96	94.32
f_5	0.7	97.14	94.32	96.96	94.32
f_5	0.9	96.96	94.32	96.96	94.32
f_6	2	96.05	93.89	96.41	94.32
f_6	5	96.77	94.32	96.96	94.32
f_6	10	96.96	94.32	96.96	94.32
f_6	20	96.96	94.32	96.96	94.32
f_6	50	96.96	94.32	96.96	94.32

5. Concluding remarks

In this work, a general reasoning model has been presented. While the classical FRM uses a single rule, this proposal works with the information provided for all rules fired with the pattern to be classified.

This general process includes a *weighting function* applied to the association degrees of the pattern with the different classes, and an *aggregation function* for the information provided by the different rules in the RB. We have analyzed six proposals for the latter and two for the weighting function in terms of the inferred results.

It has been shown that the FRM that considers only the winner rule wastes the information provided by the places of overlapping fuzzy subsets, and the FRM proposals improve the generalization capability of the FRBCS.

Furthermore, we have presented a genetic learning process to obtain the best values for the FRM parameters, and in this way, we have adapted the FRM to

Table 11
WINE (1)

f	p	Type (a) rules				Type (b) rules			
		TR	TE	TR	TE	TR	TE	TR	TE
		g_1		g_2		g_1		g_2	
f_0		99.53	88.36			98.88	91.94		
f_1		97.55	92.81	99.68	91.29	97.89	92.05	98.23	91.50
f_2		85.29	70.01	94.24	77.19	95.10	87.42	96.26	88.28
f_3	2	89.66	76.22	97.64	81.26	96.26	88.28	98.57	90.15
f_3	5	93.77	79.40	97.64	82.38	98.57	90.91	98.72	91.29
f_3	10	95.24	80.21	97.64	82.61	98.72	91.29	98.87	91.96
f_3	20	96.70	81.61	97.64	82.96	98.87	91.96	98.88	91.94
f_3	50	97.02	81.96	97.64	83.99	98.88	91.94	98.88	91.94
f_4	0.3	76.38	62.49	92.65	70.16	92.98	82.98	95.93	87.31
f_4	0.5	67.12	59.11	87.95	66.20	87.49	75.50	95.27	85.09
f_5	0.1	89.68	74.94	97.33	79.52	96.58	88.07	98.10	89.10
f_5	0.3	94.54	78.36	97.64	83.13	98.23	89.53	98.55	91.35
f_5	0.5	96.32	80.97	97.64	83.13	98.72	91.78	98.87	91.29
f_5	0.7	97.48	81.29	97.64	83.05	98.72	91.29	98.72	91.61
f_5	0.9	97.64	81.56	97.64	83.67	98.88	92.29	98.88	91.94
f_6	2	94.43	79.07	97.64	81.70	98.57	90.83	98.72	91.53
f_6	5	96.70	81.28	97.95	82.64	98.72	91.53	98.88	91.94
f_6	10	97.48	82.32	98.11	84.42	98.88	91.94	98.88	91.94
f_6	20	97.48	84.42	98.27	84.42	98.88	91.94	98.88	91.94
f_6	50	98.11	84.42	99.53	88.00	98.88	91.94	98.88	91.94

the problem and RB considered. The outcomes suggest that this process can improve the behaviour of the system.

Apart from showing the improvement of the FRBCS generalization property regardless of the abstraction step, we have pointed out the improvement provided by the integration of the proposed FRMs in a learning process for obtaining a cooperative RB.

With respect to a proposal or suggestion on using these FRMs, the results have shown that it is suitable and necessary to make a prior study for the aggregation operator to use in a specific classification problem. The two aggregation functions f_5 and f_6 have shown good results. Both are in the range between the arithmetic mean and the maximum. We find different behaviours with different parameters, and it would be interesting to study their parameters in any application. Adapting the parameters of the FRMs proposed, we can obtain an FRBCS with a better test classification percentage than the FRBCS based on classical FRM, although it is also necessary to study the parameters to be learnt for every application.

Table 12
WINE (2)

f	p	Type (c) rules			
		TR	TE	TR	TE
		g_1		g_2	
f_0		98.54	91.94		
f_1		97.91	92.73	98.23	91.94
f_2		97.58	92.05	97.74	91.94
f_3	2	97.74	91.94	98.08	92.02
f_3	5	98.07	92.76	98.23	92.29
f_3	10	98.23	92.29	98.39	92.29
f_3	20	98.39	92.29	98.54	91.94
f_3	50	98.54	91.94	98.54	91.94
f_4	0.3	96.76	90.97	97.74	91.94
f_4	0.5	95.77	92.49	97.74	91.94
f_5	0.1	97.74	92.05	97.74	91.94
f_5	0.3	98.22	92.43	98.22	92.70
f_5	0.5	98.38	92.97	98.53	91.97
f_5	0.7	98.53	92.30	98.54	92.29
f_5	0.9	98.54	91.94	98.54	91.94
f_6	2	98.39	92.24	98.23	91.53
f_6	5	98.23	91.53	98.39	91.94
f_6	10	98.39	91.94	98.23	91.94
f_6	20	98.23	91.94	98.54	91.94
f_6	50	98.54	91.94	98.54	91.94

Our future work will be centered on the design of an FRM to select the information to be aggregated. Besides that, we will study the integration of the generating process with the FRM in depth to obtain an appropriate set of cooperative rules according to the FRM selected for the FRBCS.

Appendix A. Generating method

We begin with a set of input–output data pairs (the training data set) with the following structure:

$$
\begin{aligned}
E^1 &= (e_1^1, \ldots, e_N^1, o^1),\\
E^2 &= (e_1^2, \ldots, e_N^2, o^2),\\
&\vdots\\
E^p &= (e_1^p, \ldots, e_N^p, o^p),
\end{aligned}
$$

Table 13
PIMA (1)

f	p	Type (a) rules				Type (b) rules			
		TR	TE	TR	TE	TR	TE	TR	TE
		g_1		g_2		g_1		g_2	
f_0		89.51	64.88			85.81	73.23		
f_1		83.97	72.11	90.83	70.97	81.94	72.21	83.93	72.60
f_2		72.81	57.47	83.65	59.12	82.57	71.36	84.77	72.71
f_3	2	76.99	57.16	84.49	58.73	84.77	72.81	85.84	72.29
f_3	5	81.63	56.87	84.52	58.33	85.71	72.50	85.81	72.81
f_3	10	82.99	57.71	84.52	58.02	85.81	72.81	85.78	73.43
f_3	20	84.38	57.60	84.52	58.54	85.78	73.43	85.85	73.23
f_3	50	84.59	58.85	84.56	59.95	85.78	73.33	85.78	73.44
f_4	0.3	69.91	55.20	82.95	57.69	81.45	70.44	84.14	72.19
f_4	0.5	68.21	54.69	81.80	56.34	80.30	70.03	83.58	71.37
f_5	0.1	78.66	59.63	85.78	64.56	84.03	71.46	85.53	72.39
f_5	0.3	83.37	62.01	85.81	60.88	85.60	72.08	85.60	72.60
f_5	0.5	85.67	62.73	85.81	60.98	85.74	72.80	85.78	72.70
f_5	0.7	86.97	63.46	85.81	61.19	85.88	72.80	85.92	73.01
f_5	0.9	87.55	64.48	85.81	61.80	85.81	73.22	85.85	73.22
f_6	2	81.04	56.46	84.63	57.09	85.64	72.30	85.92	73.23
f_6	5	83.48	57.91	84.73	59.46	85.92	73.33	85.88	73.53
f_6	10	84.80	59.35	84.56	59.87	85.85	73.53	85.81	73.43
f_6	20	84.52	59.66	85.60	61.73	85.81	73.43	85.81	73.33
f_6	50	85.39	61.42	91.32	67.38	85.81	73.33	85.81	73.23

where o^h is the class label for the pattern E^h.

The task here is to generate a set of fuzzy rules from the training data set that describes the relationship between the system variables and determines a mapping D between the feature space S^N and the class set $C = \{C_1, \ldots, C_M\}$.

The method consists of the following steps:

- *Fuzzifying the feature space.* Finding the domain intervals of the attributes and partitioning each domain into X_i regions ($i = 1, \ldots, N$). A membership function is adopted for each fuzzy region. In our experiments we use membership functions with triangular shapes.
- *Generating fuzzy rules from given data pairs.* For each training data $E^h = (e_1^h, \ldots, e_N^h, o^h)$, we have
  To determine the membership degrees of e_i^h in different input fuzzy subsets.
  To assign the input $e_1^h, \ldots, e_N^h$ to the region with the maximum membership degree.
  To produce a fuzzy rule from E^h, with the if-part that represents the selected fuzzy region and the consequent with the class determined by o^h. (It does not repeat the fuzzy rules.)

Table 14
PIMA (2)

f	p	Type (c) rules			
		TR	TE	TR	TE
		g_1		g_2	
f_0		81.94	74.16		
f_1		79.22	72.14	80.24	72.92
f_2		79.22	73.24	80.20	73.13
f_3	2	80.20	73.13	81.25	73.74
f_3	5	81.45	73.74	81.70	74.06
f_3	10	81.70	74.06	81.94	74.26
f_3	20	81.94	74.26	81.94	74.27
f_3	50	81.94	74.27	81.94	74.16
f_4	0.3	79.02	73.24	80.24	73.14
f_4	0.5	78.67	72.83	80.13	73.35
f_5	0.1	80.48	73.14	81.28	73.64
f_5	0.3	81.21	73.95	81.73	74.47
f_5	0.5	81.70	73.95	81.70	74.16
f_5	0.7	81.80	74.16	82.01	74.16
f_5	0.9	82.01	74.06	82.01	74.15
f_6	2	81.66	73.54	82.05	74.06
f_6	5	81.91	74.16	81.94	74.47
f_6	10	81.94	74.47	81.91	74.58
f_6	20	81.91	74.68	81.98	74.27
f_6	50	81.98	74.27	81.98	74.16

Appendix B. Classification results

Appendix B gives the Classification results for IRIS (1), IRIS (2), WINE (1), WINE (2) and PIMA (1), PIMA (2). (For details see Tables 9–14).

References

[1] S. Abe, R. Thawonmas, A fuzzy classifier with ellipsoidal regions, IEEE Transactions on Fuzzy Systems 5 (3) (1997) 358–368.
[2] C. Alsina, E. Trillas, L. Valverde, On some logical connectives for fuzzy set theory, Journal of Mathematical Analysis and Applications 93 (1983) 149–163.
[3] J.E. Baker, Reducing bias and inefficiency in the selection algorithms, in: Proceedings of the Second International Conference on Genetic Algorithms, ICGA'87, Hillsdale, 1987, pp. 14–21.
[4] A. Bárdossy, L. Duckstein, Fuzzy rule-based modeling with applications to geophysical, biological and engineering systems, CRC Press, Boca Raton, 1995.
[5] J.C. Bezdek, S.K. Pal (Eds.), Fuzzy Models for Pattern Recognition, Methods that Search for Structures in Data, IEEE Press, New York, 1992.

[6] Z. Chi, J. Wu, H. Yan, Handwritten numeral recognition using self-organizing maps and fuzzy rules, Pattern Recognition 28 (1) (1995) 59–66.
[7] Z. Chi, H. Yan, T. Pham, Fuzzy algorithms with applications to image processing and pattern recognition, World Scientific, Singapore, 1996.
[8] Z. Chi, H. Yan, ID3-Derived fuzzy rules and optimized defuzzification for handwritten numeral recognition, IEEE Transactions on Fuzzy Systems 4 (1) (1996) 24–31.
[9] O. Cordón, F. Herrera, A. Peregrin, Applicability of the fuzzy operators in the design of fuzzy logic controllers, Fuzzy Sets and Systems 86 (1997) 15–41.
[10] O. Cordón, M.J. del Jesus, F. Herrera, E, López, Selecting fuzzy rule-based classification systems with specific reasoning methods using genetic algorithms, in: Proceedings of the Seventh International Fuzzy Systems Association World Congress (IFSA'97), Prague, 1997, pp. 424–429.
[11] D. Dubois, H. Prade, A review of fuzzy set aggregation connectives, Information Sciences 36 (1985) 85–121.
[12] H. Dyckhoff, W. Pedrycz, Generalized means as model of compensative connectives, Fuzzy Sets and Systems 14 (1984) 143–154.
[13] L. Fu, Rule generation from neural networks, IEEE Transactions on Systems, Man and Cybernetics 24 (8) (1994) 1114–1124.
[14] D. E. Goldberg, Genetic algorithms in search, optimization, and machine learning, Addison-Wesley, Reading, MA, 1989.
[15] A. González, R. Pérez, Completeness and consistency conditions for learning fuzzy rules, Fuzzy Sets and Systems 96 (1998) 37–51.
[16] F. Herrera, M. Lozano, J.L. Verdegay, Dynamic and heuristic fuzzy connectives-based crossover operators for controlling the diversity and convergence of real-coded genetic algorithms, International Journal of Intelligent Systems 11 (1996) 1013–1040.
[17] F. Herrera, M. Lozano, J.L. Verdegay, Tackling real-coded genetic algorithms: Operators and tools for the behavioural analysis, Artificial Intelligence Review 12 (4) (1998) 265–319.
[18] H. Ishibuchi, K. Nozaki, H. Tanaka, Distributed representation of fuzzy rules and its application to pattern classification, Fuzzy Sets and Systems 52 (1992) 21–32.
[19] H. Ishibuchi, K. Nozaki, N. Yamamoto, H. Tanaka, Construction of fuzzy classification systems with rectangular fuzzy rules using genetic algorithms, Fuzzy Sets and Systems 65 (1994) 237–253.
[20] H. Ishibuchi, T. Morisawa, T. Nakashima, Voting schemes for fuzzy-rule-based classification systems, in: Proceedings of the Fifth IEEE International Conference on Fuzzy Systems (FUZZ-IEEE'96), 1996, pp. 614–620.
[21] L.I. Kuncheva, On the equivalence between fuzzy and statistical classifiers, International Journal of Uncertainty Fuzziness and Knowledge-Based Systems 4 (3) (1996) 245–253.
[22] D.P. Mandal, C.A. Murthy, S.K. Pal, Formulation of a multivalued recognition system, IEEE Transactions on Systems, Man, and Cybernetics 22 (4) (1992) 607–620.
[23] Z. Michalewicz, Genetic algorithms + data structures = evolutions programs, Springer, Berlin, 1992.
[24] S. Pal, D.P. Mandal, Linguistic recognition system based on approximate reasoning, Information Sciences 61 (1992) 135–161.
[25] K.S. Ray, J. Ghoshal, Approximate reasoning approach to pattern recognition, Fuzzy Sets and Systems 77 (1996) 125–150.
[26] V. Uebele, S. Abe, M.-S. Lan, A neural-network-based fuzzy classifier, IEEE Transactions on Systems, Man, and Cybernetics 25 (2) (1995) 353–361.
[27] L.-X. Wang, J.M. Mendel, Generating fuzzy rules by learning from examples, IEEE Transactions on Systems, Man, and Cybernetics 22 (6) (1992) 1414–1427.
[28] S.M. Weiss. C.A. Kulikowski, Computer systems that learn, Morgan Kaufmann, Los Altos, CA, 1991.

[29] R.R. Yager, D.P. Filev, Parameterized and-like and or-like OWA operators, International Journal General Systems 22 (1994) 297–316.

[30] L.A. Zadeh, Fuzzy sets and their applications to classification and clustering, in: J. van Ryzin (Ed.), Classification and Clustering, Academic Press, New York, 1977, pp. 251–299.