A Multi-objective Evolutionary Algorithm with an Interpretability Improvement Mechanism for Linguistic Fuzzy Systems with Adaptive Defuzzification

Antonio A. Márquez, Francisco A. Márquez and Antonio Peregrín

***Abstract*— In this paper we propose a multi-objective evolutionary algorithm with a mechanism to improve the interpretability in the sense of complexity for Linguistic Fuzzy Rule based Systems with adaptive defuzzification. The use of parameters in the defuzzification operator introduces a series of values or associated weights to each rule, which improves the accuracy but increases the system complexity and therefore has an effect on the system interpretability. To this end, we use maximizing the accuracy as an usual objective for the evolutionary process, and we define objectives related with interpretability, using three metrics: minimizing the classical number of rules, the number of rules with weights associated and the average number of rules triggered by each example. The proposed method was compared in an experimental study with a single objective accuracy-guided algorithm in two real problems showing that many solutions in the Pareto front dominate those obtained by the single objective-based one.**

I. Introduction

At present, the problem of finding a balance between interpretability and accuracy in Linguistic Fuzzy Rule-Based Systems (FRBS) has led to increased interest in methods that take into account both aspects [1] , [2]. Of course, the ideal would meet both criteria to a great extent, but because they are contradictory characteristics, this is not usually possible. One way to achieve it is to improve the accuracy of the system while maintaining the interpretability as far as possible to an acceptable level [2].

In this context, the adaptive inference system and adaptive defuzzification methods have proven to be two important elements that could easily improve the accuracy of the system [3]- [5]. This is because they look for the best way to infer and defuzzify the contribution of each rule [4] and promote the cooperation between the rules [3]. These can also specially adapt the behaviour of fuzzy operators to the Rule Base (RB) or learning the fuzzy operators and RB together [5], obtaining a positive synergy between the two elements that allows the system to achieve a high level of accuracy with a minor loss of interpretability.

For this model, in order to take into account the interpretability issues it is necessary to use a measure to quantify the interpretability of the FRBS with adaptive defuzzification method. This kind of measure could be used as an additional objective to optimize the interpretability of the adaptive parameters when the learning process is carried out. In the literature, many authors work on the difficult trade-off between accuracy and interpretability of FRBSs, obtaining linguistic models not only accurate but also interpretable. We may distinguish two kinds of approaches to manage the interpretability [6, 7]:

1. Global Interpretability (system structure): The complexity of the model (usually measured as number of rules, variables, labels per variable, etc.).

2. Local Interpretability (understanding of the model): A measure of semantic interpretation (inference system used, defuzzification method, conjunction operator, fuzzy partitions, etc.)

Therefore, the adaptive defuzzification methods [4] introduce a loss of interpretability in principle in the local sense, with implications for overall meaning, while the parameters associated with each rule increase the complexity of the system structure with the inclusion of weights. It is therefore desirable to minimize the elements that increase the system complexity.

In this work, we focus our attention on measuring the interpretability of the FRBS with adaptive defuzzification. We propose, first, a mechanism to improve the interpretability in these systems. This mechanism is based on two indexes needed to eliminate the weight of the rule and remove some rules in the same evolutionary learning process of adaptive defuzzification. Another index proposed in this paper, based on the average number of rules that are triggered by each example, will also improve the interpretability [8].

Secondly, we propose to use a Multi-objective Genetic Algorithm (MOEA) [9, 10] in order to obtain a set of accurate and interpretable linguistic fuzzy models using the adaptive defuzzification with three objectives (maximize accuracy (minimize error) and minimize the two interpretability indices discussed above). In this way, we obtain a set of solutions with a different balance between accuracy and interpretability [11]-[14].

To show the good performance of the proposed method, it is compared with a single objective accuracy-guided adaptive defuzzification algorithm [14] by applying both of them to initial linguistics models obtained from automatic

This work was supported by Projects TIN2008-06681-C06-06 and P07-TIC-03179.

A.A. Márquez, F. A. Márquez and A. Peregrín are with the Department of Information Technologies, University of Huelva, 21819 Palos de la Fra. Huelva, Spain (e-mail: amarquez@dti.uhu.es ; alfredo.marquez@dti.uhu.es; peregrin@dti.uhu.es).

learning methods. Two real world problems with different complexities were considered, showing that the solutions of the accuracy based algorithm are dominated by those obtained by our model. It can be seen that both objectives required are certainly contradictory as the obtained Pareto fronts clearly moves from the most accurate solutions to the most interpretable ones.

The paper is structured as follows: Section 2 describes the adaptive Defuzzification used, its components and effects. Section 3 presents the foundations of mechanism and index to improve the interpretability with adaptive defuzzification methods. Section 4 presents the multi-objective model, describing the main characteristics and genetic operators considered. Section 5 is devoted to an experimental study that used the techniques described above in two real problems, to finally present the conclusions of the study developed.

II. ADAPTIVE DEFUZZIFICATION METHODS

There are various tendencies in the development of adaptive defuzzification methods reported in the literature. These employ one or more parameters in their expression to modify the behaviour of the defuzzifier or, in most cases, to achieve higher accuracy.

Following the studies developed in [4], and because of its easy implementation and good performance, in the present work we considered using a method based on Adaptive Defuzzification (1):

$$y_0 = \frac{\sum_i^N f(h_i) \cdot V_i}{\sum_i^N f(h_i)}, \qquad (1)$$

where h_i is the matching degree between the input variables and the rule antecedent fuzzy sets, $f(h_i)$ is a functional of matching degree and V_i represents a characteristic value of the fuzzy set inferred from rule R_i , the Maximum Value (MVi) or the Centre of Gravity (CG). As shown, it is an expression of defuzzification method acting in Mode B, i.e. the first defuzzifier individual contribution of each rule is inferred and then the final result is computed, namely by a weighted sum.

Specifically, we considered the use of a functional term product $f(h_i) = h_i \cdot \alpha_i$ where α_i corresponds to a parameter for each rule R_i, i = 1 to N, as well as the Centre of Gravity (CG) as characteristic value, due to its computational efficiency and similar results in other types of functions [4]. The expression of the Adaptive Defuzzification method is shown in (2)

$$y_0 = \frac{\sum_i^N h_i \cdot \alpha_i \,.\, CG_i}{\sum_i^N h_i \,.\, \alpha_i}, \qquad (2)$$

The role of the individual parameter is interpreted as a modulation of the matching influence, which can be improved or attenuated. We should note that this modulation is only linear for the product case. Particularly, the study of the effect of α zone is as follows

$$\circ\, \alpha_i \cdot h_i,\ \alpha \in [1,\infty)\text{:empowerment of } h_i, \qquad (3)$$
$$\circ\, \alpha_i \cdot h_i,\ \alpha \in [0,1]\text{: penalty of } h_i.$$

The product functional term with a different parameter for each rule has the effect of weighted rules [15]. The α_i value associated with rule R_i acquires the meaning of how significant or important that rule is for the inference process. An improved accuracy is the system modelling goal when using this kind of rule. The following is an example of a set of weighted rules, where the weights are α_i :

R_1 : If X_{11} is A_{11} and …. and X_{1m} is A_{1m} then Y is B_1 with α_1

R_2 : If X_{21} is A_{21} and …. and X_{2m} is A_{2m} then Y is B_2 with α_2

...

R_n : If X_{n1} is A_{n1} and …. and X_{nm} is A_{nm} then Y is B_n with α_n

The rule weight adaptation process will produce a rule subset with better cooperation among the rules composing it [16]. This fact has shown to be of special interest when the rule set has been generated using a quick data-driven fuzzy rule generation method. These methods usually look for the best individual rule performance, and generate a linguistic RB with a low cooperation degree. Using the product functional and a parameter learning process will be equivalent to looking for a subset of rules with the best global cooperation.

Overall, the influence of rule weights on the interpretability of fuzzy systems is usually discussed. Some authors consider they can be equivalently replaced by modifications in the membership functions in order to avoid negative effects on the interpretability [16], while others claim the importance of weights as a degree of certainty and their importance in some problems [17], [18]. The product functional term with a different parameter for each rule has the effect of weighted rules. This value associated with the rule indicates the importance of that rule for the inference process.

III. MECHANISM FOR INTERPRETABILITY IMPROVEMENT IN LINGUISTIC FUZZY SYSTEMS WITH ADAPTIVE DEFUZZIFICATION METHODS

In this section, we propose a mechanism to improve the interpretability and several metrics to measure it when an adaptive defuzzification method with a product functional term is used in a linguistic fuzzy system.

As stated in the previous section, the use of adaptive defuzzification methods with functional product type has an effect equivalent to the use of rules with weights [15]. These weights in the rules have a negative effect on interpretability [6, 7] and extend the system structure complexity. For this reason, in order to reduce this negative effect, the mechanism and the metrics used should take these weights into account. At this point, we should remark that the mechanism and metrics are based on the influence of weights in the rules.

A. Mechanism to improve the interpretability

The mechanism to be described is based on two concepts:

1. First, those rules with weights close to 0 represent a low influence of that rule and therefore could indicate a dispensable rule, continuing the evolutionary learning from the rest of the system without it. Deleting a rule in the learning phase allows the evolutionary model to adjust the remaining weights to ignore that rule. Results obtained working in this way are different than those obtained by eliminating the rules with low weight after the evolutionary process, because lower weights are also important for the systems accuracy.
2. Secondly, the weight values close to 1 are those in which the rule is important and we consider that they could be used without any weight, and so remove this value in order to reduce the complexity of the system. The greater the number of rules without a weight, the better the system interpretability.

To apply these concepts, the parameters of adaptive defuzzification method used will perform in the range [0,1] (see expression 2 and 3).

On the other hand, we establish two thresholds U_0, U_1 that define when rules are removed or act without weight, respectively (see Figure 1).

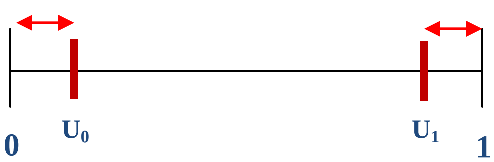

Fig. 1. Range of parameters for adaptive defuzzification and thresholds for mechanism for improvement.

B. Metrics proposed

As a result of improved mechanisms described earlier, we propose the following three metrics:

- Number of final rules ($\#R_F$))

This metric is based on the first idea: those rule weights close to 0 (between U_0 and 0) represent a low influence of that rule and therefore could indicate a dispensable rule, continuing the evolutionary process learning the rest of the system without it.

The expression of this index is:

$$\# R_F = \# R - (\text{number of rule weights close to } 0) \qquad (4)$$

where # R is the number of initial rules of the system.

- Number of rules with weight ($\#R_W$))

This metric however, is based on the second idea: weight values close to 1 (between U_1 and 1) are those in which the rule is important and could be used without any weight, and so remove this value, thus reducing the system complexity.

The expression of this index is:

$$\# R_W = \# R - (\text{number of rule weights close to } 1) \qquad (5)$$

- Average number of rules triggered by each example (MR_{TG})

The reduction in the number of final rules and the number of rules with weight improves the interpretability of the system. However, the interpretability of the system depends on the number of rules triggered at the same time, i.e., a maximum of seven rules is lesser interpretable than four. So we define a new index that will measure the average number of rules triggered by each example.

The expression of this index is:

$$MR_{TG} = \frac{\sum_{j}^{M} R^{j}_{TG}}{M}, \qquad (6)$$

where M is the number of examples and # R^{j}_{TG} is the number of rules triggered by the example j.

C. A global Interpretability index based on the aggregation of two metrics

In the present work, we propose an aggregation of two of the metrics ($\#R_W$ and MR_{TG}) in a global index based on the arithmetic mean, which is denoted as $R_W\_MR_{TG}$ index.

First, the proposed indexes are normalized between 0-1. The aggregation operator should consider this fact:

$$R_P\_MR_{TG} = \frac{\#R_P + MR_{TG}}{2}, \qquad (7)$$

The value of $R_W\_MR_{TG}$ ranges between 0 (the highest level of interpretability) and 1 (the lowest level of interpretability).

IV. MULTI-OBJECTIVE EVOLUTIONARY MODEL PROPOSED

This section describes the proposed multi-objective algorithm which uses the mechanism to improve the interpretability in adaptive defuzzification systems described in the previous Section.

The proposed algorithm uses three objectives: minimizing two indexes, # R_F and $R_W\_MR_{TG}$, to improve interpretability, and minimizing one index, the error, in order to improve the accuracy. This time we use an evolutionary model based on the popular NSGA-II [19]. In the following subsections we state the main components and parameters of this algorithm.

A. Coding scheme and initial population

In this paper, we use a real coding scheme, where m is the number of parameters α_i, one for each of the RB R_i. Each takes values in the interval [0,1].

$$C = (\alpha_1, \ldots, \alpha_m) \mid \alpha_i \in \{0, 1\}$$

The initial population is obtained as follows: An individual of the initial population has all the genes initially set to 1 in order to begin the evolutionary process with all the rules without weight. The remaining individuals of the initial population are created randomly.

B. Objectives and thresholds

As was discussed above, in this algorithm we use these objectives to minimize:

- The index of interpretability (# R_F) representing the number of final rules in the system.
- The index of interpretability ($R_W\_MR_{TG}$) representing the mean arithmetic proposed.
- The Mean Square Error (MSE) which measures the accuracy of the system.

$$\text{MSE}\ (S[i]) = \frac{\frac{1}{2} \sum_{k=1}^{N} (y_k - S[i](x_k))^2}{N} \qquad (7)$$

where *S [i]* denotes the fuzzy model the inference system of which uses the minimum t-norm as conjunction operator, the inference operator is minimum t-norm, and the adaptive defuzzification method is the one shown in expression (2). This measure uses a set of system evaluation data formed by P pairs of numerical data $Z_k = (x_k, y_k)$, k=1,..,P, with x_k being the values of the input variables, and y_k being the corresponding values of the associated output variables.

The thresholds considered for the indexes # R_F and #R_W are the following:

- If the parameter value is ≥ 0.9, it is considered that this parameter is 1 and therefore this rule has no weight.
- If the parameter value is ≤ 0.1, we evaluate this parameter like 0 and the rule will be eliminated.

Both thresholds have been chosen empirically by testing.

C. Multi-objective Genetic Algorithm population

NSGA-II [19] is one of the most known and used MOEA in the literature for solving multi-objective problems. The offspring population is generated from the current population through selection, crossover and mutation. The next generation is built from the current population and the offspring until it reaches the stop condition in this work, number of evaluations. The NSGA-II algorithm has two features that make it one of the main and most important MOEA: One is the assignment of fitness based on Pareto ranking and crowding operator, and the other is the procedure for updating each generation through elitism.

V. EXPERIMENTS

To assess the goodness of the proposed approach, two real world problems with different complexities (different number of variables and available data) to be solved are considered (these data sets are available at, http://www.keel.es/) [20]:

- Estimating the maintenance costs of medium voltage lines in a town (ELE).
- Predicting the Abalone Age (ABA).

In both cases, the well-known ad hoc data-driven learning algorithm of Wang and Mendel [21] is applied to obtain an initial set of candidate linguistic rules. To do so, we will consider triangular-shaped strong fuzzy partitions. Once the initial RB is generated, the proposed algorithm can be applied.

Methods considered for the experiments are:

- D: adaptive defuzzification method using a mono-objective genetic algorithm and considering only the sole purpose of accuracy [4].
- D_I: is used by the adaptive defuzzification with the proposed improvement on interpretability and considering three objectives to minimize: MSE, #RF and $R_W\_MR_{TG}$.

A. Experimental Set-up

We consider a 5-fold cross-validation model, i.e. 5 random partitions of data each with 20%, and the combination of 4 of them (80%) as training and the remaining one as test. For each of the 5 data partitions, the methods were run 6 times, showing for each problem the averaged results of a total of 30 runs.

Linguistic partitions considered consist of triangular shaped linguistic terms, 5 for the electrical problem and 3 for the Abalone. Minimum t-norm was used as operator of conjunction and implication. In the case of MOEA (D_I) the averaged values are calculated considering the most accurate solution from each Pareto obtained. In this way, D_I can be compared with the single objective method D.

The values of the input parameters considered by D are: population size of 61, 1000 generations, 0.6 as crossover probability and 0.2 as mutation probability per chromosome. The values of the input parameters considered by D_I are: population size of 200, external population size of 200, 1000 generations and again 0.2 as mutation probability.

B. Results and Analysis

Table I shows the results obtained with WM, where #R stands for the number of rules, $MSE_{tra/tst}$ for the averaged error obtained over the training/test data, $\#R_F$, R_W and MR_{TG} for the interpretability index. This method obtains the initial knowledge bases that will be used by D and D_I. WM takes to index $\#R_W$ the highest level of interpretability representing that all rules do not contain weights. This index affects $R_W\_MR_{TG}$ index proposed too.

TABLE I
RESULTS OBTAINED BY WM METHOD

Dataset	**#R**	**MSE_{tra}**	**MSE_{tst}**	$\#R_F$	$R_W\_MR_{TG}$	$\#R_W$	MR_{TG}
ELE	65	56136	56359	65	0.38	0	10.7
ABA	68	8.407	8.422	68	0.42	0	15.6

The results obtained by both adaptive evolutionary defuzzification methods are shown in Table II, using the most accurate solution for the MOEA. We also show *t*, which represents the results of applying a Student t-test (with 95 percent confidence) to ascertain whether differences in the performance of the best results are significant when compared with those of the other algorithm in the table. The interpretation of the t column is:

* represents the best averaged result.

\+ means that the best result has better performance than that of the related row.

Analysing the results shown in Table 2, we can highlight the following facts:

- The proposed method D_I obtains the best results in training and test compared with D in both problems, when using the most accurate solution for D_I. In the ELE problem, D_I improves by about 5% and 4% in training and test, respectively, and in the ABA problem it obtains an improvement of around 6 %
- In terms of interpretability, the solution included in Table II, while only being the most accurate for D_I, clearly obtains more interpretable models. D_I results improve all the interpretability indices employed. The two new ones, the number of rules without weights and the medium number of rules triggered by an example, have shown their usefulness with significant improvements for both problems.

Figure 2 shows the Pareto front obtained with D_I method, and the solution obtained by D in the same data partition and seed of ABA. We can observe that the obtained Pareto front is quite wide. In fact, the number of non dominated solutions is always equal to the external population size. Moreover, the solution obtained with D is dominated by several solutions from D_I. Furthermore, there is no overfitting in the results obtained with the proposed method.

The Pareto front obtained allows selecting solutions with different degrees of accuracy and interpretability. Figure 2 represents that an improvement in any interpretability index produces lack of precision and an improvement in the precision produces lack of interpretability. This figure clearly shows that both targets (accuracy and interpretability) are actually contradictory. In the extremes of the Pareto front, an improvement in one objective represents a small loss in the other objective. On the contrary, in the mid part of the Pareto front, improvements in one objective deteriorate the other objectives.

Figure 3 presents an illustrative RB obtained with D_I. We can observe the rules eliminated, the weights of the rules and the rules without weight (rules with weight equal 1).

VI. CONCLUSION

Adaptive defuzzification using weighting factors in the degree of matching through a product is a simple mechanism to improve the accuracy of linguistic fuzzy models significantly, but has the disadvantage of increasing the system complexity, resulting from the effect of adding different factors or weights in each of the rules of the knowledge base. In order to decrease this effect, in this paper we introduce a mechanism for reducing complexity using thresholds for those weights, so when the weight has a high level, the level is eliminated and when the weight has a low level, the rule is removed.

Using a multi-objective evolutionary algorithm to learn the parameters leads to greater precision, fewer rules, fewer rules with weights and fewer average number of rules triggered per example. Thus, interesting results are obtained by significantly reducing the complexity compared with conventional adaptive defuzzification, maintaining and also improving the accuracy, as shown in the experimental study. Moreover, future works will consider also including other RB learning methods and interpretability indexes such as inconsistency, redundancy and similarity rules. Furthermore, we pretend to study the influence of the thresholds in the interpretability and accuracy.

TABLE II

RESULTS OBTAINED FOR THE TWO PROBLEMS

Dataset	Method	MSE_{tra}	t-test	MSE_{tst}	t-test	$\#R_F$	$R_W\_MR_{TG}$	$\#R_W$	MR_{TG}
ELE	D	32791	+	35862	+	65	0.87	65	10.7
	D_I	32439	*	35484	*	45.3	0.46	34.3	6.0
ABA	D	4.823	+	4.826	+	68	0.91	68	15.6
	D_I	4.790	*	4.809	*	22.5	0.51	11.1	4.3

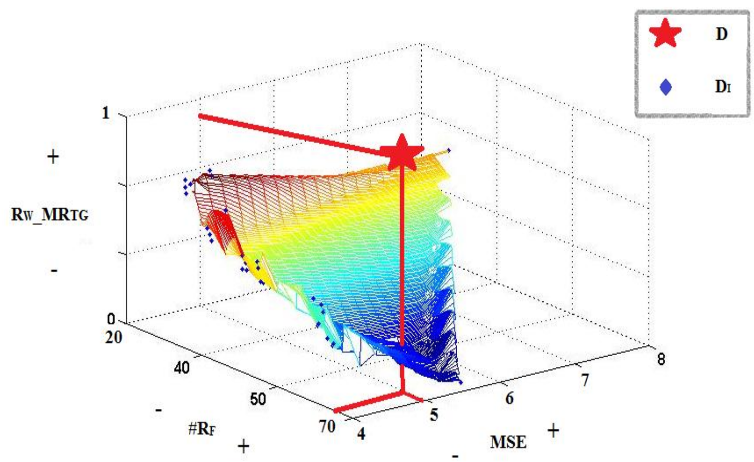

Fig. 2. Example of the Pareto front with D_I for the Abalone problem

.

Delete this rule	MSE tra	31789	#RF	45	MRTG 5.9
	MSE tst	34891	#RP	34	RP_MRTG 0.46

Rules	Weights D_I method	Rules	Weights D_I method	Rules	Weights D_I method
R1	0.6	R23	0.2	R45	1.0
R2		R24		R46	
R3	1.0	R25		R47	
R4	0.3	R26	0.3	R48	0.5
R5	1.0	R27	0.6	R49	0.8
R6	0.8	R28	1.0	R50	0.7
R7		R29	0.8	R51	1.0
R8		R30	0.7	R52	0.5
R9	0.5	R31	0.2	R53	0.5
R10	0.4	R32		R54	1.0
R11		R33	0.2	R55	1.0
R12		R34	1.0	R56	
R13	1.0	R35		R57	
R14	0.6	R36	0.3	R58	1.0
R15	0.3	R37	0.5	R59	0.6
R16	0.3	R38	0.7	R60	
R17	1.0	R39	0.8	R61	0.8
R18	0.7	R40		R62	0.7
R19	0.4	R41	0.4	R63	0.4
R20		R42		R64	
R21	0.8	R43		R65	
R22	0.6	R44	0.8		

Fig. 3. An illustrative RB obtained with D_I for ELE problem

REFERENCES

[1] J. Casillas, O. Cordón, F. Herrera and L. Magdalena. *Interpretability issues in fuzzy modeling*. Springer-Verlag, 2003.

[2] J. Casillas, O. Cordón, F. Herrera and L. Magdalena. *Accuracy improvements in linguistic fuzzy modeling*. Springer-Verlag, 2003.

[3] J. Alcalá-Fdez, F. Herrera, F. Márquez, and A. Peregrín, "Increasing fuzzy rules cooperation based on evolutionary adaptive inference systems", *Int J Intell Syst*, vol. 22 no. 9, pp. 1035–1064, 2007.

[4] O. Cordón , F. Herrera , F. A. Márquez and A. Peregrín, "A study on the evolutionary adaptive defuzzification methods in fuzzy modeling", International Journal of Hybrid Intelligent Systems. Vol. 1, no. 1, pp. 36–48, 2004.

[5] F. A. Márquez, A. Peregrín, and F. Herrera, "Cooperative evolutionary learning of linguistic fuzzy rules and parametric aggregation connectors for Mamdani fuzzy systems", *IEEE Trans Fuzzy Syst*, vol. 15, no. 6, pp. 1162–1178, 2007.

[6] J.M. Alonso and L. Magdalena, "A Conceptual Framework for Understanding Fuzzy Systems" in Proc. of IFSA World Congress (IFSA-EUSFLAT'09) , Lisbon, Portugal, 2009, pp. 119-124.

[7] M. Zhou and J. Q. Gan, "Low-level interpretability and high-level interpretability: a unified view of data-driven interpretable fuzzy system modeling", *Fuzzy Sets and Systems*, vol. 159, no. 23, pp. 3091–3131, 2008.

[8] F. Cheong and R. Lai, "Constraining the optimization of a fuzzy logic controller using an enhanced genetic algorithm*", IEEE Trans. Syst., Man, Cybern. -Part B:Cybernetics*, vol. 30, no. 1, pp. 31-46, 2000.

[9] C. A. Coello, D. A. Van Veldhuizen, and G. B. Lamont, editors. *Evolutionary algorithms for solving multi-objective problems*. Kluwer Academic Publishers, 2002.

[10] K. Deb. *Multi-objective optimization using evolutionary algorithms*. John Wiley & Sons, NY, USA, 2001.

[11] M.J. Gacto, R. Alcalá, and F. Herrera, "Adaptation and Application of Multi-Objective Evolutionary Algorithms for Rule Reduction and Parameter Tuning of Fuzzy Rule-Based Systems," *Soft Computing*,vol. 13, n. 5, pp. 419–436, 2009.

[12] H. Ishibuchi and T. Yamamoto, "Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining", *Fuzzy Sets and Systems* , vol. 141, no.1, pp. 59–88, 2004.

[13] A.A. Márquez, F.A. Márquez, and A. Peregrín. "Cooperation between the Inference System and the Rule Base by using Multiobjective Genetic Algorithms". *Proc of HAIS'08 Int Conf of Hybrid Artificial Intelligence Syst*, Burgos, 2008, pp. 739–746.

[14] A.A. Márquez, F.A. Márquez and A. Peregrín, "Rule Base and Inference System Cooperative Learning of Mamdani Fuzzy Systems with Multiobjective Genetic Algorithms". Proc of IFSA World Congress (IFSA-EUSFLAT'09), Lisbon, Portugal, 2009, pp. Pág. 1045-1050.

[15] J.S. Cho and D.J. Park, " Novel fuzzy logic control based on weighting of partially inconsistent rules using neural network", Journal of Intelligent and Fuzzy Systems, Vol 8, pp. 99-100, 2000.

[16] D. Nauck and R. Kruse, "How the learning of rule weights affects the interpretability of fuzzy systems," in *Proc. Seventh IEEE Int. Conf.Fuzzy Syst.*, 1998, pp. 1235–1240.

[17] H. Ishibuchi and T. Nakashima, "Effect of rule weights in fuzzy rulebased classification systems," *IEEE Trans. Fuzzy Syst.*, vol. 9, no. 4, pp. 506–515, 2001.

[18] H. Ishibuchi and T. Yamamoto, "Rule weight specification in fuzzy rule-based classification systems," *IEEE Trans. Fuzzy Syst.*, vol. 13, no. 4, pp. 428–435, 2005.

[19] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. "A fast and elitist multiobjective genetic algorithm: NSGA-II". *IEEE Trans Evolut Comput*, vol. 6, no 2, pp. 182–197 , 2002.

[20] J. Alcala-Fdez, L. Sánchez, S. García, M.J. del Jesús, S. Ventura, J.M. Garrell, J. Otero, C. Romero, J. Bacardit, V.M. Rivas, J.C. Fernández, and F. Herrera. *Keel: A software tool to assess evolutionary algorithms for data mining problems*. Soft Computing, vol.13, nº3, p.p. 307–318, 2009.

[21] L.X. Wang, and J.M. Mendel. "Generating fuzzy rules by learning from examples", *IEEE Trans Syst, Man, Cybernetics*, vol. 22, no. 6, pp. 1414–1427, 1992.