Rule Base and Inference System Cooperative Learning of Mamdani Fuzzy Systems with Multiobjective Genetic Algorithms

Antonio A. Márquez Francisco A. Márquez Antonio Peregrín

Information Technologies Department, University of Huelva
Huelva, Spain
Email: {amarquez, alfredo.marquez, peregrin}@dti.uhu.es

Abstract*—In this paper, we present an evolutionary multiobjective learning model achieving positive synergy between the Inference System and the Rule Base in order to obtain simpler, more compact and still accurate linguistic fuzzy models by learning fuzzy inference operators together with Rule Base. The Multiobjective Evolutionary Algorithm proposed generates a set of Fuzzy Rule Based Systems with different trade-offs between interpretability and accuracy in linguistic fuzzy modeling, allowing the designers select the one that involves the most adequate equilibrium for the desired application.*

Keywords— Linguistic fuzzy modeling, interpretability-accuracy trade-off, Multiobjective genetic algorithms, adaptive Inference System, adaptive defuzzification, rule learning.

1 Introduction

The main objective in system modelling is to develop reliable and understandable models. Interpretability and accuracy are usually contradictory requirements in the design of linguistic fuzzy models (FMs). Recent research into genetic fuzzy systems has focused on methods aimed at generating Fuzzy Rule-Based Systems (FRBS) with an appropriate trade-off between accuracy and interpretability [1, 2].

Two important tasks in the design of a linguistic FM for a particular application are the derivation of the linguistic Rule Base (RB) and the setup of the Inference System and defuzzification method. In the framework of the trade-off between interpretability and accuracy in fuzzy modeling, adaptive Inference Systems and defuzzification methods have acquired greater importance [3, 4].

Recently, the use of Multiobjective Evolutionary Algorithms (MOEA) has been applied to improve the aforementioned trade-off between interpretability and accuracy of linguistic fuzzy systems [5, 6, 7, 8, 9, 10, 11]. Some of them obtain the complete Pareto (the set of non-dominated solutions with different trade-offs) by selecting or learning the set of rules which best represents the example data, i.e., improving the system accuracy and decreasing the FRBS complexity. In [7, 8, 11] the authors also propose tuning the membership functions together with the rule selection to obtain simpler yet still accurate linguistic FMs.

Following these ideas on the advantage of using parametric operators and MOEAs to improve the trade-off between interpretability and accuracy, in [9] we presented a MOEA capable of learning the fuzzy inference operators (including inference and defuzzification) and of performing rule selection for Mamdani linguistic fuzzy systems. The proposed model aimed to achieve a positive synergy, that is, cooperation between the fuzzy operators and the RB to improve accuracy while at the same time simplifying the RB to improve interpretability.

Our main objective in this work is to include a new and highly important element in the learning process, the complete RB. Thus, we propose a MOEA capable of generating a set of FRBS (each with a high degree of cooperation between the RB and the Inference System) with varying optimal trade-offs between accuracy and complexity, so as to obtain compact and accurate linguistic fuzzy models by learning fuzzy operators and RB.

To do this, Section 2 describes the parametric fuzzy operators, Section 3 shows the RB learning used in this work, Section 4 is devoted to describing the MOEA learning proposal, Section 5 develops an experimental study, and finally, Section 6 presents some concluding remarks.

2 Adaptive Fuzzy Operators

In this section we describe the adaptive Inference System as well as the adaptive defuzzification method used in our learning proposal.

2.1 Adaptive Inference System

Linguistic FRBSs for system modeling use IF - THEN rules of the following form:

$$R_i : \text{If } X_{i1} \text{ is } A_{i1} \text{ and ... and } X_{im} \text{ is } A_{im} \text{ then } Y \text{ is } B_i$$

with i = 1 to N, where N stands for the number of rules of the RB, X_{i1} to X_{im} and Y for the input and output variables respectively, and A_{i1} to A_{im} and B_i for the involved antecedents and consequent labels, respectively.

The expression of the Compositional Rule of Inference in fuzzy modeling with punctual fuzzification is the following: $\mu_{B'}(y) = I(C(\mu_{A1}(x_1), \ldots, \mu_{Am}(x_m)), \mu_B(y))$, where $\mu_{B'}(\cdot)$ is the membership function of the inferred consequent, $I(\cdot)$ is the implication operator, $C(\cdot)$ is the conjunction operator, $\mu_{Ai}(x_i)$ are the values of the matching degree of each input of the system with the membership functions of the rule antecedents, and $\mu_B(\cdot)$ is the consequent of the rule.

ISBN: 978-989-95079-6-8

The two components, the conjunction (C(·)) and the implication operator (I(·)) are suitable for parametrization in order for the Inference System to be adapted. Our previous studies in [3] show that models based on the adaptive conjunction is a more valuable option than those based on the adaptive implication operator. Hence, we selected the adaptive conjunction in this study in order to insert parameters in the Inference System.

Taking into account the aforementioned studies in [3], we have selected the Dubois adaptive t-norm with a separate connector for every rule, the expression for which is shown in (1).

$$T_{Dubois}(x,y,\alpha) = \frac{x \cdot y}{Max(x,y,\alpha)}, \quad (0 \le \alpha \le 1) \qquad \textbf{(1)}$$

This adaptive t-norm showed the highest accuracy in previous studies, compared with Frank and Dombi t-norms and is more efficiently computed. The use of an adaptive t-norm for the antecedent connection seeks better performance than traditional t-norms. Dubois t-norm performs between minimum ($\alpha = 0$) and algebraic product ($\alpha = 1$).

2.2 Adaptive Defuzzification Interface

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

Following the studies developed in [13], in this work we consider applying the defuzzification function to the fuzzy set inferred by each rule (getting a characteristic value) and computing them by a weighted average operator, because of its fine performance, efficiency and easier implementation. This way of working is named FITA (First Infer, Then Aggregate) [12].

We also consider the use of a product functional term of the matching degree between the input variables and the rule antecedent fuzzy sets (h_i), $f(h_i) = h_i \cdot \beta_i$ where β_i corresponds to one parameter for each rule R_i, i=1 to N, in the RB, as it is more efficiently computed and obtains similar results to other functions [13]. The adaptive defuzzification formula selected is shown in (2).

$$y_0 = \frac{\sum_{i}^{N} h_i \cdot \beta_i \cdot V_i}{\sum_{i}^{N} h_i \cdot \beta_i}, \qquad \textbf{(2)}$$

where V_i represents a characteristic value of the fuzzy set inferred from rule R_i, the Maximum Value or the Gravity Center (GC), the latter being the one selected in this paper.

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.

3 Rule Base Learning

The linguistic RB learning used in this work is based on the ad-hoc data driven methodology named COR [14]. This methodology manages a set of consequent label sets (one per rule). Instead of selecting the consequent with the best performance in each subspace as usual (Wang and Mendel [15]), the COR methodology considers the possibility of using another consequent, different from the best, which allows the FRBS to be more accurate thanks to having a RB with best cooperation. For this purpose, COR performs a combinatorial search among the candidate rules looking for the set of consequents which globally achieves the best accuracy.

COR consists of two stages:

1) *Construction of the search space*—This obtains a set of candidate consequents for each rule.

2) *Selection of the most cooperative fuzzy rule set*—This performs a combinatorial search among these sets seeking the combination of consequents with the best global accuracy.

In order to perform this combinatorial search, an *explicit enumeration* or an *approximate search technique* can be considered. In this work, we use a search technique because it is effective and quick.

4 Rule Base and Inference System Cooperative Learning with Multiobjective Algorithms

This Section describes the evolutionary multiobjective model proposed in this work. As was previously mentioned, our objective is to obtain a set of fuzzy systems with different trade-offs between accuracy and interpretability, using adaptive inference and defuzzification, and Rule Base learning (including rule selection). To do this, we exploit two specific MOEAs considering a threefold coding scheme (coding of rules and coding of the parameters of the Inference Systems and Defuzzification). We adopted two of the most representative second generation MOEAs, SPEA2 [16] and NSGA-II [17], as two general purpose MOEAs for performing the cooperative adaptation of the fuzzy operators and fuzzy rule learning.

4.1 SPEA2 and NSGA-II

The SPEA2 algorithm [16] (*Strength Pareto Evolutionary Algorithm for Multiobjective Optimization*) is one of the most well-known techniques for solving multiobjective problems. It is characterized by the following two aspects: a *fitness* assignment strategy, which takes into account both dominating and dominated solutions for each individual, and a density function, estimated by employing the nearest neighbourhood, which guides the search more efficiently.

NSGA-II algorithm [17] is another of the most well-known and frequently-used MOEAs for general multi-objective optimization in the literature. It is a parameterless

ISBN: 978-989-95079-6-8

approach with several interesting principles: a binary tournament selection based on fast non-dominated sorting, an elitist strategy and a crowding distance method to estimate the diversity of a solution.

4.2 Questions related to the MOEAs.

The evolutionary model uses a chromosome with threefold coding scheme (C_C+C_D+C_R) where:

- C_C encodes the α_i parameters of the conjunction connective. They are N real coded parameters (genes), one for each rule, R_i ,of the linguistic RB. Each gene can take any value in the interval [0, 1], that is, between the minimum and the algebraic product. This is represented by the C_C part of the chromosome shown in Figure 1.

- C_D encodes the β_i parameters of the defuzzification. They are N real coded parameters, one for each rule, of the linguistic RB. Each gene can take any value in the interval [0, 10]. This interval has been selected according to the study developed in [13]. It allows attenuation as well as enhancement of the matching degree. This is represented by the real part C_D shown in Figure 1.

- C_R encodes the learning Rule Base. It is an integer string of N genes, each one representing a candidate rule consequent of the initial RB. Furthermore, depending on whether a rule is selected or not, the value '-1' is assigned to the corresponding gene. This is represented by the integer part of the chromosome shown in Figure 1.

The initial population is randomly initialized in the fuzzy operators part with the exception of a single chromosome:

- C_C with the N genes is initiated to 0 in order to make Dubois t-norm equivalent to Minimum t-norm initially.
- C_D also with the N genes is initiated to 1 with the objective of beginning like the standard WCOA method.

The initial population in the fuzzy rule part, C_R, is initialized following these two exceptions:

- A single chromosome with the N rules obtained by the WM-method [15], that is, with all the genes initialized to correspondent consequent.

- Default chromosomes randomly initiated with all rules activated. In this case, in order to achieve solutions with a high accuracy we should not lose rules that could present a positive cooperation once their FM parameters have been evolved. The best way to do this is to start with solutions that select all the possible rules. This favors a progressive extraction of bad rules (those that do not improve with the tuning of parameters).

The crossover operator employed by the fuzzy operators part is BLX-0.5 [18] while the one used for the rule learning part is HUX [19].

Finally, four offspring are generated by combining the two from the C_R part with the two from the operators part (the two best replace their parents). The mutation operator changes a gene value at random in the C_R and operators part (one in each part) with probability 0.2.

In this work, to obtain an optimal set of FRBS with different trade-offs, the fitness, based on the interpretability (using the number of rules) and the accuracy (using the error measure), must be minimized.

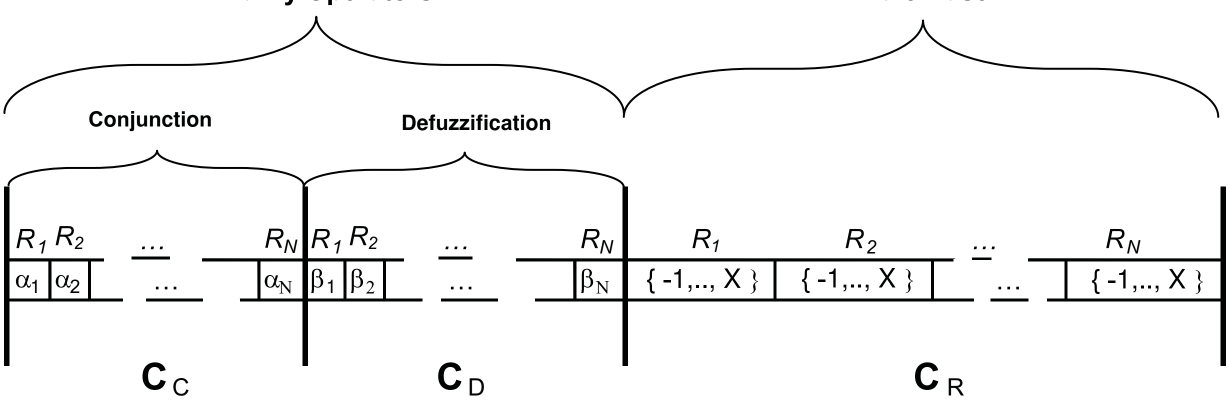

Fig. 1. Coding scheme for the MOEA with N rules

ISBN: 978-989-95079-6-8

5 Experimental Study

In order to analyse the practical behaviour of the proposed methods, we built several FMs in a real-world problem [20] of four input variables consisting in estimating the maintenance costs of medium voltage lines in a town. Methods considered for the experiments are briefly described in Table 1, where WM and COR methods are considered as reference. S denotes the methods that perform rule selection. If rule selection is performed after another method we give it a "+" denotation (eg. COR+S). However, if rule selection is performed jointly with another method we denote this in subscript (eg. COR_S). S–C–D means rule selection and fuzzy operators learning together. $SPEA2_{S\text{-}C\text{-}D}$, and $NSGA\text{-}II_{S\text{-}C\text{-}D}$ are the methods that learn the fuzzy operators and the rule selection together, while $SPEA2_{COR_{S\text{-}C\text{-}D}}$ and $NSGA\text{-}II_{COR_{S\text{-}C\text{-}D}}$ are methods that also learn the RB as previously mentioned.

Table 1. Methods considered for comparisson

Ref.	Method	Description
[15]	WM	Wang & Mendel algorithm
[21]	WM + S	Wang & Mendel and then Rule Selection
[9]	WM + S–C–D	Wang & Mendel and then Rule Selection and Adaptive Fuzzy Operators
[14]	COR	COR
	COR + S	COR and then Rule Selection
	COR_S	COR with Rule Selection
	$COR_{S\text{–}C\text{–}D}$	COR with Rule Selection and Adaptive Fuzzy Operators
[9]	$SPEA2_{S\text{-}C\text{-}D}$	SPEA2 algorithm with Rule Selection and Adaptive Fuzzy Operators
-	$SPEA2_{COR_{S\text{-}C\text{-}D}}$	SPEA2 algorithm with COR with Rule Selection and Adaptive Fuzzy Operators
[9]	$NSGA\text{-}II_{S\text{-}C\text{-}D}$	NSGA-II algorithm with Rule Selection and Adaptive Fuzzy Operators
-	$NSGA\text{-}II_{COR_{S\text{-}C\text{-}D}}$	NSGA-II algorithm with COR with Rule Selection and Adaptive Fuzzy Operators

5.1 Application Selected and Comparison Methodology

The application selected to test the evolutionary model is the aforementioned electrical distribution problem [20] that has a data set of 1059 cities with four input variables and a single output. The RB is composed of 65 linguistic rules achieved with the Wang and Mendel method [15]. The fuzzy partition used for inputs and output has 5 labels.

We considered a 5-fold cross-validation model, i.e., 5 random partitions of the data each with 20% (4 of them with 211 examples, and one of them with 212 examples), using the combination of 4 of them (80%) as training, and the remaining one as a test. We achieved a total of 30 trials for each evolutionary process, as the learning methods were run 6 times for each one of the data partitions. We show the average values of the medium square error (MSE) as a usual performance measure, computed considering the most accurate solution from each Pareto obtained with the multiobjective algorithm. This method of working was also employed in [9] in order to compare the single objective methods with the multiobjective ones based on considering the accuracy objective only, letting us see that the Pareto fronts are not only wide but also optimal, so similar solutions obtained with the WM + S–C–D or $COR_{S\text{–}C\text{–}D}$ must appear in the final Pareto. The MSE is computed with expression (3),

$$MSE\ (FM)_B = \frac{\frac{1}{2}\sum_{k=1}^{P}(y_k - FM(x_k))^2}{P} \quad \textbf{(3)}$$

where *FM* denotes the fuzzy model the Inference System of which uses the Dubois t-norm as conjunction operator showed in expression (1), the inference operator is minimum t-norm, and the adaptive defuzzificación 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 MOEAs population size was fixed at 200. The external population size of the $SPEA2_{S\text{-}C\text{-}D}$ and $SPEA2_{COR_{S\text{-}C\text{-}D}}$ was 61.

5.2 Results and Analysis

To compare the results obtained we also used non-parametric tests, according to the recommendations made in [22]. The results obtained are shown in Table 2, where #R is the average number of rules, MSE_{tra} and MSE_{tst} are the average MSE for training and test respectively, and *Wilcoxon-test* is the result of applying a *Wilcoxon* signed-ranks test [23] (with 95% confidence), with the following interpretation: * represents the best average result (control algorithm); + means that the best result has better performance than that of the corresponding row, while sign (=) means it is similar to the best result. As we have mentioned, Demšar [20] recommends a set of simple, safe and robust non-parametric tests for statistical comparisons of algorithms, one of which is the Wilcoxon signed-ranks test [23]. This is analogous to the paired t-test in non-parametrical statistical procedures.

Table 2 only shows the best result for each MOEAs for accuracy. Analysing the results we can highlight the two following points:

- The learning of the RB allows remarkable improvement in accuracy: Looking at Table 2, we can observe that $SPEA2_{COR_{S\text{-}C\text{-}D}}$ improves the accuracy of $SPEA2_{S\text{-}C\text{-}D}$ and $NSGA\text{-}II_{COR_{S\text{-}C\text{-}D}}$ improves $NSGA\text{-}II_{S\text{-}C\text{-}D}$. In spite of the fact that the number of rules shown in Table 2 is slightly larger, it must be taken into account that the results shown are the ones with the highest accuracy along the Pareto front. Looking at Table 3, we can observe similar values for accuracy in solutions with a lower number of rules from 40 to 32, so $NSGA\text{-}II_{COR_{S\text{-}C\text{-}D}}$ truly obtains better fuzzy systems (for accuracy and

ISBN: 978-989-95079-6-8

interpretability) than NSGA-II$_{S\text{-}C\text{-}D}$. Thus, the cooperation between the rules and the fuzzy operators improves the results of the evolutionary multiobjective proposal without rule learning.

- The solution with the best accuracy obtained with NSGA-II$_{COR_{S\text{-}C\text{-}D}}$ and SPEA2$_{COR_{S\text{-}C\text{-}D}}$ shows a similar accuracy to the single objective evolutionary model COR$_{S\text{–}C\text{–}D}$ (as shown in the non-parametric test) with a significant reduction in the number of rules, particularly for NSGA-II$_{COR_{S\text{-}C\text{-}D}}$. Consequently, the proposed method achieves more interpretable models with similar accuracy. We also notice that the best accuracy in Table 2 is obtained by the single objective model COR$_{S\text{–}C\text{–}D}$. The difference is small, but we can deduce that the evolutionary multiobjective methods are not achieving the most accurate solution. This fact suggested in [9] the design of more specific multiobjective algorithms in order to get even better solutions than the generic MOEAs SPEA2 and NSGA-II. Figure 2 shows the Pareto progress for each evolutionary algorithm (SPEA2$_{COR_{S\text{-}C\text{-}D}}$ and NSGA-II$_{COR_{S\text{-}C\text{-}D}}$) where we can observe the Pareto movement for each generation. Because to the adaptive fuzzy operators search space is large, we consider it may be necessary to focus the search process on the Pareto zone with highest accuracy, so that the same accuracy can be achieved as with single objective evolutionary algorithms based on accuracy.

Table 2. Results obtained

Method	**#R**	**MSE$_{tra}$**	**Wilcoxon-test**	**MSE$_{test}$**	**Wilcoxon-test**
WM	65	56135.75	+	56359.42	+
WM + S	40.9	41517.01	+	44064.67	+
WM + S–C–D	52.8	22640.95	+	26444.43	+
COR	65	50710.80	+	54584.76	+
COR + S	44.7	40763.48	+	43228.38	+
COR$_S$	43	39530.19	+	41060.99	+
COR$_{S\text{–}C\text{–}D}$	50	**20123.39**	*	**23323.72**	*
SPEA2$_{S\text{-}C\text{-}D}$	**38,60**	24021,41	+	29333,72	+
SPEA2$_{COR_{S\text{-}C\text{-}D}}$	41,10	21254,70	+	24079,32	=
NSGA-II$_{S\text{-}C\text{-}D}$	38,90	23364,63	+	28174,76	+
NSGA-II$_{COR_{S\text{-}C\text{-}D}}$	40,27	**20689,86**	=	**23346,34**	=

Table 3. A Pareto front example obtained from NSGA-II$_{COR_{S\text{-}C\text{-}D}}$

#R	**MSE$_{tra}$**	**MSE$_{test}$**	**#R**	**MSE$_{tra}$**	**MSE$_{test}$**	**#R**	**MSE$_{tra}$**	**MSE$_{test}$**
40	20036,37	22771,09	32	22354,40	24819,15	25	30133,70	42029,38
39	20198,77	22768,15	31	23266,76	25226,77	24	32481,97	43640,18
38	20388,66	22675,47	30	24347,59	27510,49	23	35058,65	43578,74
37	20783,97	23251,03	29	24994,05	26904,67	22	39128,61	53587,74
36	20999,33	23435,43	28	26190,94	29391,74	21	43276,59	58873,29
34	21719,71	24057,71	27	27445,85	30597,32	20	47428,26	69762,81
33	21901,55	24051,67	26	28706,22	33035,11	20	47428,26	69762,81

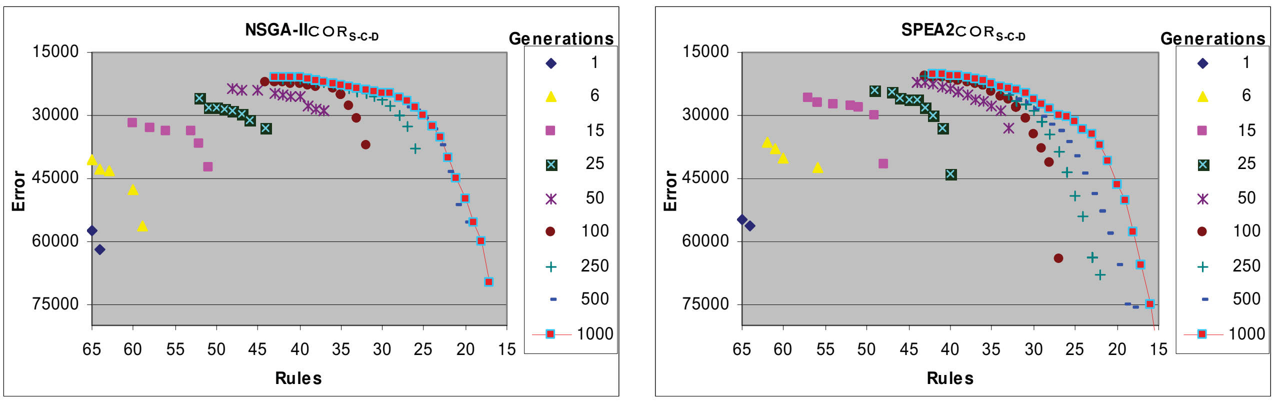

Fig. 2. Example of the Pareto front for NSGA-II$_{COR_{S\text{-}C\text{-}D}}$ and SPEA2$_{COR_{S\text{-}C\text{-}D}}$

.

ISBN: 978-989-95079-6-8

6 Conclusions

In the framework of the trade-off between accuracy and interpretability, the use of MOEAs gives a set of solutions with different levels of conciliation between both features. In this work we have proposed a multiobjective evolutionary learning model where the adaptive fuzzy operator parameters are learnt together with the RB. This fact allows both elements to cooperate, improving the accuracy as well as the interpretability.

The results obtained have shown that the use of MOEAs can represent a way to obtain a set of FRBSs in a single run with optimal trade-off between accuracy and interpretability. In terms of future work, some improvements may be developed in order to guide the search towards the desired Pareto zone with higher accuracy (right and central zone) where the FRBSs obtained are perhaps more interesting in more applications.

By focusing the search process we can reduce the effort of the search, and a better precision in the non-dominated solutions can be obtained, because the search effort is concentrated on a reduced zone of the Pareto, such that the density of the obtained solutions is higher. An improvement could be a change in the MOEAs used or a change in the non-dominated definitions in order to give more weight to the objective of accuracy.

Acknowledgment

Paper Supported by Projects TIN2008-06681-C06-06 and P07-TIC-03179.

References

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

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

[3] J. Alcala-Fdez, F. Herrera, F.A. Márquez, et al. Increasing Fuzzy Rules Cooperation Based On Evolutionary Adaptive Inference Systems. *Int. J. of Intelligent Systems* 22(9): 1035-1064, 2007.

[4] F.A. Márquez, A. Peregrín, F. Herrera. Cooperative Evolutionary Learning of Fuzzy Rules and Parametric Aggregation Connectors for Mamdani Linguistic Fuzzy Systems. *IEEE Trans. on Fuzzy Syst*. 15(6): 1162-1178, 2007.

[5] H. Ishibuchi, T. Yamamoto. Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining. *Fuzzy Sets and Syst*. 141: 59-88, 2004.

[6] K. Narukawata, Y. Nojima, H. Ishibuchi. Modification of evolutionary Multiobjective optimization algorithms for Multiobjective design of fuzzy rule-based classification systems. *Proc. 2005 IEEE Int. Conf. on Fuzzy Syst*. 809-814. Reno , 2005

[7] R. Alcalá, J. Alcala-Fdez, M.J. Gacto, et al. A Multi-Objective Evolutionary Algorithm for Rule Selection and Tuning on Fuzzy Rule-Based Systems. *Proc. of the 16th IEEE Int. Conf. on Fuzzy Syst. FUZZ-IEEE'07*, 1367-1372, London , 2007.

[8] R. Alcalá, M.J. Gacto, F. Herrera, et al. A Multi-Objective Genetic Algorithm for tuning and Rule Selection to obtain Accurate and Compact Linguistic Fuzzy Rule-Based Systems. *Int. J. of Uncertainty, Fuzziness and Knowledge-Based Syst.,* 15(5): 539-557, 2007.

[9] A.A. Márquez, F.A. Márquez, A. Peregrín. Cooperation between the Inference System and the Rule Base by using Multiobjective Genetic Algorithms. *Proceedings of Third Internacional of Hybrid Artificial Intelligence System,(HAIS'08)* 739-746 Burgos, Spain, 2008.

[10] M. Cococcioni, P. Ducange, B. Lazzerini, et al. A Pareto-based Multi-objective Evolutionary approach to the identification of Mamdani Fuzzy Systems. *Soft Computing* 11(11): 1013-1031, 2007

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

[12] J.J. Buckley, Y. Hayashi. Can approximate reasoning be consistent?. *Fuzzy Sets and Syst.,* 65(1): 13-18, 1994.

[13] O. Cordón, F. Herrera, F.A. Márquez, et al. A Study on the Evolutionary Adaptive Defuzzification Methods in Fuzzy Modelling. *Int. J.of Hybrid Intelligent Syst*, 1(1): 36-48, 2004.

[14] J. Casillas, O. Cordón, F. Herrera. COR: A methodology to improve ad hoc data-driven linguistic rule learning methods by inducing cooperation among rules. *IEEE Trans. Syst., Man, Cybern.—Part B: Cybern*. 32(4): 526-537, 2002.

[15] L.X. Wang, J.M. Mendel. Generating fuzzy rules by learning from examples. *IEEE Trans. on Syst., Man, and Cybernetics* 22(6): 1414-1427, 1992.

[16] E. Zitzler, M. Laumanns, L. Thiele. SPEA2: Improving the strength pareto evolutionary algorithm for multiobjective optimization. *Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems (EUROGEN'01*), 95-100, 2001.

[17] K. Deb, A. Pratap, S. Agarwal, et al. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. on Evolutionary Computation* 6(2): 182-197 , 2002.

[18] L.J. Eshelman, J.D. Schaffer. Real-coded genetic algorithms and interval-schemata. *Found. of Genetic Algorithms* 2: 187-202, 1993.

[19] L.J. Eshelman. The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination. *Found. of Genetic Algorithms* 1: 265-283 , 1991.

[20] O. Cordón, F. Herrera, L. Sánchez. Solving electrical distribution problems using hybrid evolutionary data analysis techniques. *Appl. Intell.,* 10, 5-24,1999.

[21] J. Casillas, O. Cordon, M.J. del Jesus, et al. Genetic tuning of fuzzy rule deep structures preserving interpretability and its interaction with fuzzy rule set reduction. *IEEE Trans. on Fuzzy Syst*. 13(1):13–29, 2005.

[22] J. Demšar, Statistical comparisons of classifiers over multiple data sets, *J. Mach. Learn. Res*. 7: 1-30, 2006.

[23] D.J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, CRC Press, Boca Raton, FL, 2003.

ISBN: 978-989-95079-6-8