A Genetic Algorithm for Tuning Fuzzy Rule-Based Classification Systems with Interval-Valued Fuzzy Sets

J. Sanz, A. Fernández, H. Bustince and F. Herrera

Abstract— **Fuzzy Rule-Based Classification Systems are a widely used tool in Data Mining because of the interpretability given by the concept of linguistic label. However, the use of this type of models implies a degree of uncertainty in the definition of the fuzzy partitions. In this work we will use the concept of Interval-Valued Fuzzy Set to deal with this problem. The aim of this contribution is to show the improvement in the performance of linguistic Fuzzy Rule-Based Classification Systems afterward the application of a cooperative tuning methodology between the tuning of the amplitude of the support and the lateral tuning (based on the 2-tuples fuzzy linguistic model) applied to the linguistic labels modeled with Interval-Valued Fuzzy Sets.**

I. Introduction

Computational Intelligence methods have shown to be useful tools to solve complex problems in classification tasks. Among them, Fuzzy Rule-Based Classification Systems (FRBCSs) [13] are widely employed since they provide a good performance together with a high interpretability of the rules due to the use of linguistic labels. Furthermore, these systems offer the possibility of mixing different kinds of information as the one given by experts or the one obtained by mathematical models or empirical measures.

In FRBCSs the definition of the membership functions used to represent the linguistic labels is truly significant. When defining the fuzzy partitions we can use expert knowledge or we can simply proceed by establishing an homogeneous partition over the input space. In both cases, the theory of Interval-Valued Fuzzy Sets (IVFSs) [15], [3] allows us to model the possible ignorance inherent to the definition of membership functions. We must point out that the interval membership of an element to a set provides a lower bound and an upper bound for the punctual value of the membership of the element to the set. In [16], we have shown that the use of IVFSs in FRBCSs is useful in the framework of classification with imbalanced data-sets.

In addition to the previous issue, sometimes fuzzy partitions are not well fitted to the context, as they remain fixed during the rule generation process. Therefore, it seems necessary to carry out a post-processing step to tune the linguistic labels modeled, in this case, by means of IVFSs.

The objective of this work is to improve the performance of FRBCSs by means of the cooperation between IVFSs and a post-processing tuning step. With this aim, in addition to the use of the IVFSs model in the Fuzzy Reasoning Method (FRM) (afterward of the rule generation process), we propose a cooperative tuning methodology to manage both the amplitude of the support and the position of the linguistic labels through a lateral displacement (based on the 2-tuples fuzzy linguistic model [10]). We study the behavior of our methodology with 16 data-sets selected from UCI repository [2].

The work is arranged as follows: in Section II we start explaining the rule generation algorithm employed. Then, we present the FRBCS with the linguistic labels modeled by IVFSs in Section III. The approaches to perform the genetic tuning are introduced in Section IV. Next, in Section V, we show the experimental study carried out and we finish the work with some concluding remarks in Section VI.

J. Sanz and H. Bustince are with the Department of Automatics and Computation, University of Navarre, Campus Arrosadia s/n, 31006 Navarre, Spain (phone: +34 948 166048; emails: {bustince, joseantonio.sanz}@unavarra.es).

A. Fernández and F. Herrera are with the Department of Computer Science and Artificial Intelligence, CITIC-UGR (Research Center on Information and Communications Technology) University of Granada, 18071 Granada, Spain (emails: {alberto, herrera}@decsai.ugr.es).

II. Fuzzy Rule Based Classification Systems And Fuzzy Learning Method

FRBCSs are a very useful tool in Data Mining, since they allow the inclusion of all the available information in system modeling, both the one that comes from expert knowledge and the one from empirical measures and mathematical models, deriving on a very interpretable model and therefore allowing the knowledge representation to be understandable for the system users.

Any classification problem consists of m training patterns $x_p = (x_{p1}, \ldots, x_{pn}, y_p)$, $p = 1, 2, \ldots, m$ from M classes where x_{pi} is the ith attribute value ($i = 1, 2, \ldots, n$) of the p-th training pattern.

In this work we use fuzzy rules of the following form for our FRBCSs:

$$\text{Rule } R_j: \quad \text{If} \quad x_1 \text{ is } A_{j1} \text{ and } \ldots \text{ and } x_n \text{ is } A_{jn} \quad \text{then} \quad \text{Class} = C_j \text{ with } RW_j,$$

where R_j is the label of the jth rule, $x = (x_1, \ldots, x_n)$ is an n-dimensional pattern vector, A_{ji} is an antecedent fuzzy set, C_j is a class label, and RW_j is the rule weight. We use triangular membership functions as antecedent fuzzy sets.

Fuzzy learning methods are the basis to build a FRBCS. The algorithm used in this work is the method proposed in [4], that we have called the Chi et al.'s rule generation.

To generate the fuzzy rule base, this FRBCSs design method determines the relationship between the variables of the problem and establishes an association between the space of the features and the space of the classes by means of the following steps:

1) *Establishment of the linguistic partitions.* Once the domain of variation of each feature A_i is determined, the fuzzy partitions are computed.
2) *Generation of a fuzzy rule for each example* $x_p = (x_{p1}, \ldots, x_{pn}, C_p)$. To do this it is necessary:
   2.1 To compute the matching degree $\mu(x_p)$ of the example to the different fuzzy regions using a conjunction operator (usually modeled with a minimum or product T-norm).
   2.2 To assign the example x_p to the fuzzy region with the greatest membership degree.
   2.3 To generate a rule for the example, whose antecedent is determined by the selected fuzzy region and whose consequent is the label of class of the example.
   2.4 To compute the rule weight.

We must remark that rules with the same antecedent can be generated during the learning process. If they have the same class in the consequent we just remove one of the duplicated rules, but if they have a different class only the rule with the highest weight is kept in the rule base.

III. Fuzzy Rule-Based Classification Systems With Interval-Valued Fuzzy Sets

In this section we present the model that employs IVFSs to represent the linguistic labels of FRBCSs. The use of IVFSs allows to handle the uncertainty associated with the ad-hoc construction of fuzzy partitions and, in this way, it is possible to increase the performance of the system.

IVFSs [3] are an extension of the theory of fuzzy sets [18]. These sets were born in 1975 with the work of Sambuc [15]. Later, in the eighties, Gorzalczany called these sets IVFSs for the first time [9].

We denote by $L([0,1])$ the set of all closed subintervals of the closed interval $[0,1]$; that is:

$$L([0,1]) = \{\mathbf{x} = [\underline{x}, \overline{x}] | (\underline{x}, \overline{x}) \in [0,1]^2 \text{ and } \underline{x} \leq \overline{x}\}$$

$L([0,1])$ is a partially ordered set with respect to the relation $\leq_L$ defined in the following way; given $\mathbf{x}, \mathbf{y} \in L([0,1])$:
$\mathbf{x} \leq_L \mathbf{y}$ if and only if $\underline{x} \leq \underline{y}$ and $\overline{x} \leq \overline{y}$.
With this order relation the smallest element is $[0,0]$ and the largest is $[1,1]$.

Definition 1: An Interval-Valued Fuzzy Set A on the universe $U \neq \emptyset$ is a mapping $A : U \rightarrow L([0,1])$ such that the membership degree of $u \in U$ is given by $A(u) = [\underline{A}(u), \overline{A}(u)] \in L([0,1])$, where $\underline{A} : U \rightarrow L([0,1])$ and $\overline{A} : U \rightarrow L([0,1])$ are functions defining the lower and upper bounds of the membership interval $A(u)$, respectively.

We generate the initial knowledge base by means of the rule learning algorithm of Chi et al.'s [4], explained in the previous section. We use the initial knowledge base as starting point to obtain the IVFSs: we take the fuzzy sets generated by this algorithm as the lower bounds and we add the upper bound to each set. In this manner, we will study the influence of the IVFSs in the FRM, not in the rule generation process.

We build the upper bound in this way: it is centered in the maximum of the membership function of the fuzzy partition and the amplitude of its support is 50% greater than the one of the lower bound (being symmetrical in both sides). Figure 1 shows an example of a linguistic variable represented by 3 labels (IVFSs) in the initial state. The solid lines represent the lower bounds ($\underline{A_j}$) and the dashed lines represent the upper bounds ($\overline{A_j}$).

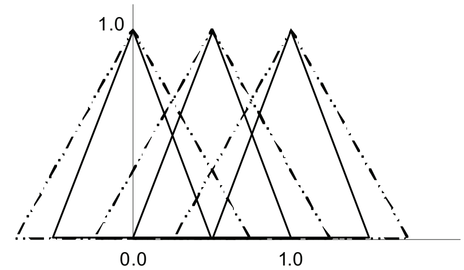


Fig. 1. Example if the IVFSs employed.

Furthermore, as we work with IVFSs, the rule weight will be compounded by a tuple (PCF_{Lj}, PCF_{Uj}) computed using the *Penalized Certainty Factor* (PCF) (see [14]):

$$PCF_{Lj} = \frac{\sum\limits_{x_p \in Class C_j} \underline{A_j}(x_p) - \sum\limits_{x_p \notin Class C_j} \underline{A_j}(x_p)}{\sum\limits_{p=1}^{m} \underline{A_j}(x_p)} \quad (1)$$

$$PCF_{Uj} = \frac{\sum\limits_{x_p \in Class C_j} \overline{A_j}(x_p) - \sum\limits_{x_p \notin Class C_j} \overline{A_j}(x_p)}{\sum\limits_{p=1}^{m} \overline{A_j}(x_p)} \quad (2)$$

As the lower bound of each IVFS is the fuzzy set created by the rule learning algorithm given by Chi et al.'s [4], PCF_{Lj} is equal to RW_j.

In this work we employ the FRM of the winning rule, this FRM among others can be seen in [5]. However, the use of IVFSs implies the following two changes in the FRM:

- *Matching degree between the antecedent of the rule and the example*: We apply the product T-norm both to the lower bound and the upper bound. So, we have an interval.

$$\mu_L A_j(x_p) = T(\underline{A_{j1}}(x_{p1}), \ldots, \underline{A_{jn}}(x_{pn})), \quad j = 1, \ldots, L. \quad (3)$$

$$\mu_U A_j(x_p) = T(\overline{A_{j1}}(x_{p1}), \ldots, \overline{A_{jn}}(x_{pn})), \quad j = 1, \ldots, L. \quad (4)$$

- *Association degree*: We take the mean between the product of the matching degree by the rule weight associated to the lower bound and the product of the matching degree by the rule weight associated to the upper bound.

$$b_j^k = \frac{\mu_L A_j(x_p) * PCF_L j^k + \mu_U A_j(x_p) * PCF_U j^k}{2}$$
$$k = 1, \ldots, M, \quad j = 1, \ldots, L. \quad (5)$$

At this point we already have a single value associated to the class. According to this, we can apply the rest of the method in the same way than in the general FRM.

IV. Linguistic Label Tuning Through A Genetic Algorithm

Membership functions are usually obtained by normalization process or defined by experts. In all cases they are not well-suited to the context of each variable due to they remain fixed during the rule generation process. For this reason it is necessary to perform a post-processing step in order to tune the fuzzy partitions to adapt the behavior of FRBCSs for each single problem.

In this section we introduce three approaches to carry on the tuning. We start describing the model in which we tune the amplitude of the upper bound of the IVFSs, next we explain the genetic tuning based on the 2-tuples fuzzy linguistic model [10] to make the lateral displacement and then we introduce the methodology in which the previous genetic tuning models cooperate. Finally, we describe the use of the CHC algorithm [7] which has been selected to perform the genetic tuning process.

A. Tuning Of The Amplitude Of The Linguistic Labels

As we have stated in Section III we represent the linguistic labels of the FRBCS by means of IVFSs. We build the upper bound of each IVFS with a fixed amplitude. However, as the expert has not the same uncertainty defining the different membership functions, the amplitude of the upper bound does not need to be the same for all linguistic labels. Therefore, we perform a post-processing step, by means of a genetic algorithm, in which we tune the amplitude of the upper bound of the IVFSs. It is necessary to remark that the amplitude of the lower bound will remain fixed.

The modification of the amplitude is given by a number within the interval [0, 1], that is, from the overlapping of both bounds (value 0) to twice the amplitude of the upper with respect to the lower bound (value 1). The amplitude of the upper bound will be uniformly increased according to intermediate values being ‘0.5’ the initial situation of the FRBCS, that is, when the amplitude of the upper bound is 50% greater than the one of the lower bound. These situations are depicted in Figure 2.

B. Lateral Tuning Of Linguistic Labels Based On The 2-tuples Model

In our initial model we have fixed the position of the different labels, such that for each value of the input space of each variable the sum of the membership degrees of the different labels is 1. This labels’ distribution does not need to be optimal as the data distribution does not need to be uniform. Therefore, we use the genetic tuning based on the 2-tuples fuzzy linguistic model [10] (adjusting the previous proposal in this topic [1] for our problem) to make the lateral displacements of the linguistic labels. In this manner, we could improve the performance of the FRBCS.

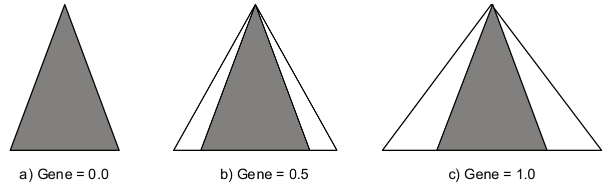


Fig. 2. Gene values representation in the genetic amplitude tuning. a) Upper and lower bounds are overlapped. b) Initial situation. c) Upper bound amplitude is twice than the one of the lower bound

The symbolic translation of a linguistic term is a number within the interval [-0.5, 0.5) that expresses the bounds of the domain of a label when it is moving between its two lateral labels. If the number is negative the displacement will be done to the left and, if the number is positive, to the right. An example is illustrated in Figure 3 where we show the symbolic translation of a label represented by the pair (S_2,-0.3) together with the lateral displacement of the corresponding membership function.

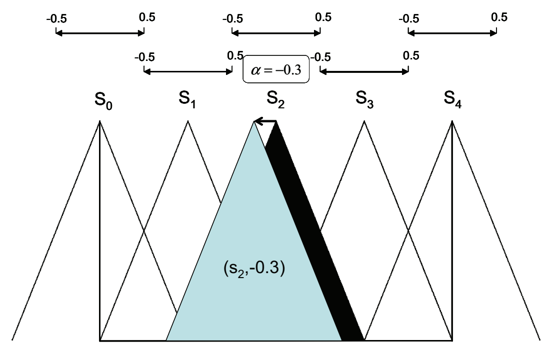


Fig. 3. 2-tuples based tuning model.

C. Cooperative Tuning For Both The Amplitude And The Lateral Displacement

Along this section, we have defined two posible tuning approaches of the membership functions separately. However, in this work we want to go one step further and to propose a tuning methodology in which we tune simultaneously both the the amplitude and the the position of the linguistic labels. To do so, we need to define an unique representation inside the genetic algorithm for both possibilities.

The motivation for this proposal lies in the following premise: the use of each previously defined tunings separately can make that the system reach a sub-optimal model. With this cooperative methodology, the genetic algorithm search engine will be able to work at the same time with both characteristics, what should lead us to discover better adapted solutions to the problem and therefore more accurate ones.

D. CHC Algorithm

In order to apply the genetic tuning, we will consider the use of CHC algorithm [7], which presents a good trade-off between diversity and convergence, being a good choice in complex problems. The components needed to design this process are explained below:

1) *Coding Scheme:* A real coding is considered in all the models, where the representation of each gene of the chromosome depends on the model:
   - *Amplitude tuning:* Each gene represents the amplitude modification as we have explained in Subsection IV-A. So, the chromosome length is equal to the number of labels times the number of variables.
   - *Lateral tuning:* Each gene represents the lateral displacement as we have explained previously. So, the chromosome length is equal to the number of labels times the number of variables.
   - *Cooperative tuning:* Each chromosome will be composed by two parts, one to make the amplitude tuning and the other one to make the lateral tuning. The representation of each part is the same as we have explained for the previous tuning models.
2) *Chromosome Evaluation:* The fitness function is the accuracy rate.
3) *Initial Gene Pool:* Depending on the model, we initialize the population in a different way:
   - *Amplitude tuning:* The initial pool is obtained with the first individual having all genes with value 0.5 (the initial FRBCS). The second and the third individuals have all genes with values 0 and 1 respectively, whereas the remaining individuals are generated at random in [0, 1].
   - *Lateral tuning:* The initial pool is obtained with the first individual having all genes with value 0.0 (the initial FRBCS), whereas the remaining individuals are generated at random in [-0.5, 0.5].
   - *Cooperative tuning:* In this model we initialize three individuals to cover the same situations than in the amplitude tuning. The first part of each chromosome of these three individuals is initialized having all genes with value 0.0 and the remainder of each chromosome will be initialized as in the amplitude tuning. The remaining individuals will have initialized all the genes randomly.
4) *Crossover Operator:* We consider the Parent Centric BLX (PCBLX) operator [11], which is based on the BLX-α. Figure 4 depicts the behavior of these kinds of operators.

   PCBLX is described as follows. Let us assume that $X = (x_1 \cdots x_n)$ and $Y = (y_1 \cdots y_n)$, $(x_i, y_i \in [a_i, b_i] \subset \Re, i = 1 \cdots n)$, are two real-coded chromosomes that are going to be crossed. PCBLX operator generates the two following offspring:

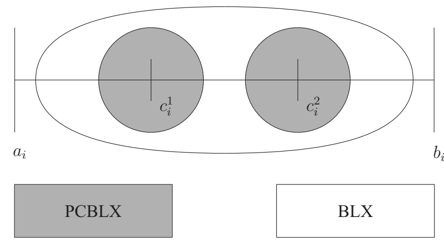


Fig. 4. Scheme of the behavior of the BLX and PCBLX operators

   - $O_1 = (o_{11} \cdots o_{1n})$, where o_{1i} is a randomly (uniformly) chosen number from the interval $[l_i^1, u_i^1]$, with $l_i^1 = max\{a_i, x_i - I_i\}$, $u_i^1 = min\{b_i, x_i + I_i\}$, and $I_i = \mid x_i - y_i \mid$.
   - $O_2 = (o_{21} \cdots o_{2n})$, where o_{2i} is a randomly (uniformly) chosen number from the interval $[l_i^2, u_i^2]$, with $l_i^2 = max\{a_i, y_i - I_i\}$ and $u_i^2 = min\{b_i, y_i + I_i\}$.

   On the other hand, the incest prevention mechanism will be only considered in order to apply the PCBLX operator. In our case, two parents are crossed if half their Hamming distance is above a predetermined threshold, L. Since we consider a real coding scheme, we have to transform each gene considering a Gray Code (binary code) with a fixed number of bits per gene ($BITSGENE$), that is determined by the system expert. In this way, the threshold value is initialized as:

$$L = (\#Genes \cdot BITSGENE)/4.0$$

   where $\#Genes$ stands for the total length of the chromosome. Following the original CHC scheme, L is decremented by one ($BITSGENE$ in this case) when there are no new individuals in the next generation.

   In order to work with the cooperative tuning methodology, we realize the cross in the following way: we cross the parts of the chromosome representing the same kind of tuning among them. In this process we generate four offspring and we select the two best ones.
5) *Restarting approach:* When the threshold value is lower than zero, all the chromosomes are regenerated randomly. Furthermore, the best global solution found is included in the population to increase the convergence of the algorithm.

V. Experimental Study

In this study, our aim is to analyze the behavior of the FRBCS by the combination of the IVFS model and the post-processing step. Furthermore, we want to check whether the behavior of the synergy between the genetic amplitude tuning and the lateral tuning improves the behavior of both tuning approaches when they are performed separately.

In this section, first we describe the experimental set-up together with the parameters employed in the study. Next we

TABLE I

SUMMARY DESCRIPTION FOR THE EMPLOYED DATA-SETS.

Data-set	#Ex.	#Atts.	#Class.
Balance	625	4	3
Bupa	345	6	2
Cleveland	297	13	5
Ecoli	336	7	8
Glass	214	9	6
Haberman	306	3	2
Iris	150	4	3
Magic	1902	10	2
New-Thyroid	215	5	3
Page-blocks	548	10	5
Penbased	1099	16	10
Pima	768	8	2
Ring	768	8	2
Vehicle	846	18	4
Wine	178	13	3
Wisconsin	683	9	2

introduce the statistical tests used and we conclude with the empirical results achieved.

A. Experimental Set-Up

We have analyzed the performance of the different proposals in 16 data-sets selected from UCI repository [2]. Table I summarizes the following characteristics of each data-set: number of examples (#Ex.), number of attributes (#Atts.) and number of classes (#Class.). We must point out that the *Magic, Page-blocks, Penbased* and *Ring* data-sets are stratified-sampled to the 10% to improve the learning process efficiency and we have removed the missing values of *Cleveland and Wisconsin* after the partionate.

To carry out the different experiments we consider a *5-folder cross-validation model*, i.e., 5 random partitions of data with a 20%, and the combination of 4 of them (80%) as training and the remaining one as test. For each data-set we consider the average results of the five partitions.

The FRBCS configuration is the following one: 3 labels per fuzzy partition, product T-norm as conjunction operator, together with the PCF heuristic [14] for the computation of the rule weight and FRM of the winning rule.

All the genetic tuning models presented in Section IV employ populations composed by 50 individuals, 30 bits per gen in order to perform the *gray* codification and the number of evaluations is 5.000 times the number of attributes.

B. Statistical Test For The Performance Comparison

In this work, we use some hypothesis validation techniques in order to give a statistical support to the result analysis. We will use a non parametric test, because the initial conditions that guarantee the reliability of the parametric tests cannot be fulfilled [6], [8]. We employ the Wilcoxon rank test [17] as non parametric statistical procedure to make comparisons between two algorithms; we use the Iman-Davenport test [17] to detect statistical differences among a group of results and the Holm post-hoc test [12] to find the algorithms that reject the equality hypothesis with respect to a selected control method.

C. Empirical Results

The objective of this study is double, on one hand, we want to show the goodness of our proposal and, on the other hand, we want to analyze whether the cooperative tuning methodology is better than the simple tuning approaches when they are performed separately.

Table II shows the results achieved by the different approaches, both in training and in test in each data-set. This table presents two different groups of results. The first one covers the approaches in which the data base is composed by standard fuzzy sets, that is, the results achieved by the basic Chi et al.'s algorithm and the results provided by the initial FRBCS post processed with lateral tuning. The second group is formed by the approaches in which the data base is composed by IVFSs, that is, the results of the models of amplitude, lateral and cooperative tuning.

Usually, tuning models present an overfitting behavior, especially when the tuning is guided by the accuracy rate. However, as we can observe from Table II, our cooperative tuning methodology present a good balance between the performance in training and test, which is a desirable property in this kind of methodology.

In Figure 5 we show the ranking of the different methods. The value given to each method is calculated assigning the position of each algorithm depending on the performance for each data-set and then computing the mean value. We can observe that our proposal of cooperative tuning is the best, followed by the lateral tuning applied to the FRBCS with IVFS and for the lateral tuning applied to the basic FRBCS.

We carry out a Iman-Davenport test in which we find significative differences among the results since the "p-value" obtained is close to zero. For this reason, we can apply a post-hoc test (the Holm test in this case) to compare our cooperative tuning methodology against the remaining methods.

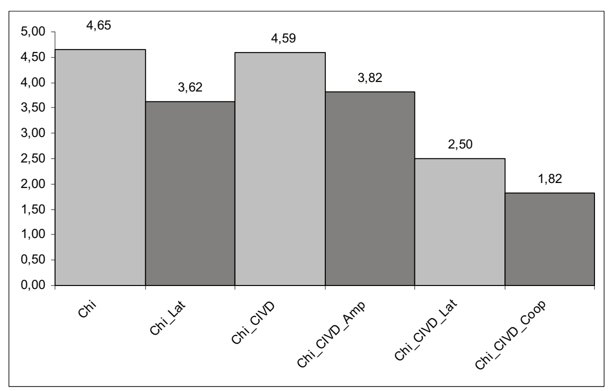


Fig. 5. Ranking of the different proposals.

From the results of this Holm test (Table III) it is shown that our cooperative tuning methodology is statistically better than the remaining approaches. Furthermore, we carry out a Wilcoxon test (Table IV) in which we show that our methodology enhances the performance of the lateral tuning applied to the FRBCS with IVFSs.

TABLE II

Results in Train (Tr.) and Test (Tst) of the different methods.

Data-set	Chi		Chi_GTS		Chi_IVFS		Chi_IVFS_Amp		Chi_IVFS_GTS		Chi_IVFS_Coop	
	$Tr.$	Tst	$Tr.$	Tst	$Tr.$	Tst	$Tr.$	Tst	$Tr.$	Tst	$Tr.$	Tst
Balance	91.62	89.92	92.18	89.76	91.10	90.24	92.18	90.88	92.30	**91.52**	92.30	90.08
Bupa	60.73	57.68	77.45	58.84	59.71	58.26	61.09	57.68	68.87	59.42	75.85	**61.16**
Cleveland	92.22	36.01	94.76	40.75	90.02	**53.21**	92.81	50.50	94.00	51.51	95.86	**53.21**
Ecoli	76.18	72.64	71.26	64.84	69.98	67.58	77.37	71.45	68.35	61.28	88.35	**82.16**
Glass	66.28	57.95	75.09	60.30	55.94	51.87	69.21	59.86	77.21	**64.50**	78.73	63.57
Haberman	74.57	72.88	78.42	72.56	73.83	**73.53**	74.73	72.88	75.88	72.88	78.51	73.22
Iris	92.94	92.67	98.32	94.67	93.61	92.00	97.31	94.67	97.98	**96.00**	98.49	94.00
Magic	75.98	74.87	83.64	78.86	73.31	72.66	77.06	75.08	81.63	78.92	83.37	**79.13**
New-Thyroid	86.32	84.65	98.25	**94.88**	81.75	80.93	86.43	85.12	90.29	86.05	99.18	**94.88**
Pageblocks	92.73	91.42	94.38	93.06	91.50	90.88	92.87	91.42	93.14	91.24	94.42	**93.61**
Penbased	98.66	94.27	99.50	93.09	95.31	92.27	98.95	94.36	99.11	**95.55**	99.59	95.09
Pima	75.45	72.53	81.94	73.69	69.71	67.97	76.30	72.39	78.55	74.08	81.55	**74.22**
Ring	59.53	52.70	96.92	85.81	51.17	50.41	59.66	52.43	93.81	85.81	96.14	**88.78**
Vehicle	65.85	60.88	78.78	**66.55**	54.87	51.55	68.99	60.05	74.08	63.24	77.89	66.08
Wine	98.73	92.67	100.00	93.81	98.02	94.90	99.86	94.33	100.00	94.38	100.00	**94.92**
Wisconsin	98.17	90.49	99.27	93.71	97.65	96.05	98.39	96.05	98.68	**96.19**	99.08	95.75
Mean	81.62	74.64	88.76	78.45	77.97	74.02	82.70	76.20	86.49	78.91	89.96	**81.24**

Taking into account all of these results, we have shown the good synergy between both tuning models in a cooperative methodology exploiting the advantages that each tuning approach provides separately.

TABLE III

Holm Test to compare all the methodologies. The cooperative tuning methodology is selected as the control method.

i	Algorithm	z	p	α/i	Hypothesys ($\alpha = 0, 05$)
5	Chi	4.400	1.082E-5	0.01	Rejected for Chi_IVFS_Coop
4	Chi_IVFS	4.308	1.644E-5	0.0125	Rejected for Chi_IVFS_Coop
3	Chi_IVFS_Amp	3.117	0.002	0.017	Rejected for Chi_IVFS_Coop
2	Chi_GTS	2.796	0.005	0.025	Rejected for Chi_IVFS_Coop
1	Chi_IVFS_GTS	1.054	0.292	0.05	Not rejected

TABLE IV

Wilcoxon Test to compare the cooperative tuning methodology against the lateral tuning applied to the FRBCS with IVFSs.

Comparison	R^+	R^-	Hypothesys ($\alpha = 0.05$)	p-value
Chi_IVFS_Coop vs. Chi_IVFS_GTS	118.0	35.0	Rejected for Chi_IVFS_Coop	0.049

Finally, in order to illustrate the effect of the cooperative tuning methodology in the fuzzy partitions, Table V depicts by columns the results of the tuning process for the *iris* data-set. The first column refers to the variable studied, the second column shows the fuzzy partition in the initial FRBCS with IVFSs generated by the Chi et al.'s algorithm. Next, we present the final values of the genes after the tuning process and the last column depicts the representation of each fuzzy partition afterward of the post-processing step. First of all, the final values of the genes confirms the necessity of the contextualization of each linguistic label as no fuzzy partition remains in its initial state. Furthermore, we show that for the first and second variables all linguistic labels are displaced slightly to the right while their amplitudes do not suffer large variations except in two cases, one for each variable. Regarding the two last variables, we show that their amplitude tend to shrink considerably and in both variables one linguistic label is displaced a lot from its initial position. In conclusion, since the amplitude of the upper bound several linguistic labels decrease, we can asseverate that the initial fuzzy partitions are quite well defined. This fact is especially significant in the two last variables which, knowing the features of this data-set, are enough to discriminate well among all the classes.

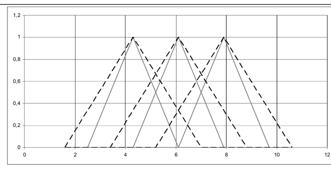
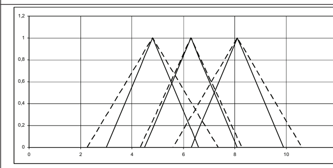
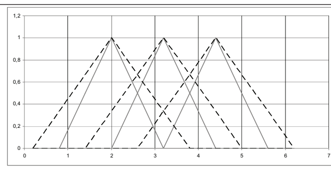
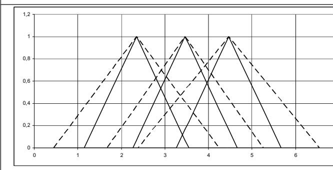
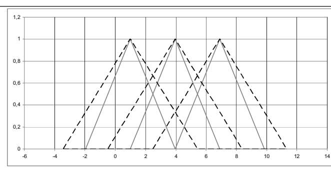
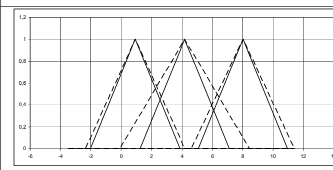
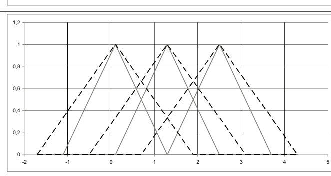
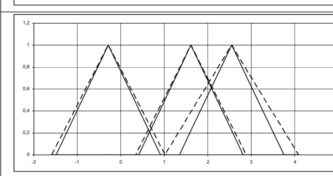
VI. Conclusions

In this work we have proposed the application of a post-processing genetic model applied to the FRBCSs with IVFS that performs simultaneously both the tuning of the amplitude of the upper bound of the IVFS and the lateral tuning, based on the 2-tuples fuzzy linguistic model. The aim of this methodology is to perform a good cooperation of both approaches in order to improve the performance of the initial fuzzy system.

The achieved empirical results shown that, on one hand, tuning models improve their respective initial systems (both the basic FRBCS and the FRBCS with IVFSs) and, on the other hand, our cooperative tuning methodology is statistically better than the tuning of the amplitude and the lateral tuning applied to the FRBCS with IVFSs. In this way, we handle in a proper way the uncertainties of the system by means of IVFSs. The higher performance of our cooperative methodology seems to prove the goodness of our methodology, reaching a better tuning of the membership functions to the context of each problem.

TABLE V

FUZZY PARTITIONS OF THE *IRIS* DATA-SET BEFORE AND AFTER OF THE COOPERATIVE TUNING METHODOLOGY.

<table>
<tr><th></th><th>Initial partition</th><th>Gene values for each linguistic label (LL)</th><th>Final partition</th></tr>
<tr><td>Sepal length</td><td></td><td><table><tr><td>Tuning</td><td>LL. 1</td><td>LL. 2</td><td>LL. 3</td></tr><tr><td>GTS</td><td>0.278</td><td>0.109</td><td>0.112</td></tr><tr><td>Amp</td><td>0.207</td><td>0.048</td><td>0.197</td></tr></table></td><td></td></tr>
<tr><td>Sepal width</td><td></td><td><table><tr><td>Tuning</td><td>LL. 1</td><td>LL. 2</td><td>LL. 3</td></tr><tr><td>GTS</td><td>0.289</td><td>0.218</td><td>0.054</td></tr><tr><td>Amp</td><td>0.291</td><td>0.247</td><td>0.367</td></tr></table></td><td></td></tr>
<tr><td>Petal length</td><td></td><td><table><tr><td>Tuning</td><td>LL. 1</td><td>LL. 2</td><td>LL. 3</td></tr><tr><td>GTS</td><td>-0.026</td><td>0.078</td><td>0.381</td></tr><tr><td>Amp</td><td>0.053</td><td>0.224</td><td>0.076</td></tr></table></td><td></td></tr>
<tr><td>Petal width</td><td></td><td><table><tr><td>Tuning</td><td>LL. 1</td><td>LL. 2</td><td>LL. 3</td></tr><tr><td>GTS</td><td>-0.323</td><td>0.260</td><td>0.041</td></tr><tr><td>Amp</td><td>0.043</td><td>0.029</td><td>0.140</td></tr></table></td><td></td></tr>
</table>

Acknowledgment

This work was supported in part by the Spanish Ministry of Science and Technology under projects TIN2008-06681-C06-01 and TIN2007-65981.

References

[1] R. Alcalá, J. Alcalá-Fdez and F. Herrera: A proposal for the Genetic Lateral Tuning of Linguistic Fuzzy Systems and its Interaction with Rule Selection *IEEE Transactions on Fuzzy Systems*, vol. 15, no. 4, pp. 616–635, (2007)

[2] A. Asuncion and D. Newman: UCI machine learning repository, 2007. University of California, Irvine, School of Information and Computer Sciences. `http://www.ics.uci.edu/~mlearn/MLRepository.html`

[3] H. Bustince, J. Montero, M. Pagola, E. Barrenechea and D. Gomez: A survey of Interval-Valued Fuzzy Sets. *In: Handbook of Granular Computing*, Addison Wesley, pp. 491 – 516, (2008)

[4] Z. Chi, H. Yan and T. Pham: Fuzzy algorithms with applications to image processing and pattern recognition. *World Scientific* (1996)

[5] O. Cordón, M. J. del Jesus, and F. Herrera: A proposal on reasoning methods in fuzzy rule-based classification systems. *International Journal of Approximate Reasoning*, vol. 20, no. 1, pp. 21–45, (1999)

[6] J. Demšar: Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, vol. 7, pp. 1–30, (2006)

[7] L. J. Eshelman: The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination. *In: Foundations of Genetic Algorithms*, Morgan Kaufman, pp. 265–283, (1991)

[8] S.García and F. Herrera: An extension on "statistical comparisons of classifiers over multiple data sets" for all pairwise comparisons. *Journal of Machine Learning Research*, vol. 9, pp. 2677–2694, (2008)

[9] M. B. Gorlzakczany: A method of inference in approximate reasoning based on interval-valued fuzzy sets. *Fuzzy Sets and Systems*, vol. 21, pp. 1–17, (1987)

[10] F. Herrera and L. Martínez: A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems*, vol. 8, no. 6, pp. 746–752, (2000)

[11] F. Herrera, M. Lozano and A. M. Sánchez: A taxonomy for the crossover operator for real-coded genetic algorithms: An experimental study. *International Journal of Intelligent Systems*, vol. 18, pp. 309–338, (2003)

[12] S. Holm: A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, vol. 6, pp. 65–70, (1979)

[13] H. Ishibuchi, T. Nakashima and M. Nii: Classification and modeling with linguistic information granules: Advanced approaches to linguistic Data Mining. *Springer-Verlag* (2004)

[14] H. Ishibuchi and T. Yamamoto: Rule weight specification in fuzzy rule-based classification systems. *IEEE Transactions on Fuzzy Systems*, vol. 13, pp. 428–435, (2005)

[15] R. Sambuc: Function ϕ-flous, application a l'aide au diagnostic en pathologie thyroidienne. *Ph.D. thesis*, University of Marseille. (1975)

[16] J. Sanz, A. Fernandez, H. Bustince and F. Herrera: A First Study on the Use of Interval-valued Fuzzy Sets with Genetic Tuning for Classification with Imbalanced Data-sets. *Lecture Notes on Computer Science*, vol. 5572, pp. 581 – 588, (2009)

[17] D. Sheskin: *Handbook of parametric and nonparametric statistical procedures*. Chapman & Hall/CRC, second edition, (2006)

[18] L. A. Zadeh: Fuzzy sets. *Information and Control* vol. 8, no. 3, pp. 338–353, (1965)