Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Genetic Tuning for Improving Wang and Mendel's Fuzzy Database

E. R. R. Kato, O. Morandin Jr., M. Sgavioli, B. D. Muniz
Department of Computer Science
Federal University of São Carlos (UFSCar)
São Carlos, Brasil
{kato, orides, mayra_sgavioli, bruno_muniz }@dc.ufscar.br

***Abstract*—This paper aims to use the Genetic Tuning, an evolutionary approach to improving an existing fuzzy database of a fuzzy system, which the rule base is generated by example, through fitness function manipulation. Evolutionary Computing is shown efficient for solving problems where the solution space is too large and its complete analyses is computationally unfeasible. The evolutionary solution will be applied in the adjustment of fuzzy sets generated by the Wang and Mendel's method. Wang and Mendel's method high performance has been clearly demonstrated. Although, inducing genetic tuning, a better set of parameters of the database can be achieved. Our proposal is to find a better representation of the input data attributes and find a minimum distance between the original output of the data set and the output given by the Wang and Mendel's method keeping a good interpretability of the sets and improving the accuracy of the system. To analyze the performance we will use tree different data set, considering Real Coding Scheme to chromosome codification.**

***Keywords*—genetic tuning, fuzzy database, Wang and Mendel's method.**

I. INTRODUCTION

In 1965, Lotfi A. Zadeh introduced the fuzzy sets theory, and it has been applied to different areas and identified as a good solution to predict one or more outputs in accordance with a set of entries. The fact that the input data may be in one or more sets at the same time with different membership degrees differs the fuzzy logic from the classical logic.

A fuzzy system is capable to answers questions relative to the information stored in their base [1].

Various approaches have been presented for the generation of a rule base and the fuzzy sets from numerical data. Neural Networks [2] and genetic algorithms [3, 4] have been used. Other method called "*ad-hoc data driven*" has been proposed. The Wang and Mendel's method (WM's Method) is a widely used ah-hoc data driven method.

Li-Xin Wang and Jerry M. Mendel established in 1992 a method for generating fuzzy rules from examples [5]. This method is able to combine numerical and linguistic information, culminating in a set of rules to be used in a Fuzzy Rule Based System (FRBS).

A FRBS is composed of the knowledge base and the Inference systems. The knowledge base contains the fuzzy rules base and the fuzzy database, which is composed by the fuzzy sets.

Tuning is one of the most widely used approaches to improve the performance of FRBS and consist by the manipulating the shape of the membership function of the set that composes an already defined database [6].

Some works had previously been developed to improve the WM's Method. Reference [7] proposed a modification of the WM's Method to improve the rule cooperation. In [8], the authors introduced a methodology to improve "*ad-hoc data driven*" methods by inducing cooperation among rules. In [6] an evolutionary lateral tuning of membership function was proposed.

The aim of this work is to use genetic tuning to find a minimum distance between the original outputs from a data set from the outputs given by the WM's method and increase the performance and the accuracy of a fuzzy system.

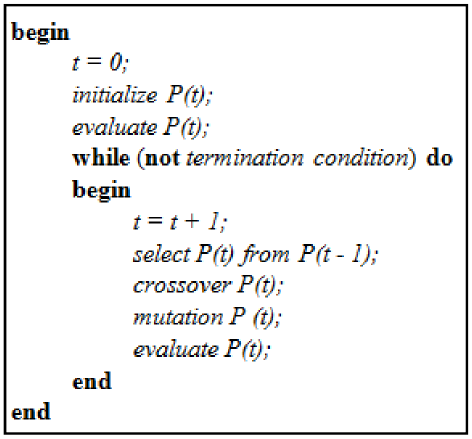
II. GENETIC ALGORITHMS

The Genetic Algorithms were created in the 60's by John Holland. It is a search algorithm based on the mechanisms of genetics and natural selection.

There are three types of operators that are the basis of genetic algorithms: Selection, Crossover and Mutation. The Selection selects individuals from population for reproduction according to the fitness function that represents how much the individual is suitable.

Based on natural selection, genetic algorithms perform a search in the space of solutions by crossover and mutation of possible solutions, generating new solutions that are evaluated. Each solution is called chromosome.

The crossover generates a new chromosome by combining two others. The mutation consists of changing a single chromosome. While the crossover combines two parent chromosomes, mutation creates one child from one parent. The Genetic Algorithm structure can be seen in Fig. 1.



```
begin
    t = 0;
    initialize P(t);
    evaluate P(t);
    while (not termination condition) do
    begin
        t = t + 1;
        select P(t) from P(t - 1);
        crossover P(t);
        mutation P (t);
        evaluate P(t);
    end
end
```

Figure 1. The Genetic Algorithm structure [6].

III. WANG AND MENDEL'S METHOD

The method proposed by Li-Xin Wang and Jerry M. Mendel in 1992 aims to create a base of fuzzy rules combined. The type of information to build a base of fuzzy rules can be classified into two types: numerical, obtained by sensors, measures, and language, obtained by human experts. This method is widely known for its simplicity and good performance. The rule structure considered is the Mamdani type rule with m input variables and one output.

The method consists of five steps:

- Divide the input data and output data into fuzzy regions;
- Generate fuzzy rules from the data;
- Resolve conflicts between the rules generated from the product of the membership degree of each rule;
- Create a combined fuzzy rule base;
- Determine a mapping based using a defuzzification processes;

This work will follow until the third step considering the generation of the database and the rule base. For a given set of input and output data pairs:

A. *Step1 - Divide the input and output spaces into fuzzy regions.*

Considering the domain intervals of the input and output variables, divide each one in to a number of regions. A kind of membership function and a name to each fuzzy set has to be chosen.

B. *Step 2 – Generate fuzzy rules*

For each data pairs determine the membership degrees in each different region and assign to the data the region with maximum degree. The rule is obtained by the input and output regions assigned.

C. *Step 3 – Assign a degree to each rule*

A degree is assigned to each rule to resolve conflict between rules that have the same antecedent part. The degree is calculate by the product of the membership degree of the input data and the output data. The rule considered for the final rule base is the one with higher degree.

IV. GENETIC TUNING

The aim of a tuning method is to find a better set of parameters changing the database components [9].

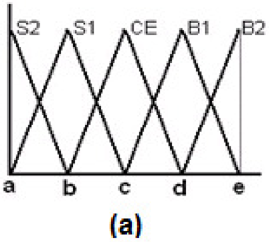
Basic tuning changes the membership functions of an already defined database by fitting the parameters defining their shapes [10]. Genetic tuning uses Genetic Algorithm to manipulate and evaluate the fuzzy sets of a previously defined database of a FRBS.

A genetic tuning considers a chromosome with genes representing all the parameters involved. [6]

In this work we propose to improve the membership function of a database generated by the WM's method.

The real coding scheme [11] is the representation used to codification the chromosomes. The chromosome is encoded by the definition points of each triangular set that composes the database.

The tuning approach could loss the interpretability of the sets involved. Therefore, it is important to apply constraints. Fig. 2 show the limits used in this work to preserve the interpretability.



(a)

$$a \in \left[a - \frac{b-a}{2},\ a + \frac{b-a}{2}\right]$$

$$b \in \left[b - \frac{b-a}{2},\ b + \frac{c-b}{2}\right]$$

$$c \in \left[c - \frac{c-b}{2},\ c + \frac{d-c}{2}\right]$$

$$d \in \left[d - \frac{d-c}{2},\ d + \frac{e-d}{2}\right]$$

$$e \in \left[e - \frac{e-d}{2},\ e + \frac{e-d}{2}\right]$$

(b)

Figure 2. (a)Fuzzy sets. (b) Interpretability constraints.

V. EXPERIMENTS AND RESULTS

In order to illustrate the results obtained by this method, the Auto MPG, Boston Housing and Forest Fires data sets will be considered. The data set were obtained from the Machine Learning Repository from the University of California, Irvine (UCI). Only continuous attributes were used. Table I shows the attributes used for each data set.

TABLE I. DATA SETS USED FOR TESTS

Data Set	Attributes		Number of Instances
	Input	*Output*	
AutoMPG	Displacement	MPG	392
	Horsepower		
	Weight		
	Acceleration		
Boston Housing	CRIM	MEDV	506
	INDUS		
	NOX		
	RM		
	AGE		
	DIS		
	TAX		
	PTRATIO		
	B		
	LSTAT		
Forest Fires	FFMC	Area	517
	DMC		
	DC		
	ISI		
	Temperature		
	RH		
	Wind		

When the WM's method is applied the fuzzy sets and fuzzy rules are generated. We used triangular shaped because of its simplicity with similar results with other shapes [6]. The fuzzy sets generated for the AutoMPG data set are shown in Fig. 3.

The fuzzy rule base for AutoMPG has 147 rules. The Boston Housing rule base has 307 and Forest Fires rule base has 219 rules.

The tests were performed using a standard GA with the 20 individuals in the initial population randomly generated according the interpretability limits shown in Fig. 2. Crossover and Mutation rates are 70% and 5%, respectively. In this work elitism were also implemented with 5% rate. The maximum number of iterations was fixed in 300, but after 100 iterations the values didn't change.

For running the tests, we used the half of number of instances of each data set. The lines were selected randomly. In each line was used the Mamdani inference method, the WM's fuzzy rule base and the fuzzy sets contained in the chromosomes.

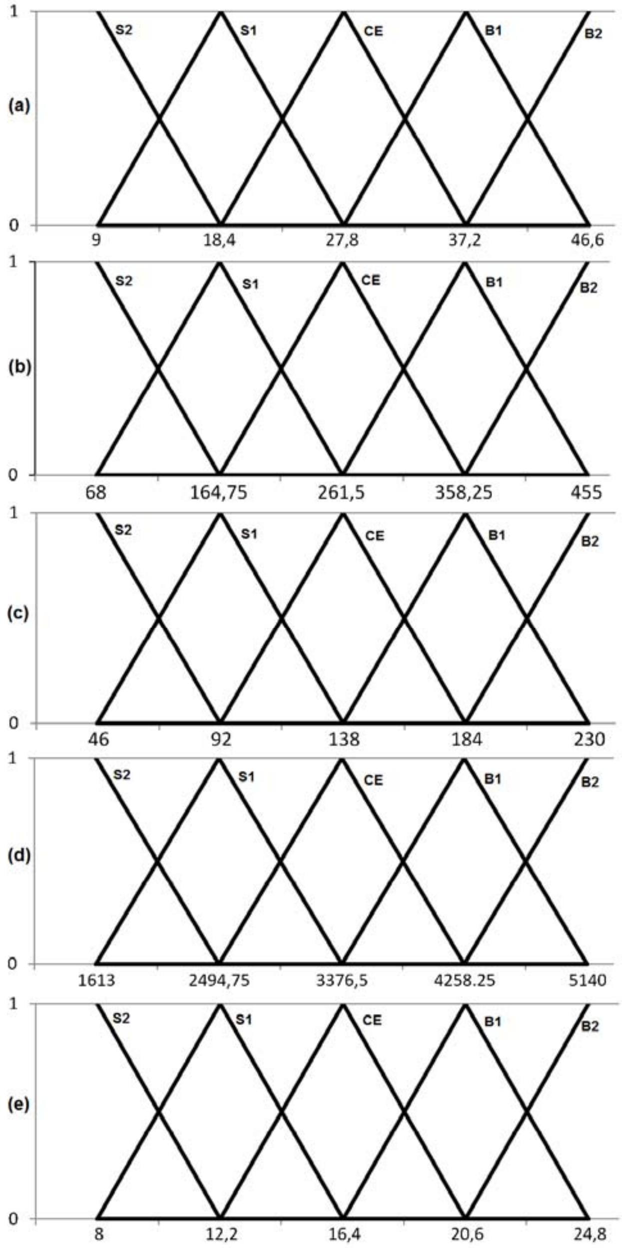


Figure 3. AutoMPG fuzzy sets generate by the WM's method. (a) MPG (b) Displacement (c) Horsepower, (d) Weight and (e) Acceleration.

The defuzzification method generates the output and this output is compared with the original output of the data line.

Following the genetic algorithm defined in Fig. 1, the initial population of 20 individuals is evaluated. The defuzzification method of *Mean of Maxima Weighted* [12], given by (1), where n is the number of rules fired, VM is the maximum value of all output of the rule R_k and G is the firing degree of the rule R_k.:

$$f_0 = \sum_{i=0}^{n} \frac{VM\ (R_k) * G(R_k)}{G(R_k)} \qquad (1)$$

The metric used for evaluation the chromosome was the *mean absolute error* (*MAE*), applied only to the instances that activate rules of the Rule Base. The *MAE* is an average of the absolute errors, where f_i is the prediction calculated by (1) and y_i the true value and is given by (2).

$$MAE = \frac{1}{n} \sum_{i=0}^{n} |f_i - y_i| \qquad (2)$$

According to (2), the lower the value of fitness, the closer the chromosome is of the expected output, and thus is better adapted.

The main cycle of the genetic algorithm is the crossover and mutation operations. The one-point crossover was chosen, which divides each parent chromosome in a half. For each pair of chromosomes a pair of attribute is switched. In AutoMPG data set, the attributes *Displacement* and *Horsepower* of the parent chromosome 1 is recombined with the attributes representing the *Weight* and *Acceleration* of the parent chromosome 2, and vice-versa, creating two children. And this is done for 70% of the population. The type of mutation used was non-uniform.

Individuals are evaluated again after the crossover and mutation operations. The elitism is performed and the best individual goes to the next generation. The stop condition is checked, this being the number of iterations. If the condition is not reached, the algorithm enters into iteration, selecting the new population and applying again the operators.

With the development of the method of Genetic Tuning, it was possible to obtain an improved chromosome. The values of the original chromosome and the tuned chromosome are shown in Table II.

TABLE II. ORIGINAL AND TUNED CHROMOSSOME VALUES

Data Set	Original	Tuned
AutoMPG	3.695473	2.377778
Boston Housing	5.2216716	2.2999992
Forest Fires	11.699314	5.1144443

The fuzzy sets of the AutoMPG data set after the genetic tuning step are shown in Fig.4. The chromosome generated by WM's method and the tuned chromosome are shown respectively, in Table III. and Table IV for the AutoMPG data set, in Table V and Table VI for the Boston Housing and Table VII and Table VIII for the Forest Fires.

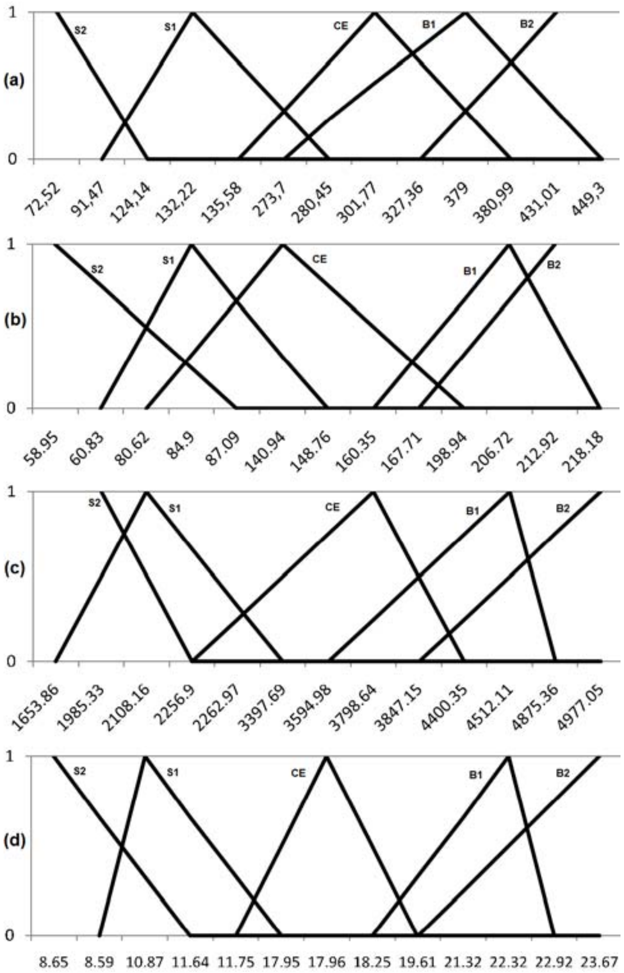


Figure 4. AutoMPG fuzzy sets after the tuning step. (a) Displacement (b) Horsepower, (c) Weight and (d) Acceleration

VI. CONCLUSIONS

In this work we have adopt genetic tuning to improve the WM's database. Our results have shown that genetic tuning improves the database and short the distances between the original output of the data set and the Wang and Mendel's output for fuzzy systems which the rule base is generated by example.

Thus, we can conclude that the tuning step is necessary to obtain better results when the WM's method is used to generate fuzzy rules and fuzzy sets from numerical data because improve the accuracy of the fuzzy system involved.

TABLE III. ORIGINAL CHROMOSOME FOR AUTOMPG DATA SET.

	S2			S1			CE			B1			B2		
x1	68	68	164.75	68	164.75	261.5	164.75	261.5	358.25	261.5	358.25	455	358.25	455	455
x2	46	46	92	46	92	138	92	138	184	138	184	230	184	230	230
x3	1613	1613	2494.75	1613	2494.75	3376.5	2494.75	3376.5	4258.25	3376.5	4258.25	5140	4258.25	5140	5140
x4	8	8	12.2	8	12.2	16.4	12.2	16.4	20.6	16.4	20.6	24.8	20.6	24.8	24.8

TABLE IV. CHROMOSOME GENERATED BY THE GENETIC TUNING FOR AUTOMPG DATA SET.

	S2			S1			CE			B1			B2		
x1	72.52	72.52	124.14	91.47	132.22	280.45	135.58	301.77	380.99	273.7	379.0	449.3	327.36	431.01	431.01
x2	58.95	58.95	87.09	60.83	84.9	148.76	80.62	140.94	198.94	160.35	206.72	218.18	167.71	212.92	212.92
x3	1985.33	1985.33	2262.97	1653.86	2108.16	3397.69	2256.9	3798.64	4400.35	3594.98	4512.11	4875.36	3847.15	4977.05	4977.05
x4	8.65	8.65	11.64	8.59	10.87	17.95	11.75	17.96	21.32	18.25	22.32	22.92	19.61	23.67	23.67

TABLE V. ORIGINAL CHROMOSOME FOR BOSTON HOUSING DATA SET.

	S2			S1			CE			B1			B2		
x1	0.006	0.006	22.249	0.006	22.249	44.491	22.249	44.491	66.734	44.491	66.734	88.976	66.734	88.976	88.976
x2	0.46	0.46	7.28	0.46	7.28	14.1	7.28	14.1	20.92	14.1	20.92	27.74	20.92	27.74	27.74
x3	0.385	0.385	0.507	0.385	0.507	0.628	0.507	0.628	0.75	0.628	0.75	0.871	0.75	0.871	0.871
x4	3.561	3.561	4.866	3.561	4.866	6.171	4.866	6.171	7.475	6.171	7.475	8.78	7.475	8.78	8.78
x5	2.9	2.9	27.175	2.9	27.175	51.45	27.175	51.45	75.725	51.45	75.725	100.0	75.725	100.0	100.0
x6	1.13	1.13	3.879	1.13	3.879	6.628	3.879	6.628	9.377	6.628	9.377	12.127	9.377	12.127	12.127
x7	187	187	318	187	318	449	318	449	580	449	580	711	580	711	711
x8	12.6	12.6	14.95	12.6	14.95	17.3	14.95	17.3	19.65	17.3	19.65	22.0	19.65	22.0	22.0
x9	0.32	0.32	99.465	0.32	99.465	198.61	99.465	198.61	297.755	198.61	297.755	396.9	297.755	396.9	396.9
x10	1.73	1.73	10.79	1.73	10.79	19.85	10.79	19.85	28.91	19.85	28.91	37.97	28.91	37.97	37.97

TABLE VI. CHROMOSOME GENERATED BY THE GENETIC TUNING FOR BOSTON HOUSING DATA SET.

	S2			S1			CE			B1			B2		
x1	3.528	3.528	26.513	3.274	25.307	44.999	25.314	45.627	69.698	47.497	70.453	89.739	70.895	87.016	87.016
x2	2.368	2.368	9.539	2.86	8.163	14.549	7.674	14.73	21.127	14.664	22.601	27.981	21.703	25.492	25.492
x3	0.966	0.966	0.871	0.461	0.73	0.901	0.708	0.7	0.952	1.089	0.768	1.336	0.758	0.518	0.518
x4	3.571	3.571	5.387	4.047	5.311	6.404	6.086	7.315	8.305	7.326	7.505	9.667	8.015	7.768	7.768
x5	4.8	4.8	28.741	4.261	31.777	52.866	28.284	54.328	76.465	55.931	77.133	104.543	80.127	95.325	95.325
x6	1.483	1.483	4.851	1.43	5.061	6.899	4.776	7.339	10.671	7.865	10.422	12.221	9.677	11.882	11.882
x7	207.033	207.033	329.259	203.716	325.808	458.581	333.276	459.286	591.102	450.831	593.733	720.322	584.257	695.22	695.22
x8	12.824	12.824	16.082	13.73	15.305	18.288	15.758	17.529	20.692	17.934	20.346	22.689	20.177	21.055	21.055
x9	6.758	6.758	120.033	16.884	116.144	206.986	114.857	203.164	298.827	199.981	300.788	400.073	316.538	382.239	382.239
x10	1.968	1.968	11.513	3.589	13.088	21.373	12.601	21.76	30.402	22.195	31.125	39.451	28.994	37.036	37.036

TABLE VII. ORIGINAL CHROMOSOME FOR FOREST FIRES DATA SET.

	S2			S1			CE			B1			B2		
x1	18.7	18.7	38.08	18.7	38.08	57.45	38.08	57.45	76.82	57.45	76.82	96.2	76.82	96.2	96.2
x2	1.1	1.1	73.65	1.1	73.65	146.2	73.65	146.2	218.75	146.2	218.75	291.3	218.75	291.3	291.3
x3	7.9	7.9	221.07	7.9	221.07	434.25	221.07	434.25	647.43	434.25	647.43	860.6	647.43	860.6	860.6
x4	0.0	0.0	14.03	0.0	14.03	28.05	14.03	28.05	42.07	28.05	42.07	56.1	42.07	56.1	56.1
x5	2.2	2.2	9.97	2.2	9.97	17.75	9.97	17.75	25.53	17.75	25.53	33.3	25.53	33.3	33.3
x6	15.0	15.0	36.25	15.0	36.25	57.5	36.25	57.5	78.75	57.5	78.75	100.0	78.75	100.0	100.0
x7	0.4	0.4	2.65	0.4	2.65	4.9	2.65	4.9	7.15	4.9	7.15	9.4	7.15	9.4	9.4

TABLE VIII. CHROMOSOME GENERATED BY THE GENETIC TUNING FOR FOREST FIRES DATA SET.

	S2			S1			CE			B1			B2		
x1	19.93	19.93	34.26	13.19	40.17	54.07	44.41	56.95	74.0	53.03	72.58	96.42	78.99	90.02	90.02
x2	32.91	32.91	47.04	-21.7	96.6	116.16	87.12	112.67	187.93	139.81	201.94	293.42	230.44	255.76	255.76
x3	46.62	46.62	126.72	-51.48	263.96	401.93	279.8	402.08	629.77	430.58	552.54	915.26	676.89	794.06	794.06
x4	0.9	0.9	7.49	-2.69	19.49	23.27	16.65	21.84	35.37	21.73	40.64	59.04	46.42	49.48	49.48
x5	4.13	4.13	6.68	1.39	13.03	17.14	11.44	14.63	22.8	16.6	24.63	33.46	28.5	30.31	30.31
x6	19.41	19.41	33.58	8.07	43.98	55.93	40.15	51.56	71.19	53.84	77.13	108.76	79.42	99.76	99.76
x7	0.78	0.78	2.44	0.09	3.18	4.72	3.7	4.88	6.25	4.61	7.12	10.32	8.1	8.52	8.52

REFERENCES

[1] R. R. Yager, S. Ovchinnikov, R.M. Tong, H. T. Nguyen, “Fuzzy Sets and Applications: Selected Papers by LA Zadeh”, Wiley, New York, 1987.

[2] D. Nauck, R. Kruse, “Neuro-fuzzy systems for function approximation, Fuzzy Sets and Systems”, 101, 261–271, 1999.

[3] O. Cordón, F. Herrera, “A three-stage evolutionary process for learning descriptive and approximate fuzzy logic controller knowledge bases from examples”, Int. Journal of Approximate Reasoning 17:4, 369–407, 1997.

[4] F. Herrera, M. Lozano, J. L. Verdegay, “Generating Fuzzy Rules from Examples Using Genetic Algorithms”, Fifth International Conference of Information Processing and Management of Uncertainty in Knowledge-Based Systems, p. 675-680, Paris, France, 1994.

[5] L. X. Wang, J. M. Mendel, “Generating fuzzy rules by learning from examples”, IEEE Transactions on Systems, Man and Cybernetics, vol.22, no.6, p.1414-1427, 1992.

[6] R. Alcalá, J. Alcalá-Fdez, F. Herrera, “A Proposal for the Genetic Lateral Tuning of Linguistic Fuzzy Systems and Its Interaction With Rule Selection”, vol. 15, no.4, 616-635, 2007.

[7] J. Casillas, O. Cordón, F. Herrera, “Improving the Wang and Mendel's fuzzy rule learning method by inducing cooperation among rules”, Proceedings of the 8th Information Processing and Management of Uncertainty in Knowledge-Based Systems Conference, Madrid, Spain, 1682-1688, 2000.

[8] J. Casillas, O. Cordón, F. Herrera, “COR: a methodology to improve ad hoc data-driven linguistic rulelearning methods by inducing cooperation among rules”, IEEE Transactions On Systems, Man, And Cybernetics—Part B: Cybernetics, vol.32, no.4, 526-537, 2002.

[9] R. Alcalá, J. Casillas, J. L. Castro, A. González, F. Herrera, “A Multicriteria Genetic Tuning for Fuzzy Logic Controllers”, Mathware & Soft Computing 8, Universitat Politècnica de Catalunya (UPC), Catalunya, Spain, 2001.

[10] J. Casillas, O. Cordón, M.J. Del Jesus, F. Herrera, “Genetic tuning of fuzzy rule-based systems integrating linguistic hedges”. Proceedings of the 9th IFSA World Congress and 20th NAFIPS International Conference, Vancouver (Canada), 1570-1574, 2001.

[11] F. Herrera, M. Lozano, J. L. Verdegay, “Tackling realo-codec genetic algorithms: Operators and tools for the behavior analysis”, Artificial Intelligence Reviwe 12, 265-319, 1998.

[12] G. J. Klir, B. Yuan, “Fuzzy sets and fuzzy logic: theory and applications”. Upper Saddle River, NJ: Prentice Hall,1995.