Genetic Learning of the Knowledge Base of a Fuzzy System by Using the Linguistic 2-Tuples Representation

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— One of the problems associated to linguistic Fuzzy Modeling is its lack of accuracy when modeling some complex systems. To overcome this problem, many different possibilities of improving the accuracy of linguistic fuzzy modeling have been considered in the specialized literature, maintaining the desired trade-off between accuracy and interpretability.

Recently, a new linguistic rule representation model was presented to perform a genetic lateral tuning of membership functions. It is based on the linguistic 2-tuples representation model that allows 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.

Based on the linguistic 2-tuples representation model, in this work, we present a new method to obtain linguistic fuzzy systems by means of an evolutionary learning of the data base *a priori* (granularity and lateral displacements of the membership functions) and on the use of a basic rule generation method to obtain the whole knowledge base. In this way, the search space reduction provided by the linguistic 2-tuples representation helps to the evolutionary search technique to obtain more precise and compact knowledge bases. Moreover, we analyze this approach considering a real-world problem.

I. INTRODUCTION

One of the problems associated to linguistic Fuzzy Modeling (FM), modeling of systems building a linguistic model clearly interpretable by human beings, is its lack of accuracy when modeling some complex systems. It is due to the inflexibility of the concept of linguistic variable, which imposes hard restrictions to the fuzzy rule structure [1]. This drawback leads linguistic FM to sometimes move away from the desired trade-off between interpretability and accuracy, thus losing the usefulness of the model finally obtained.

To overcome this problem, many different possibilities to improve the accuracy of linguistic FM while preserving its intrinsic interpretability have been considered in the specialized literature [2]. A great number of these approaches share the common idea of improving the way in which the linguistic fuzzy model performs the interpolative reasoning by inducing a better cooperation among the rules composing it. This rule cooperation may be encouraged acting on three different model components: 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—.

This cooperation improvement involves an extension of the linguistic FM and can be achieved, either by means of methods considering the *learning of the RB and the DB*, or by using *post-processing mechanisms* that are applied to improve the system behavior once the RB and the DB are obtained. A particular case of post-processing was proposed in [3], where a new linguistic rule representation model was presented for the genetic tuning of the DB. This approach is based on the linguistic 2-tuples representation [4], which allows the lateral displacement of the labels considering an unique parameter per label. In this way, two main objectives were achieved:

- to obtain linguistic labels containing a set of samples with a better covering degree (accuracy improvements) maintaining their original shapes, and
- to reduce the search space respect to the classical tuning in order to easily obtain optimal models.

Such learning scheme starts from an initial DB and an initial RB that remains fixed during all the tuning process. By this reason, the learning process will be askew by the initial DB and RB, remaining fixed important aspects such as the granularity of the fuzzy partitions, or the number and the structure of the obtained rules. Therefore, it would be desirable a greater degree of cooperation between these two tasks (RB and DB learning) in order to obtain models with a good accuracy-interpretability trade-off.

With this aim, we propose an evolutionary method 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 describes the linguistic rule representation model based on the linguistic 2-tuples. The learning scheme considered for the learning of KBs is introduced in Section III. Section IV

proposes the evolutionary learning method considered in this work. Section V shows an experimental study of the method behavior applied over a real-world estimation problem. Finally, Section VI points out some concluding remarks.

II. THE LINGUISTIC 2-TUPLES REPRESENTATION

In [3], a new model of tuning of Fuzzy Rule-Based Systems (FRBSs) was proposed considering the linguistic 2-tuples representation scheme introduced in [4], 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.
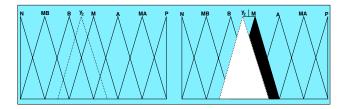

Fig. 1. Lateral Displacement of the Linguistic Label M

Figure 1 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.

The symbolic translation of a linguistic term is a number within the interval [-0.5, 0.5] that expresses the domain of a label when it is moving between its two lateral labels (interval [-1, 1]). Formally, we have the couple,

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

In [4], both the linguistic 2-tuples representation model and the needed elements for linguistic information comparison and aggregation are presented and applied to the Decision Making framework. In the context of the FRBSs, we are going to see its use in the linguistic rule representation. 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:

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

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

Figure 2 shows the concept of classical rule and linguistic 2-tuples represented rule. Analized 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".

This proposal decreases the learning problem complexity, since the 3 parameters considered per label are reduced to only 1 symbolic translation parameter.

In [3], 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 couple (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 couple "X_i is High" is initially considered.

III. LEARNING OF THE KNOWLEDGE BASE

Classically, the learning of the KB of a FRBS is based on the existence of a previous definition of the DB [5], [6]. Generally, to define this DB involves choosing a number of linguistic terms for each linguistic variable and setting the values of the system parameters by a uniform distribution of the linguistic terms. However, this operation mode makes the DB to have a significant influence on the performance of the FRBS finally obtained, since the obtained RB depends on the DB goodness.

In other works, the DB and the RB are simultaneously obtained [7]. In this way, they have the possibility of generating better definitions of the KB, but they deal with a larger search space that makes very difficult to obtain an optimal solution.
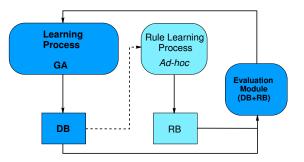

Fig. 3. Learning scheme of the KB

Other way to generate the whole KB consists of obtaining the DB and the RB separately, based on the DB learning *a priori* [8]–[14] (see Figure 3). 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.

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.

  We consider simple triangular membership functions.
- A quick *Ad-hoc data-driven method* to derive the RB [6] 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.

IV. EVOLUTIONARY ALGORITHM FOR LEARNING OF THE KNOWLEDGE BASE

The automatic definition of fuzzy systems can be considered as an optimization or search process and nowadays, Evolutionary Algorithms, particularly GAs, are considered as the more known and used global search technique. Moreover, the genetic coding that they use allow them to include prior knowledge and to use it leading the search up. For this reason, Evolutionary Algorithms have been successfully applied to learn fuzzy systems in the last years, giving way to the appearance of the so called Genetic Fuzzy Systems [15], [16].

Evolutionary Algorithms in general and, GAs in particular, has been widely used in the tuning of FRBSs. In this work, we will consider the use of GAs to design the proposed learning algorithm. CHC [17] 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 "Population-based Selection" approach. M parents and their corresponding offspring are combined to select the best M 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 the evolutionary learning process are explained. They are: DB codification, chromosome evaluation, initial gene pool, crossover operator and restarting approach. A scheme of the algorithm is shown in Figure 4 considering the Wang and Mendel Method.
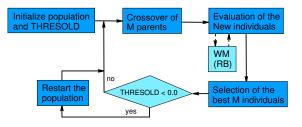

Fig. 4. Scheme of the algorithm

A. 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):

$$C_2 = (\alpha_1^1, \ldots, \alpha_{L^1}^1, \ldots, \alpha_1^N, \ldots, \alpha_{L^N}^N)$$

B. Chromosome Evaluation

To evaluate a determined chromosome we will apply the well-known rule generation method of Wang and Mendel [6] on the DB coded by such chromosome. To do that, each membership function is previously displaced to its new position in the interval defined between the vertexes of its two lateral labels. 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 (to penalize a large 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.

C. 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.

D. 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. Figure 5 depicts the behavior of these kinds of operators. 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 $I = | x_i - y_i |$. 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.
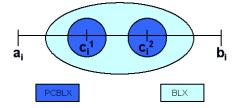

Fig. 5.   Scheme of the behavior of the BLX and PCBLX operators

On the other hand, CHC makes use of an incest prevention mechanism, i.e., two parents are 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.

E. Restarting approach

To get away from local optima, a restarting mechanism is considered [17] when the threshold value L is lower than zero. 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].

V. 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
[6]	**WM**	*Ad-hoc* Data-Driven Method
[5]	**COR**	Cooperative Rules by Using the Best-Worst Ant System
[10]	**GA+WM**	Granularity + Scaling Factors + Contexts + RB by DB learning *a priori* and Using WM
[8]	**GA+COR**	Granularity + Scaling Factors + Contexts + RB by DB learning *a priori* and 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 considered 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. 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

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

- Although the proposed method (learning of the RB, granularity and lateral displacements) can learn partitions presenting a high granularity, the RB is obtained by means of a generation method to learn few rules (57.5,

---

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.

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	–

41.1 or 30.7 from the 6561 possible rules if the input partitions consider nine labels, 9x9) 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 function reduces the search space respect to the classical learning of membership functions which usually considers 3 or 4 parameters (in the case of triangular or trapezoidal membership functions). Therefore, this learning approach represents the ideal framework to be combined with other learning schemes, since these combinations greatly increase the problem search space. It is the case of the combination with the learning of the RB and the granularity of the linguistic partitions, the GLD+WM method. Furthermore, the linguistic models so obtained are interpretable in a high level since the original shapes of the initial membership functions are maintained.

Figures 6 and 7 respectively 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 'l1' to 'lL^i'. Nevertheless, such labels have associated a linguistic meaning determined by an expert. In this way, if the 'l1' label of the 'X1' variable represents 'LOW', 'l1+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 is the case of Figure 6, where the new labels could maintain their initial meanings.
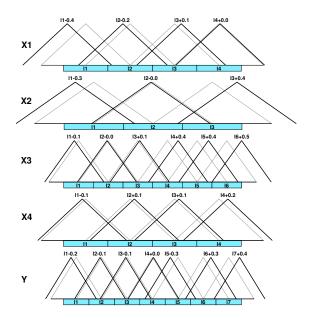
Fig. 6. DB without lateral displacements (in gray) and DB with lateral displacements (in black) of a model obtained by GLD+WM with $\alpha = 5$

X1	X2	X3	X4	Y	X1	X2	X3	X4	Y
I1 -0.41	I1 -0.31	I1 -0.14	I1 -0.08	I1 -0.21	I3+0.15	I2 -0.04	I4+0.36	I2+0.10	I5 -0.30
I1 -0.41	I1 -0.31	I2 -0.04	I1 -0.08	I2 -0.06	I3+0.15	I2 -0.04	I4+0.36	I3+0.10	I6+0.30
I2 -0.16	I1 -0.31	I1 -0.14	I1 -0.08	I1 -0.21	I3+0.15	I2 -0.04	I5+0.38	I2+0.10	I6+0.30
I2 -0.16	I1 -0.31	I1 -0.14	I2+0.10	I2 -0.06	I3+0.15	I2 -0.04	I5+0.38	I3+0.10	I7+0.43
I2 -0.16	I1 -0.31	I2 -0.04	I1 -0.08	I2 -0.06	I3+0.15	I3+0.42	I4+0.36	I2+0.10	I5 -0.30
I2 -0.16	I1 -0.31	I2 -0.04	I2+0.10	I3 -0.06	I3+0.15	I3+0.42	I4+0.36	I3+0.10	I6+0.30
I2 -0.16	I2 -0.04	I2 -0.04	I1 -0.08	I2 -0.06	I3+0.15	I3+0.42	I6+0.50	I2+0.10	I6+0.30
I2 -0.16	I2 -0.04	I2 -0.04	I2+0.10	I3 -0.06	I3+0.15	I3+0.42	I6+0.50	I3+0.10	I7+0.43
I2 -0.16	I2 -0.04	I3+0.06	I1 -0.08	I3 -0.06	I4+0.02	I2 -0.04	I2 -0.04	I1 -0.08	I3 -0.06
I2 -0.16	I2 -0.04	I3+0.06	I2+0.10	I4+0.04	I4+0.02	I2 -0.04	I2 -0.04	I2+0.10	I3 -0.06
I3+0.15	I2 -0.04	I2 -0.04	I1 -0.08	I2 -0.06	I4+0.02	I2 -0.04	I2 -0.04	I3+0.10	I4+0.04
I3+0.15	I2 -0.04	I2 -0.04	I2+0.10	I3 -0.06	I4+0.02	I2 -0.04	I2 -0.04	I4+0.20	I6+0.30
I3+0.15	I2 -0.04	I2 -0.04	I3+0.10	I4+0.04	I4+0.02	I2 -0.04	I3+0.06	I1 -0.08	I4+0.04
I3+0.15	I2 -0.04	I3+0.06	I1 -0.08	I3 -0.06	I4+0.02	I2 -0.04	I3+0.06	I2+0.10	I4+0.04
I3+0.15	I2 -0.04	I3+0.06	I2+0.10	I4+0.04	I4+0.02	I2 -0.04	I3+0.06	I3+0.10	I5 -0.30
I3+0.15	I2 -0.04	I3+0.06	I3+0.10	I5 -0.30	I4+0.02	I2 -0.04	I3+0.06	I4+0.20	I6+0.30

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

Fig. 7. RB and displacements of a model obtained by GLD+WM with $\alpha = 5$

VI. CONCLUSIONS

This work presents a new method for learning KBs by means of an *a priori* evolutionary learning of the DB (granularity and translation parameters). It makes use of the linguistic 2-tuples rule representation model presented in [3]. In the following, we present our conclusions and further works:

- The used learning scheme and the 2-tuples rule representation model allow an important reduction of the search space, obtaining more precise and compact 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 different methods presented in this work. At this moment, we are working on this issue to present a complete study of these kinds of techniques.

The use of different FRBS learning schemes considering the 2-tuples rule representation model seems to be a good approach to obtain more compact and precise models. As further work, we propose other smart combinations, e.g., applying more powerful rule generation approaches as could be an iterative rule learning, the cooperative rule learning, etc.

REFERENCES

[1] A. Bastian, "How to handle the flexibility of linguistic variables with applications," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 3, no. 4, pp. 463–484, 1994.
[2] J. Casillas, O. Cordón, F. Herrera, and L. Magdalena, Eds., *Accuracy improvements in linguistic fuzzy modeling*. Springer-Verlag, 2003.
[3] 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.
[4] 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.
[5] 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*, 2005, in press.
[6] L. Wang and J. Mendel, "Generating fuzzy rules by learning from examples," *IEEE Trans. Syst., Man, Cybern.*, vol. 22, no. 6, pp. 1414–1427, 1992.
[7] 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.
[8] 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.
[9] O. Cordón, F. Herrera, and P. Villar, "Generating the knowledge base of a fuzzy rule-based system by the genetic learning of the data base," *IEEE Trans. Fuzzy Syst.*, vol. 9, no. 4, pp. 667–674, 2001.
[10] 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.
[11] B. Filipic and D. Juricic, "A genetic algorithm to support learning fuzzy control rules from examples," in *Genetic Algorithms and Soft Computing*, F. Herrera and J. L. Verdegay, Eds. Physica-Verlag, 1996, pp. 403–418.
[12] W. Pedrycz, "Associations and rules in data mining: A link analysis," *Int. Journal of Intelligent Systems*, vol. 19, no. 7, pp. 653–670, 2004.
[13] D. Simon, "Sum normal optimization of fuzzy membership functions," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 4, pp. 363–384, 2002.
[14] Y. Teng and W. 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.
[15] 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.
[16] 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.
[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. 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.
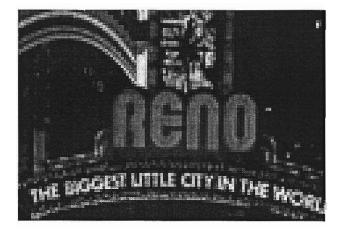
Proceedings of

The 14th IEEE International Conference on Fuzzy Systems

FUZZ-IEEE 2005

"Biggest Little Conference in the World"

May 22-25, 2005
Reno, Nevada, USA

Pattern Recognition III

Fuzzy Association Rules

Model Selection and Rule Extraction III