A New Genetic Fuzzy System Based on Linguistic 2-Tuples to Learn Knowledge Bases

Rafael Alcalá and Jesús Alcalá-Fdez and Francisco Herrera
Department of Computer Science and Artificial Intelligence
University of Granada
18071 Granada, Spain
E-mail: {alcala, jalcala, herrera}@decsai.ugr.es

José Otero
Department of Computer Science
University of Oviedo
Campus de Viesques, 33203 Gijón, Spain
E-mail: jotero@lsi.uniovi.es

***Abstract*—Fuzzy systems demonstrated their ability to solve different kinds of problems in various application domains. Currently, there is an increasing interest to augment fuzzy systems with learning and adaptation capabilities. Two of the most successful approaches to hybridise fuzzy systems with learning and adaptation methods have been made in the realm of Soft Computing (Genetic and Neuro-Fuzzy Systems). One of them, Genetic Fuzzy Systems, are basically fuzzy systems augmented by learning processes based on Genetic Algorithms. However, in some cases, the large search space handled by such techniques provokes the derivation of sub-optimal models.**

Recently, the linguistic 2-tuples rule representation model was presented to allow the lateral displacement of a label considering an unique parameter. It involves a reduction of the search space that eases the derivation of optimal models. In this work, we propose the use of a new Genetic Fuzzy System based on these kinds of rules, to obtain linguistic fuzzy systems by means of an evolutionary learning of the data base *a priori* (granularity and lateral displacements) and the use of a basic rule generation method to learn the whole knowledge base. We analyze this approach considering a real-world problem.

I. Introduction

Genetic Fuzzy Systems (GFSs) [1], [2] are basically fuzzy systems augmented by learning processes based on Genetic Algorithms (GAs). GAs are search algorithms, based on natural genetics, that provide robust search capabilities in complex spaces, and thereby offer a valid approach to problems requiring efficient and effective search processes [3].

A literature analysis shows that the most prominent types of GFSs are *Genetic Fuzzy Rule-Based Systems* (GFRBSs) [1], whose genetic process learns or tunes different components of a Fuzzy Rule-Based System (FRBS): the data base (DB) —containing the parameters of the linguistic partitions—, the rule base (RB) —containing the set of rules— and the whole knowledge base (KB) —containing the RB and the DB—.

Inside GFRBSs it is possible to distinguish between either methods considering the *learning of the RB and the DB*, or *post-processing mechanisms* that are applied to improve the system behavior once the RB and the DB are obtained. However, in some problems, such techniques handle a large search space which could provoke the derivation of sub-optimal models, specially when they learn the whole KB.

To solve this problem, a particular case of post-processing was proposed in [4], where a new rule representation model was presented for the genetic tuning of the DB. This approach is based on the linguistic 2-tuples representation [5], which allows the lateral displacement of the labels considering an unique parameter per label. In this way, such labels maintain their original shapes, and the search space is reduced respect to the classical tuning in order to easily obtain optimal models.

Such learning scheme starts from an initial DB (with a fixed granularity) and an initial RB that remains fixed during all the tuning process. However, it would be desirable a greater degree of cooperation between these two tasks (RB and DB learning) in order to obtain models with a better accuracy-interpretability trade-off.

With this aim, we propose a new GFRBS to obtain whole KBs based on the learning of the granularity and the linguistic 2-tuples rule representation model, which at the same time learns the optimal number of labels per variable, the lateral displacement of such labels and, from this, by means of a simple rule generation method, obtains the corresponding RB. As an example, we will analyze this technique by solving a real world problem from both, the accuracy point of view and the interpretability point of view.

This contribution is arranged as follows. The next section presents a short review of GFRBSs in order to introduce the learning scheme considered in this work. Section III describes the linguistic rule representation model based on the linguistic 2-tuples and proposes the new evolutionary learning method. Section IV shows an experimental study of the method behavior applied over a real-world estimation problem. Finally, Section V points out some concluding remarks.

II. Genetic Fuzzy Rule-Based Systems: Short Introduction

This section presents a short introduction to GFRBSs focusing on the classical systems to perform parameter optimization or rule generation and then, a new (or less common) research line is presented in order to introduce the learning scheme considered in this work. An extended description of GFSs can be found in [1], [2].

A number of papers employ an evolutionary learning process to automate the design of the KB. From the viewpoint of optimization, the KB parameters constitute the optimization space, which is transformed into a suitable genetic representation on which the search process operates.

The first step in designing a GFRBS is to decide which parts of the KB are subject to optimization by the GA. The decision on which part of the KB to adapt depends on two conflicting objectives: dimensionality and efficiency of the search. Therefore, there is an obvious trade-off between the completeness and dimensionality of the search space and the efficiency of the search. This trade-off offers different possibilities for GFS design that are considered in the following subsections.

A. Genetic tuning of data bases

Tuning of the scaling functions and fuzzy membership functions is an important task in FRBS design. Parameterized scaling functions and membership functions are adapted by the GA according to a fitness function that specifies the design criteria in a quantitative manner. Although it is also possible an *a priori genetic DB learning*, tuning processes usually assume a predefined RB and have the objective of finding a set of optimal parameters for the membership [6], [7] and/or the scaling [6] functions.

B. Genetic learning of rule bases

Genetic learning of RBs is only applied to linguistic FRBSs, assuming a predefined set of membership functions in the DB to which the rules refer to by means of linguistic labels. However, it is also possible an *a posteriori DB learning* (tuning).

Three different learning approaches can be considered to learn RBs: Michigan approach [8], Pittsburgh approach [9], and iterative rule learning approach [7]. The first one is characterized by representing an entire rule set as a chromosome. The Michigan approach considers each chromosome is a individual rule and a rule set is represented by the entire population. In the third approach, the iterative one, each chromosome codes individual rules, and a new rule is adapted and added to the rule set, in an iterative fashion, in every run of the GA.

C. Genetic learning of knowledge bases

Since genetic learning of the KB deals with heterogeneous search spaces, it encompasses different genetic representations such as variable length chromosomes, multi-chromosome genomes and chromosomes encoding single rules instead of a whole KB. The computational cost of the genetic search grows with the increasing complexity of the search space. A GFRBS that encodes individual rules rather than entire KBs is an option to maintain a flexible, complex rule space in which the search for a solution remains feasible and efficient. Again, the three learning approaches can considered: Michigan [10], Pittsburgh [11], and iterative rule learning approach [7].

D. New trends to learn KBs

In addition to the classical approaches, some new directions to apply genetic (evolutionary) techniques to FRBSs can be found in [2]: selection of fuzzy rules, feature selection, learning of KBs via genetic derivation of the DB, interpretability maintenance via multi-objective genetic processes, learning approaches considering different model structures, etc. These kinds of techniques try to find a better tradeoff between the efficiency of the learning/post-processing process and the dimensionality of the search space.

It is the case of the *learning of KBs via genetic derivation of DBs*, a recent approach involving a simpler search space than the classical learning of KBs. It consists of obtaining the DB and the RB separately, based on the DB learning *a priori* [12]–[15] (see Figure 1). This way to work allows us to learn the most adequate context for each fuzzy partition, which is necessary in different application contexts and different fuzzy rule extraction models.

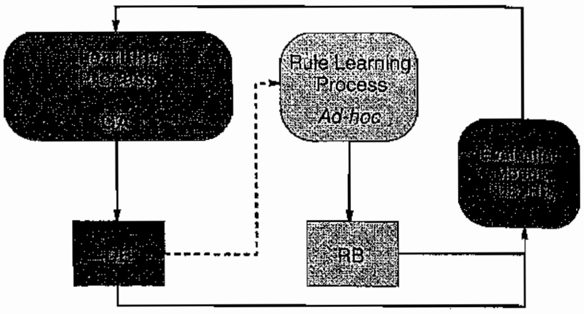


Fig. 1. Learning scheme of the KB

The learning scheme considered in this work belongs to this last group and is comprised of two main components:

- A process to learn the DB, which allows to define:
  - The number of labels for each linguistic variable.
  - The lateral displacements of such labels.

  Triangular membership functions are considered for their simplicity.
- A quick *Ad-hoc data-driven method* to derive the RB [16] considering the DB previously obtained. This method is run from each DB definition generated by a Genetic Algorithm (GA), thus, allowing the proposed hybrid learning process to finally obtain the whole definition of the KB (DB and RB) by means of the cooperative action of both methods.

III. GFRBS FOR LEARNING OF THE KB

This section presents the linguistic 2-tuples rule representation model and the evolutionary algorithm to obtain KBs based on the DB learning *a priori* and this new rule representation.

A. The Linguistic 2-Tuples Representation

In [4], a new model of tuning of FRBSs was proposed considering the linguistic 2-tuples representation scheme introduced in [5], which allows the lateral displacement of the support of a label and maintains the interpretability associated to the obtained linguistic FRBSs. This proposal also introduces a new model for rule representation based on the symbolic translation concept that is a number within the interval [-0.5, 0.5) expressing the domain of a label when it is moving between its two lateral labels. Formally, we have the pair,

$$(s_i, \alpha_i), \ \ s_i \in S, \ \ \alpha_i \in [0.5, -0.5).$$

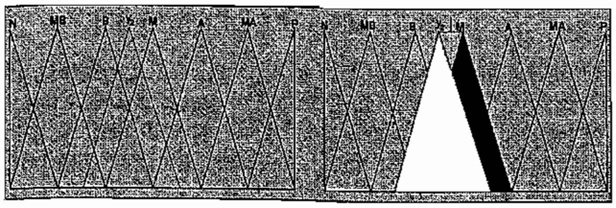


Fig. 2. Lateral Displacement of the Linguistic Label M

Figure 2 shows the lateral displacement of the label M. The new label "y_2" is located between B and M, being enough smaller than M but closer to M.

In the next, we present this approach considering a simple control problem. Let us consider a control problem with two input variables, one output variable and a DB defined from experts determining the membership functions for the following labels:

$$Error, \ \nabla Error \rightarrow \{N, Z, P\}, \ Power \rightarrow \{L, M, H\} \ .$$

Classical Rule:
R1: If the **Error** is Zero and the **Error Variation** is Positive then the **Power** is High

Rules with 2-tuples Representation:
R1: If the **Error** is (Zero, 0.3) and the **Error Variation** is (Positive, -0.2) then the **Power** is (High, -0.1)

Fig. 3. Classical Rule and Rule with 2-Tuple Representation

Figure 3 shows the concept of classical rule and linguistic 2-tuples represented rule. Analyzed from the rule interpretability point of view, we could interpret the obtained rule as:

> If the **Error** is "higher than Zero" and
> the **Error Variation** is "a little smaller than Positive"
> then the **Power** is "a bit smaller than High".

In [4], two different rule representation approaches were proposed, a global approach and a local approach. In our particular case, the learning is applied to the level of linguistic partition (considering the global approach). In this way, the pair (X_i, label) takes the same tuning value in all the rules where it is considered. For example, X_i is (High, 0.3) will present the same value for those rules in which the pair "X_i is High" is initially considered. This proposal decreases the learning problem complexity, since the 3 parameters considered per label are reduced to only 1 symbolic translation parameter.

B. Evolutionary Algorithm

In this work, we will consider the use of CHC [17] to design the proposed learning algorithm. CHC is a GA that presents a good trade-off between exploration and exploitation, being a good choice in problems with complex search spaces. This genetic model makes use of a mechanism of "Selection of Populations". N parents and their corresponding offspring are combined to select the best N individual to take part of the next population.

Considering this approach, the learning process of the DB have to define both, the granularity of the linguistic partitions and the lateral displacements of the involved labels. For this reason, a double coding scheme is considered (granularity + displacements).

In the following, the components needed to design our GFRBS are explained. They are: DB codification, chromosome evaluation, initial gene pool, crossover operator and restarting approach.

1) DB Codification: A double coding scheme ($C = C_1 + C_2$) to represent both parts, *granularity* and *translation parameters*, is considered:

- Number of labels (C_1): This part is a vector of integer numbers with size N (being N the number of system variables). The possible numbers of labels considered are the set $\{3, \ldots, 9\}$:
$$C_1 = (L^1, \ldots, L^N) \ .$$
- Lateral displacements (C_2): This part is a vector of real numbers with size $N * 9$ (N variables with a maximum of 9 linguistic labels per variable) in which the displacements of the different labels are coded for each variable. Of course, if a chromosome does not have the maximum number of labels in one of the variables, the space reserved for the values of these labels is ignored in the evaluation process. In this way, the C_2 part has the following structure (where each gene is associated to the tuning value of the corresponding label):
$$C_2 = (\alpha_1^1, \ldots, \alpha_{L^1}^1, \ldots, \alpha_1^N, \ldots, \alpha_{L^N}^N)$$

2) Chromosome Evaluation: To evaluate a determined chromosome we will apply the well-known rule generation method of Wang and Mendel [16] on the DB coded by such chromosome. Once the whole KB is obtained, the Mean Square Error (MSE) is computed and the following function is minimized:

$$F_C = w_1 \cdot MSE + w_2 \cdot NR,$$

where, NR is the number of rules of the obtained KB (in order to penalize an excessive number of rules), $w_1 = 1$ and w_2 is computed from the MSE and the number of rules of the KB generated from a DB considering the maximum number of labels (9 labels) and without considering the displacement parameters,

$$w_2 = \alpha \cdot \frac{MSE_{max-lab}}{NR_{max-lab}} \ ,$$

with α being a weighting percentage given by the system expert that determines the trade-off between accuracy and complexity. Values higher than 1.0 search for linguistic models with few rules, and values lower than 1.0 search for linguistic models with high accuracy. A good neutral choice is for example 1.0 (good accuracy and not too many rules).

For the fuzzy inference, we have selected the *minimum t-norm* playing the role of the implication and conjunctive operators, and the *center of gravity weighted by the matching* strategy acting as defuzzification operator.

3) Initial Gene Pool: The initial population will be comprised of two different parts (with the same number of chromosomes):

- In the first part, each chromosome has the same number of labels for all the problem variables and considers strong fuzzy partitions with translation parameters initialized to zero.
- In the second part, the only change is that each variable could have a different number of labels.

Since CHC has no mutation operator, the translation parameters remain unchanged and the most promising number of labels is obtained for each linguistic variable. The algorithm operates in this way until the first restarting is reached.

4) Crossover Operator: Two different crossover operators are considered depending on the two parent's scope to obtain two offspring:

- *When the parents encode different granularity levels in any variable*, a crossover point is randomly generated in C_1 and the classical crossover operator is applied on this point in both parts, C_1 and C_2 (exploration).
- *When both parents have the same granularity level per variable*, an operator based on the the concept of environments (the offspring are generated around one parent) is applied only on the C_2 part (exploitation). These kinds of operators present a good cooperation when they are introduced within evolutionary models forcing the convergence by pressure on the offspring (as the case of CHC). Particularly, we consider the Parent Centric BLX (PCBLX) operator [18], which is based on the BLX-α. The 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 generates the offspring $Z = (z_1 \cdots z_n)$, where z_i is a randomly (uniformly) chosen number from the interval $[l_i, u_i]$, with $l_i = max\{a_i, x_i - I\}$, $u_i = min\{b_i, x_i + I\}$, and $I = \mid x_i - y_i \mid$. The parents X and Y will be named differently: X will be called *female parent*, and Y will be called *male parent*. In this way, by taking X as female parent (Y as male), and then by taking Y as female parent (X as male) our algorithm generates two offspring.

  On the other hand, CHC makes use of an incest prevention mechanism, i.e., two parents are only crossed if their hamming distance divided by 2 is over a predetermined threshold, L. It will be only considered in order to apply the PCBLX operator. Since, we consider a real coding scheme (the C_2 part is going to be crossed), we have to transform each gene considering a Gray Code with a fixed number of bits per gene ($BITSGENE$) determined by the system expert. In this way, the threshold value is initialized as:

$$L = (\#GenesC_2 * BITSGENE)/4.0.$$

Following the original CHC scheme, L is decremented by one when no cross is performed in one generation. In order to avoid very slow convergence, in our case, L will be also decremented by one when no improvement is achieved respect to best chromosome of the previous generation.

5) Restarting approach: To get away from local optima, a restarting mechanism is considered [17] when the threshold value L is lower than cero. In this case, all the chromosomes set up their C_1 parts to that of the best global solution, being the parameters of their C_2 parts generated at random within the interval $[-0.5, 0.5)$. Moreover, if the best global solution had any change from the last restarting point, this is included in the population (the exploitation continues while there is convergency). This operation mode was initially proposed by the CHC authors as a possibility to improve the algorithm performance when it is applied to solve some kinds of problems [17].

IV. Experiments and Analysis of Results

To analyze the behavior of the proposed method, learning of the granularity together with the global lateral displacements, several experiments have been carried out considering a real-world problem, the estimation of the maintenance costs of the medium voltage electrical network in a town [19]. This problem handles four input variables and therefore, it involves *a large search space*. A short description of this problem can be found in the following subsection.

TABLE I
METHODS CONSIDERED FOR THE EXPERIMENTAL STUDY

Ref.	Method	Type of Learning
[16]	WM	*Ad-hoc* Data-Driven Method
[20]	COR	Cooperative Rules (Best-Worst Ant System)
[13]	GA+WM	Granularity + Scaling Factors + Contexts + RB by DB learning *a priori* and Using WM
[12]	GA+COR	As GA+WM but using COR
—	GLD+WM	Granularity + Global Lateral Displacements + RB by DB learning *a priori* and Using WM

Table I presents a brief description of the studied methods. The WM and COR algorithms are used as simple rule generation methods to obtain RBs from a predefined DB. Two methods to obtain complete KBs are considered for comparisons. They are based on the DB learning *a priori* obtaining the granularity, scaling factors and contexts (i.e., the variable domain or working range to perform the fuzzy partitioning). The proposed method (GLD+WM) and the GA+WM method integrate the WM algorithm within its own DB learning process as the mechanism to obtain the RB. GA+COR integrates the best-worst ant system-based COR algorithm to perform this task.

In the case of WM and COR, the initial linguistic partitions are comprised by *five linguistic terms* with triangular-shaped fuzzy sets giving meaning to them (number of labels by which they presented the best behavior). Finally, the following values have been considered for the parameters of each method [1]: 50 individuals and 50,000 evaluations; 0.6 and 0.2 as crossover and mutation probabilities in the case of COR, GA+COR and GA+WM. The α factor for the fitness function of GLD+WM was set to 1, 3 and 5, in order to obtain models with different levels of accuracy and simplicity. The number of bits for the Gray codification is 30 bits per gene.

A. Problem Description: Estimating the Maintenance Costs of Medium Voltage Lines

Estimating the maintenance costs of the medium voltage electrical network in a town [19] is a complex but interesting problem. Since a direct measure is very difficult to obtain, the consideration of models becomes useful. These estimations allow electrical companies to justify their expenses. Moreover, the model must be able to explain how a specific value is computed for a certain town. Our objective will be to relate the *maintenance costs of medium voltage line* with the following four variables: *sum of the lengths of all streets in the town, total area of the town, area that is occupied by buildings,* and *energy supply to the town.* We will deal with estimations of minimum maintenance costs based on a model of the optimal electrical network for a town in a sample of 1,059 towns.

To develop the different experiments in this contribution, 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. In this way, 5 partitions considering an 80% (847) in training and a 20% (212) in test are considered for the experiments.

B. Results and Analysis

For each one of the 5 data partitions, the tuning methods has been run 6 times, showing for each problem the averaged results of a total of 30 runs. Moreover, a *t-test* (with 95 percent confidence) was applied to the best averaged result in training or test by comparing one by one this result to the averaged results of the remaining methods.

The results obtained by the analyzed methods are shown in Table II, where $\#R$ stands for the number of rules, MSE_{tra} and MSE_{tst} respectively for the averaged error obtained over the training and test data, σ for the standard deviation and *t-test* represents the following information:

$\star$	Denotes the best averaged result
$-$	Denotes a significant worst behavior than the best

TABLE II

RESULTS OBTAINED BY THE STUDIED METHODS

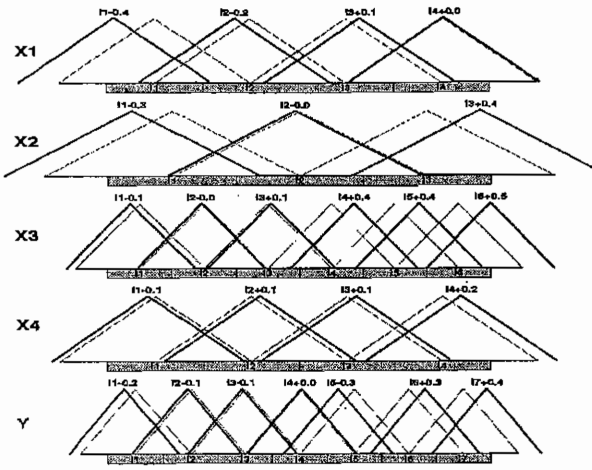
Method	#R	MSE_{tra}	σ_{tra}	t-test	MSE_{tst}	σ_{tst}	t-test
WM	65	56136	1498	$-$	56359	4686	$-$
COR	41	39640	566	$-$	41683	1599	$-$
GA+WM	51.1	23014	2143	$-$	24090	3667	$-$
GA+COR	**17.8**	20360	1561	$-$	22830	3259	$-$
GLD+WM, with $\alpha = 1$	57.5	**10218**	1044	$\star$	**12088**	1972	$\star$
GLD+WM, with $\alpha = 3$	41.2	13074	2040	$-$	15196	2757	$-$
GLD+WM, with $\alpha = 5$	30.7	16884	2822	$-$	18943	3649	$-$

Analyzing the results presented in Table II we can point out the following conclusions:

- Although the proposed method could learn partitions presenting a high granularity, the RB is obtained by means of a method to learn few rules (57.5, 41.1 or 30.7 from the 6561 possible rules if the input partitions consider nine labels, 9^4) which, together with the objective to minimize the number of rules, allow us to obtain accurate but compact models and, therefore, more interpretable. Notice that, the GA+COR method obtains the linguistic models with less number of rules. It is due to the rule simplification performed by COR during the RB learning, which results in linguistic models with too few rules.
- The method proposed in this work shows an important reduction of the mean squared error over the training and test data in a problem with a large search space, being robust to random factors. We can see that the proposed algorithm does not present significant deviations in the results, being its standard deviation one of the lowest.
- The consideration of an unique parameter per membership function reduces the search space respect to the classical learning of membership functions. Therefore, this learning approach represents the ideal framework to be combined with other learning schemes (hierarchical fuzzy rule learning, consideration of COR [20] to obtain the RB, etc). Furthermore, the linguistic models so obtained are highly interpretable since the original shapes of the initial membership functions are maintained.

Figure 4 depicts the evolved fuzzy linguistic partitions and the RB obtained by the GLD+WM method from one of the 30 runs performed with $\alpha = 5$. To easy the graphical representation, in these figures, the labels are named from '11' to '$1L^i$'. Nevertheless, such labels have associated a linguistic meaning determined by an expert. In this way, if the '11' label of the 'X1' variable represents 'LOW', '11+0.11' could be interpreted as 'a little smaller than LOW' (based on the expert opinion) or, as in the case of some classical learning approaches, maintaining the original meaning of such label. It

[1] With these values we have tried to ease the comparisons selecting standard common parameters that work well in most cases instead of searching very specific values for each method. Moreover, we have set a large number of evaluations in order to allow the compared algorithms to achieve an appropriate convergence. No significant changes were achieved by increasing that number of evaluations.



X1	X2	X3	X4	Y	X1	X2	X3	X4	Y
l1 -0.41	l1 -0.31	l1 -0.14	l1 -0.08	l1 -0.21	l3+0.15	l2 -0.04	l4+0.36	l2+0.10	l5 -0.30
l1 -0.41	l1 -0.31	l2 -0.04	l1 -0.08	l2 -0.06	l3+0.15	l2 -0.04	l4+0.36	l3+0.10	l6+0.30
l2 -0.16	l1 -0.31	l1 -0.14	l1 -0.08	l1 -0.21	l3+0.15	l2 -0.04	l5+0.38	l2+0.10	l6+0.30
l2 -0.16	l1 -0.31	l1 -0.14	l2+0.10	l2 -0.06	l3+0.15	l2 -0.04	l5+0.38	l3+0.10	l7+0.43
l2 -0.16	l1 -0.31	l2 -0.04	l1 -0.08	l2 -0.06	l3+0.15	l3+0.42	l4+0.36	l2+0.10	l5 -0.30
l2 -0.16	l1 -0.31	l2 -0.04	l2+0.10	l3 -0.06	l3+0.15	l3+0.42	l4+0.36	l3+0.10	l6+0.30
l2 -0.16	l2 -0.04	l2 -0.04	l1 -0.08	l2 -0.06	l3+0.15	l3+0.42	l6+0.50	l2+0.10	l6+0.30
l2 -0.16	l2 -0.04	l2 -0.04	l2+0.10	l3 -0.06	l3+0.15	l3+0.42	l6+0.50	l3+0.10	l7+0.43
l2 -0.16	l2 -0.04	l3+0.06	l1 -0.08	l3 -0.06	l4+0.02	l2 -0.04	l2 -0.04	l1 -0.08	l3 -0.06
l2 -0.16	l2 -0.04	l3+0.06	l2+0.10	l4+0.04	l4+0.02	l2 -0.04	l2 -0.04	l2+0.10	l3 -0.06
l3+0.15	l2 -0.04	l2 -0.04	l1 -0.08	l2 -0.06	l4+0.02	l2 -0.04	l2 -0.04	l3+0.10	l4+0.04
l3+0.15	l2 -0.04	l2 -0.04	l2+0.10	l3 -0.06	l4+0.02	l2 -0.04	l2 -0.04	l4+0.20	l6+0.30
l3+0.15	l2 -0.04	l2 -0.04	l3+0.10	l4+0.04	l4+0.02	l2 -0.04	l3+0.06	l1 -0.08	l4+0.04
l3+0.15	l2 -0.04	l3+0.06	l1 -0.08	l3 -0.06	l4+0.02	l2 -0.04	l3+0.06	l2+0.10	l4+0.04
l3+0.15	l2 -0.04	l3+0.06	l2+0.10	l4+0.04	l4+0.02	l2 -0.04	l3+0.06	l3+0.10	l5 -0.30
l3+0.15	l2 -0.04	l3+0.06	l3+0.10	l5 -0.30	l4+0.02	l2 -0.04	l3+0.06	l4+0.20	l6+0.30

#R: 32, MSE-tra: 13680.317, MSE-tst: 14857.941

Fig. 4. DB with/without lateral displacements (black/gray), RB and displacements of a model obtained by GLD+WM with $\alpha = 5$

is the case in Figure 4, where the new labels could maintain their initial meanings.

V. Conclusions

This work presents a new GFRBS for learning KBs by means of an *a priori* learning of the DB (granularity and translation parameters). It makes use of the linguistic 2-tuples rule representation model presented in [4]. In the following, we present our conclusions to subsequently present some further considerations:

- The used learning scheme together with the 2-tuples rule representation model allows an important reduction of the search space that eases the derivation of more precise and compact linguistic models.
- Since a global approach is considered and the shapes of the initial linguistic partitions are preserved, the interpretability of the obtained models is maintained to a high level respect to the classical learning of fuzzy systems.

An interesting further question is the accomplishment of a wider analysis about the existent relation between the problem complexity and the behavior of the different GFRBSs presented in this work.

Acknowledgment

This work has been supported by the CICYT Project TIC-2002-04036-C05-01.

References

[1] O. Cordón, F. Herrera, F. Hoffmann, and L. Magdalena, *GENETIC FUZZY SYSTEMS. Evolutionary tuning and learning of fuzzy knowledge bases*, ser. Advances in Fuzzy Systems - Applications and Theory. World Scientific, 2001, vol. 19.

[2] O. Cordón, F. Gomide, F. Herrera, F. Hoffmann, and L. Magdalena, "Ten years of genetic fuzzy systems: Current framework and new trends," *Fuzzy Sets and Systems*, vol. 41, no. 1, pp. 5–31, 2004.

[3] D. E. Goldberg, *The design of competent genetic algorithms: steps toward a computational theory of innovation.* Kluwer Academic Publishers, 2002.

[4] R. Alcalá and F. Herrera, "Genetic tuning on fuzzy systems based on the linguistic 2-tuples representation," in *Proc. of the 2004 IEEE International Conference on Fuzzy Systems*, vol. 1, Budapest (Hungary), 2004, pp. 233–238.

[5] F. Herrera and L. Martínez, "A 2-tuple fuzzy linguistic representation model for computing with words," *IEEE Trans. Fuzzy Syst.*, vol. 8, no. 6, pp. 746–752, 2000.

[6] J. Casillas, O. Cordón, M. del Jesús, and F. Herrera, "Genetic tuning of fuzzy rule deep dtructures preserving interpretability for linguistic modeling," *IEEE Trans. Fuzzy Syst.*, 2004, to be published.

[7] O. Cordón and F. Herrera, "A three-stage evolutionary process for learning descriptive and approximate fuzzy logic controller knowledge bases from examples," *International Journal of Approximate Reasoning*, vol. 17, no. 4, pp. 369–407, 1997.

[8] H. Ishibuchi, T. Nakashima, and T. Murata, "Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems," *IEEE Trans. Syst., Man, Cybern. B*, vol. 29, pp. 601–618, 1999.

[9] F. Hoffmann and G. Pfister, "Evolutionary design of a fuzzy knowledge base for a mobile robot," *International Journal of Approximate Reasoning*, vol. 17, no. 4, pp. 447–469, 1997.

[10] J. R. Velasco, "Genetic-based ou-line learning for fuzzy process control," *International Journal of Intelligent Systems*, vol. 13, no. 10-11, pp. 891–903, 1998.

[11] L. Magdalena and F. Monasterio-Huelin, "A fuzzy logic controller with learning through the evolution of its knowledge base," *International Journal of Approximate Reasoning*, vol. 16, no. 3-4, pp. 335–358, 1997.

[12] J. Casillas, O. Cordón, F. Herrera, and P. Villar, "A hybrid learning process for the knowledge base of a fuzzy rule-based system," in *Proc. of the 2004 International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, vol. 3, Perugia (Italy), 2001, pp. 2189–2196.

[13] O. Cordón, F. Herrera, L. Magdalena, and P. Villar, "A genetic learning process for the scaling factors, granularity and contexts of the fuzzy rule-based system data base," *Information Sciences*, vol. 136, pp. 85–107, 2001.

[14] W. Pedrycz, "Associations and rules in data mining: A link analysis," *International Journal of Intelligent Systems*, vol. 19, no. 7, pp. 653–670, 2004.

[15] Y. W. Teng and W. J. Wang, "Constructing a user-friendly ga-based fuzzy system directly from numerical data," *IEEE Trans. Syst., Man, Cybern. B*, vol. 34, no. 5, pp. 2060–2070, 2004.

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

[17] L. Eshelman, "The chc adaptive search algorithm: How to have safe serach when engaging in nontraditional genetic recombination," in *Foundations of genetic Algorithms*, G. Rawlin, Ed. Morgan Kaufman, 1991, vol. 1, pp. 265–283.

[18] 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.

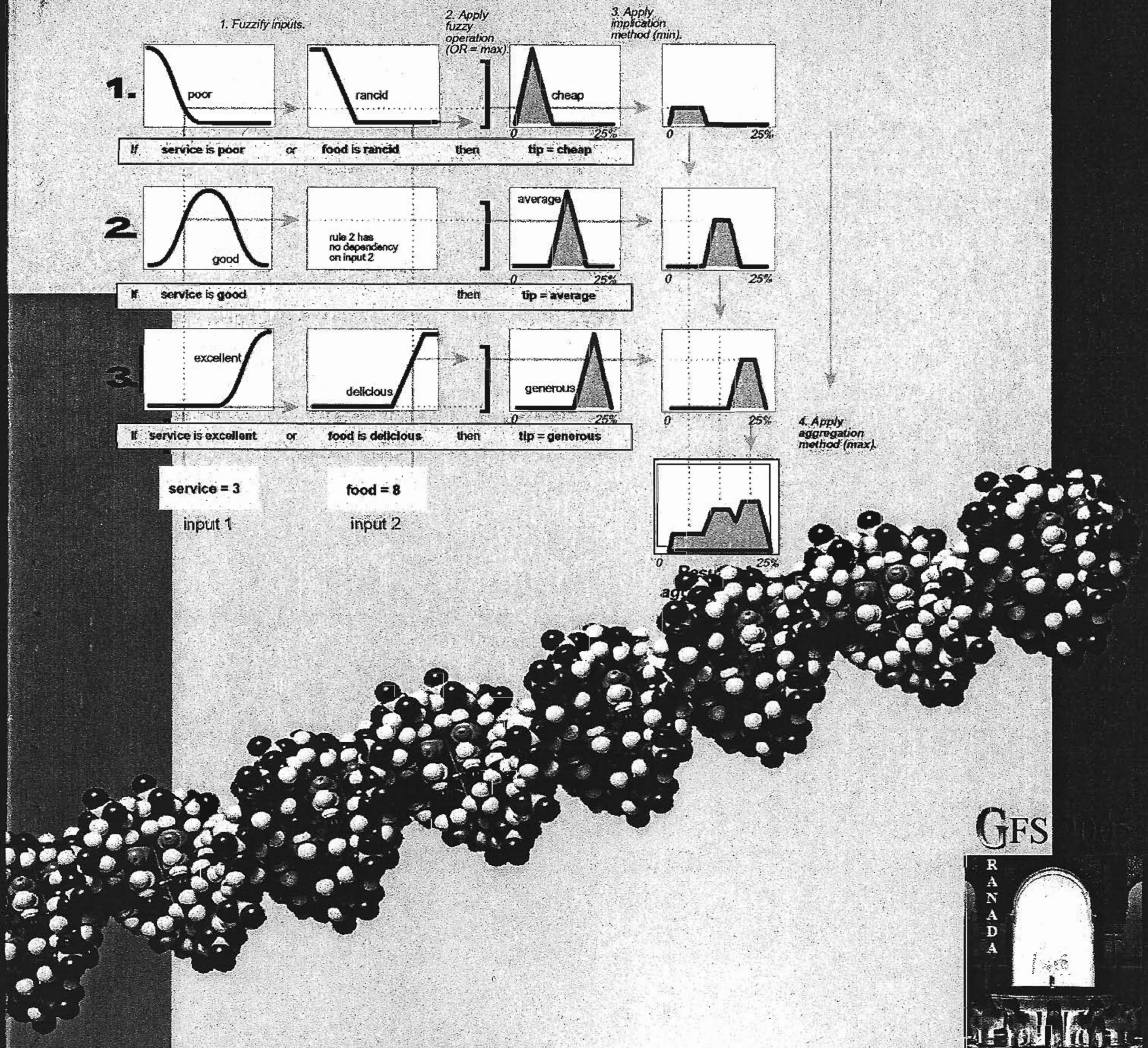
[19] O. Cordón, F. Herrera, and L. Sánchez, "Solving electrical distribution problems using hybrid evolutionary data analysis techniques," *Applied Intelligence*, vol. 10, pp. 5–24, 1999.

[20] J. Casillas, O. Cordón, I. F. de Viana, and F. Herrera, "Learning cooperative linguistic fuzzy rules using the best-worst ant systems algorithm," *International Journal of Intelligent Systems*, 2004, in press.

Proceedings of the

I Workshop on Genetic Fuzzy Systems

Granada, March 17-19, 2005

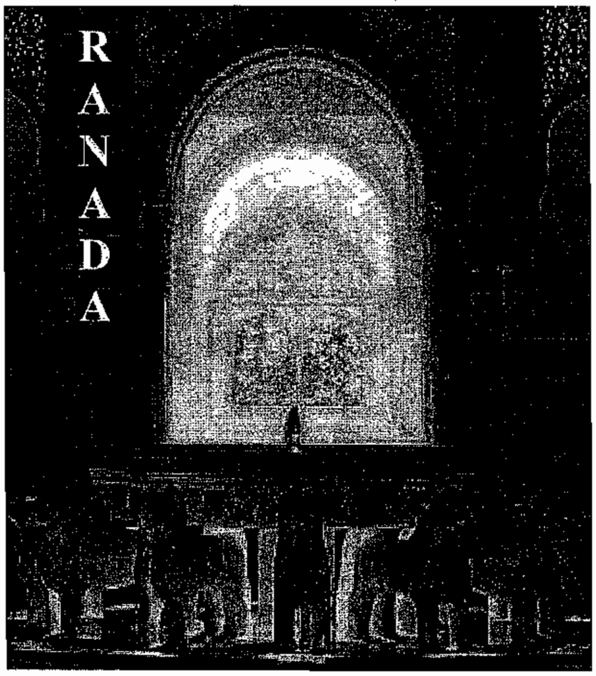


GFS 2005
GRANADA

Proceedings of the
I Workshop on Genetic Fuzzy Systems

Granada, March 17-19, 2005

GFS 2005



Organized by:

Research group on "Soft Computing and Intelligent Information Systems"
(http://sci2s.ugr.es), University of Granada

Edited by:
Rafael Alcalá
Jesús Alcalá-Fdez
Jorge Casillas
Oscar Cordón
María J. del Jesus
Francisco Herrera
Ignacio Rojas
Raúl Pérez
Pedro Villar

Printed in Spain

ISBN: 84-689-1117-8

Applications of Genetic Fuzzy Systems

Evolutionary Learning and Tuning of Fuzzy Rule-Based Systems II

Genetic Fuzzy Systems and the Interpretability-Accuracy Trade-off