Hierarchical Knowledge Bases for Fuzzy Rule-Based Systems

O. Cordón, F. Herrera, I. Zwir
Dept. of Computer Science and Artificial Intelligence
E.T.S. de Ingeniería Informática
University of Granada, 18071 - Granada, Spain
E-mail: ocordon, herrera, igor@decsai.ugr.es

Abstract

In this paper we extend the structure of the Knowledge Base of Fuzzy Rule Base Systems in a hierarchical way, in order to make it more flexible. This flexibility will allow us to have linguistic rules defined over linguistic partitions with different granularity levels, and thus to improve the modeling of those problem subspaces where the former models have bad performance.

To do so, we propose a local approach to design linguistic models which are accurate to a high degree and may be suitably interpreted. This approach will be based on the development of a Hierarchical System of Linguistic Rules learning methodology, which has been thought as a refinement of simple linguistic models which, preserves their descriptive power and introduces small changes to increase their accuracy. We also introduce an iterative extension to this method, and compare both with a previous global hierarchical method.

Keywords: Linguistic Modeling, Mamdani-type Fuzzy Rule-Based Systems, hierarchical linguistic partitions, Hierarchical Knowledge Base, rule selection, Genetic Algorithms.

[0]This research has been supported by CICYT PB98-1319.

1 Introduction

One of the most important applications of Fuzzy Rule-Based Systems (FRBSs) is *System Modeling* [1, 10]. *Linguistic Modeling* [12] is the usual type of System Modeling where the main requirement is the interpretability of the model. It also has a problem associated which is its lack of accuracy in some complex problems. This fact is due to some problems related to the linguistic rule structure considered, which are a consequence of the inflexibility of the concept of linguistic variable [15]. To deal with this problem, we extend the Knowledge Base (KB) structure of linguistic FRBSs by introducing the concept of "*layers*". In this extension, which is also a generalization, the KB is composed of a set of layers where each one contains linguistic partitions with different granularity levels and linguistic rules whose linguistic variables take values in these partitions. This KB is called Hierarchical Knowledge Base (HKB), and it is formed by a Hierarchical Data Base (HDB) and a Hierarchical Rule Base (HRB), containing linguistic partitions of the said type and linguistic rules defined over them, respectively.

In this paper, we will show results of three Linguistic Modeling approaches -developed by means of linguistic FRBSs- which allows us to learn HRBs, i.e., Hierarchical Systems of Linguistic Rules Learning Methodologies (HSLR-LMs). First, we will introduce a Two-Level HSLR-LM -whose linguistic variables are defined on a two-level HDB- which is a local approach that, as a simple models refinement, improves its accuracy without losing its interpretability to a high degree. Later, will also show results of an iterative extension of this methodology (more than two-level) and compare the results of both methods with a previous global hierarchical approach.

To do so, this paper is set up as follows. In Section 2, a description of the HKB and the relation between its components is regarded. In Section 3, two methodologies (local and global approaches) to automatically design a HKB from a generic linguistic rule generating method are introduced. In Section 4, a Linguistic Modeling process obtained from previous methodologies and a well-known inductive linguistic rule generation process is applied to solve a real-world application. Finally in Section 5, some concluding remarks are pointed out.

2 Hierarchical Knowledge Base Philosophy

The KB structure usually employed in the field of Linguistic Modeling has the drawback of its lack of accuracy when working with very complex systems. This fact is due to some problems related to the linguistic rule structure considered, which are a consequence of the inflexibility of the concept of linguistic variable [15]. A summary of these problems may be found in [2], and it is briefly enumerated as follows:

- There is a lack of flexibility in the FRBSs because of the rigid partitioning of the input and output spaces.
- When the system input variables are dependent themselves, it is very hard to fuzzy partition the input spaces.
- The homogenous partitioning of the input and output spaces when the input-output mapping varies in complexity within the space is inefficient and does not scale to high dimensional spaces.
- The size of the Rule Base (RB) directly depends on the number of variables and linguistic terms in the system. Obtaining an accurate FRBS requires a significant granularity amount, i.e., it needs of the creation of new linguistic terms. This granularity increase causes the number of rules to rise significantly, which may take the system to lose the capability of being interpretable for human beings.

Due to the inflexibility of the KB structure used in Linguistic Modeling, which as has been said is a consequence of the concept of linguistic variable, we present a more flexible KB structure that allows us to improve the accuracy of linguistic models without losing their interpretability to a high degree: the HKB. It is composed of a set of layers, and each layer is defined by its components in the following way:

$$layer(t, n(t)) = DB(t, n(t)) + RB(t, n(t))$$

with:

- $n(t)$ being the number of linguistic terms that compose the partitions of layer t.
- $DB(t, n(t))$ being the Data Base (DB) which contains the linguistic partitions with granularity level $n(t)$ of layer t.
- $RB(t, n(t))$ being the RB formed by those linguistic rules whose linguistic variables take values in the former partitions.

At this point, we should note that, in this work, we are using *linguistic partitions* with the same number of linguistic terms for all input-output variables, composed of triangular-shaped, symmetrical and uniformly distributed membership functions.

From now on and for the sake of simplicity, we are going to refer to the components of a $DB(t, n(t))$ and $RB(t, n(t))$ as *t-linguistic partitions* and *t-linguistic rules*, respectively.

This set of layers is organized as a hierarchy, where the order is given by the granularity level of the linguistic partition defined in each layer. That is, given two successive layers t and $t+1$, then the granularity level of the linguistic partitions of layer $t+1$ is greater than the ones of layer t. This causes a refinement of the previous layer linguistic partitions. As a consequence of the previous definitions, we could now define the HKB as the union of every layer t:

$$HKB = \cup_t layer(t, n(t))$$

In the remainder of this Section, we are going to study the linguistic partitions and their extension to consider them as component parts of the $DB(t,n(t))$ of the $layer(t,n(t))$. Then, we are going to describe the relation between DBs from different layers (e.g. t and $t+1$), and to develop a methodology to build them under certain requirements. Finally, we will explain how to relate these DBs with linguistic rules, i.e., to create RBs from them.

2.1 Hierarchical Data Base

In this Subsection, we are going to show how to build the HDB, bearing in mind that it is organized in a hierarchy, where the order is given by an increasing granularity level of the linguistic partitions.

To extend the classical linguistic partition, let us consider a partition P of the domain U of a linguistic variable A in the layer t :

$$P_A = \left\{S_1, ..., S_{n(t)}\right\}$$

with S_k $(k = 1,..,n(t))$ being linguistic terms which describe the linguistic variable A. These linguistic terms are mapped into fuzzy sets by the semantic function M, which gives them a meaning: $M_U : S_k \rightarrow \mu_{S_k}(u)$ [15].

We extend this definition of P allowing the existence of several partitions, each one with a different number of linguistic terms, i.e., with a different granularity level. To do so, we add the parameter $n(t)$ to the definition of the linguistic partition P, which represents the granularity level of the partitions contained in the layer t where it is defined:

$$P_A^{n(t)} = \left\{S_1^{n(t)}, ..., S_{n(t)}^{n(t)}\right\}$$

where $P_A^{n(t)} \in DB(t,n(t))$.

In order to build the HDB, we develop an strategy which satisfies two main requirements:

- To preserve all possible fuzzy set structures from one layer to the next in the hierarchy.
- To make smooth transitions between successive layers.

On the one hand, we decided to preserve all the membership function modal points, corresponding to each linguistic term, through the higher layers of the hierarchy in order to fulfill the first requirement. On the other hand, and with the aim of building a new *t+1-linguistic partition*, we just add a new linguistic term between each two consecutive terms of the *t-linguistic partition*. To do so, we reduce the support of these linguistic terms in order to keep place for the new one, which is located in the middle of them. An example of the correspondence among a 1-*linguistic partition,* a 2-*linguistic partition,* and a *3-linguistic partition,* with *n(1)=3, n(2)=5* and *n(3)=9* respectively, is shown in Figure 1.

Table 1: Hierarchy of DBs starting from 2 or 4 initial terms.

DB(t,n(t))		**DB(t,n(t))**
$DB(1,2)$		$DB(1,4)$
$DB(2,3)$		$DB(2,7)$
$DB(3,5)$	*or*	$DB(3,13)$
$DB(4,9)$		$DB(4,25)$
$\vdots$		$\vdots$
$DB(6,33)$		$DB(6,97)$
$\vdots$		$\vdots$

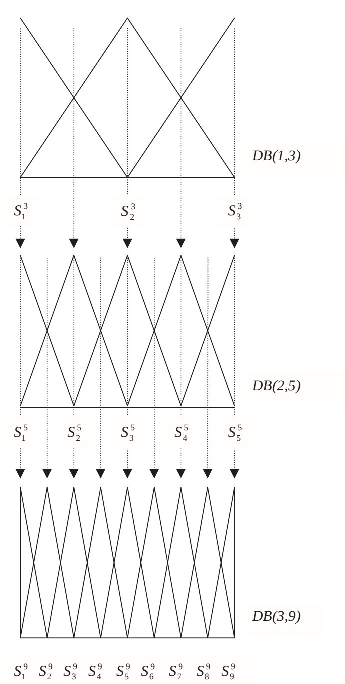

Figure 1: Three layers of linguistic partitions

which compose the HDB

Table 2: Mapping between terms from successive DBs

DB(t,n(t))		**DB(t+1,2·n(t)-1)**
$S_{k-1}^{n(t)}$	$\longrightarrow$	$S_{2k-3}^{2\cdot n(t)-1}$
		$S_{2k-2}^{2\cdot n(t)-1}$
$S_{k}^{n(t)}$	$\longrightarrow$	$S_{2k-1}^{2\cdot n(t)-1}$
		$S_{2k}^{2\cdot n(t)-1}$
$S_{k+1}^{n(t)}$	$\longrightarrow$	$S_{2k+1}^{2\cdot n(t)-1}$

As a result of the above considerations, Table 1 shows the number of linguistic terms which is needed in each *t-linguistic partition* in $DB(t, n(t))$ to satisfy the previous requirements. The values of parameter $n(t)$ represent the *t-linguistic partition* granularity levels and depend on the initial value of $n(t)$ defined in the first layer (e.g. 2 or 4 in Table 1).

Generically, we could say that a DB from a layer $t+1$ is obtained from its predecessor as:

$$DB(t, n(t)) \rightarrow DB(t+1, 2\cdot n(t) - 1)$$

which means that a *t-linguistic partition* in $DB(t, n(t))$ with $n(t)$ linguistic terms becomes a *(t+1)-linguistic partition* in $DB(t+1, 2\cdot n(t)-1)$.

In order to satisfy the previous requirements, each linguistic term $S_k^{n(t)}$ -term of order k from the *t-linguistic partition* in $DB(t, n(t))$ - is mapped into $S_{2k-1}^{2\cdot n(t)-1}$, preserving the former modal points, and a set of *n(t)-1* new terms is created, each one between $S_k^{n(t)}$ and $S_{k+1}^{n(t)}$ $(k = 1, ..., n(t)-1)$. This mapping is clearly shown in Table 2 and a graphical example is to be found in Figure 1.

In this view, we can generalize this two-level successive layer definition for $n(t)$, for all layers t in the following way:

$$n(t) = (N-1)\cdot 2^{t-1} + 1$$

with $n(1) = N$, i.e., the number of linguistic terms in the initial layer partitions.

2.2 Hierarchical Rule Base

In this Subsection we explain how to develop an RB from layer $t+1$ based on $RB(t, n(t))$, $DB(t, n(t))$ and $DB(t+1, 2\cdot n(t)-1)$, in order to create an HRB. Later, in the following Section, we are going to give a concrete method to perform this task for an Iterative Process.

The *t-linguistic RB* structure is formed by a collection of well known Mamdani-type linguistic rules:

$$R_i^{n(t)} : IF\ x_1\ is\ S_{i1}^{n(t)}\ and\ \ldots$$
$$\ldots and\ x_m\ is\ S_{im}^{n(t)}\ THEN\ y\ is\ B_i^{n(t)}$$

with $x_1, \ldots, x_m$ and y being the input linguistic variables and the output one, respectively; and with $S_{i1}^{n(t)}, \ldots, S_{im}^{n(t)}$, $B_i^{n(t)}$ being linguistic terms from different *t-linguistic partitions* of $DB(t, n(t))$, with fuzzy sets associated defining their meaning. In this contribution, we will use the Minimum t-norm in the role of conjunctive and implication operator and the *Center of Gravity weighted by the matching degree* [3] as defuzzification strategy.

The main purpose of developing an HRB is to model the problem space in a more accurate way. To do so, those *t-linguistic rules* that model a subspace with bad performance are expanded into a set of *(t+1)-linguistic rules*, which become their image in $RB(t+1, 2\cdot n(t)-1)$. This set of rules model the same subspace that the former one and replaces it.

We should note that not all *t-linguistic rules* are to be expanded. Only those *t-linguistic rules* which model a subspace of the problem with a significant error become the ones that are involved in this rule expansion process to build the $RB(t+1, 2\cdot n(t)-1)$. The remaining rules preserve their location in $RB(t, n(t))$. An explanation for this behavior could be found in the fact that it is not always true that a set of rules with a higher granularity level, performs a better modeling of a problem than another one, with a lower granularity level. Moreover, this is not true for all kinds of problems, and what is more, it is also not true for all linguistic rules that model a problem [6].

3 System Modeling with an HKB

In this part of the paper we will introduce two methodologies which develop a HKB. On the one hand in the following Subsection a local Two-Level HSLR Learning Methodology (HSLR-LM) and its iterative extension (I-HSLR-LM) are introduced. Later, HSLR is compared with a global approach (G-(I-)HSLR-LM) previously introduced by Ishibuchi et al. in [9].

3.1 A Local Approach: A Two-Level HSLR Learning Methodology (HSLR-LM)

This methodology was proposed in [7] as a strategy to improve simple linguistic models preserving their structure and descriptive power, by reinforcing only the modeling of those problem subspaces with more difficulties by a hierarchical treatment of the rules generated in these zones. Due to this reason, HSLRs are based on two hierarchical levels, i.e., a HKB of two layers.

In the following, the structure of the learning methodology and its most important components are briefly described:

1. *Hierarchical Knowledge Base Generation Process*
   (a) Generate the initial *RB(1,n(1))* from the present *DB(1,n(1))* using any inductive Linguistic Rule Generating method (LRG-method), the initial *1-linguistic partitions* given by an expert, and a training data set.
   (b) Select those bad performance *1-linguistic rules* $RB_{bad}(1,n(1))$, which are going to be expanded, making the difference from the good ones $RB_{good}(1,n(1))$, by comparing their error with the one performed by the whole rule set.
   (c) Obtain the next layer DB, $DB(2{,}2{\cdot}n(1)-1)$.
   (d) Now, for each $R_i^{n(1)} \in RB_{bad}(1,n(1))$:
      i. Select the *2-linguistic partition* terms which have a "significant intersection" with the ones in $R_i^{n(1)}$.
      ii. Combine the previously selected sets.
      iii. Extract *2-linguistic rules* from the combined selected *2-linguistic partition* terms and the use of an LRG-method. These *2-linguistic rules* are the image of the expanded linguistic rule $R_i^{n(1)}$, i.e., the candidates to be in the HRB from rule i, $(CLR(R_i^{n(1)}))$.
   (e) Obtain a joined set of candidate linguistic rules, JCLR, performing the union of the group of the new generated *2-linguistic rules* $(CLR(R_i^{n(1)}))$ and the former good performance *1-linguistic rules* $(RB_{good}(1,n(1)))$:

   $$JCLR = RB_{good}(1,n(1)) \cup (\cup_i CLR(R_i^{n(1)}))$$

   with $R_i^{n(1)} \in RB_{bad}(1,n(1))$.
2. *Hierarchical Rule Base Selection Process.* Simplify the set JCLR by using a genetic linguistic rule selection process, in order to remove the unnecessary rules from it, and to generate an HKB with good cooperation [4, 9]:

   . $HRB = Select(JCLR)$
3. *User Evaluation Process.* Evaluate the obtained model. If it is not appropriate, adapt the granularity of the initial linguistic partitions $n(1)$ and/or the threshold which determine if an *n(t)-linguistic rule* will be expanded in a set of $(2{\cdot}n(t)-1)$*-linguistic rules* α, and apply again the methodology in order to obtain a better model.

We should note that this methodology was thought as an strategy to improve simple linguistic models. Therefore, we could select any inductive LGR-method to build the HRB, based on the existence of a set of input-output data E_{TDS} and a previously defined $DB(1,n(1))$. In order to illustrate this situation, two LRG-methods have been used in [7]: the one proposed by Wang and Mendel in [14] and the one proposed by Thrift in [13].

This *Two-level HSLR-LM* was extended in [8] by considering it as an iterative process. While the former methodology was thought as a *simple descriptive refinement* of linguistic models, the *Iterative HSLR-LM (I-HSLR-LM)* is viewed as an *accurate refinement* of those models, which preserves HSLR-LM features but loses description, having linguistic rules defined over more than two layers in the HRB, in order to improve the modeling accuracy performed by the learned HSLR.

3.2 A Global HSLR Learning Methodology (G-HSLR-LM)

As said, another approach generated in the same line have been performed by Ishibuchi et al. [9] This method obtains an HSLR creating several hierarchical linguistic partitions with different granularity levels, generating the complete set of linguistic rules in each of these partitions, taking the union of all of these sets, and finally performing a genetic rule selection process on the whole rule set. For the sake of simplicity, in this Subsection we will refer to this method as a global HSLR learning methodology (G-HSLR-LM), in order to distinguish it from our local approach (HSLR-LM). Although G-HSLR-LM was designed to construct a fuzzy classification system, and the main purpose of the HSLR-LM proposed in this paper is to perform Linguistic Modeling, some interesting coincidences and differences have been found between them:

Although G-HSLR-LM was designed to construct a fuzzy classification system, and the main purpose of the HSLR-LM proposed in this paper is to perform Linguistic Modeling, some interesting coincidences and differences have been found between them:

- While HSLR-LM locally expands those rules which perform a bad modeling in some subspaces of the problem, G-HSLR-LM performs the same task in a global way, i.e., it expands all rules in all granularity levels.

- Due to the global expansion it performs, G-HSLR-LM allows the HSLR derived from it, to present both the expanded rule and some of the rules composing its image in the next layer RB, thus resulting in a reinforcement of the expanded rule. As said, since HSLR-LM directly substitutes the expanded rule by its image, there is no possibility for this reinforcement.

- Both methods perform a genetic rule selection to extract the set of rules which best cooperates between them, i.e. the HRB, but on a different rule set. We should note that, in order to allow the comparison between both hierarchical methods, the same fitness was used in the GA for both approaches.

Table 3 shows a common notation for both hierarchical methodologies in order to clarify their similarities and differences. We should remember that $CLR(R_i^{n(1)})$ stands for the image of the expanded bad linguistic rule $R_i^{n(1)}$, which joined with the former good performance *1-linguistic rules* constitute the set of candidate linguistic rules to be in the final HRB.

Table 3: Local and Global Selection Processes

HSLR-LM	$HRB = Selection$ $(RB_{good}(t, n(t)) \cup (\cup_i CLR(R_i^{n(1)})))$
G-HSLR-LM	$HRB = Selection$ $(RB(t, n(t)) \cup RB(t+1, n(t+1))))$

4 Examples of Application: Experiments and Analysis of Results

With the aim of analyzing the behavior of the proposed methodology, a real-world electrical engineering distribution problems in Spain have been selected [5, 11]. The concern of this problem is to relate some characteristics of certain village with the actual length of low voltage line contained in it. It would be preferable that the solutions obtained verify another requirement: they have not only to be numerically accurate in the problem solving, but must be able to explain how a specific value is computed for a certain village or town. That is, it is interesting that these solutions are interpretable by human beings to some degree.

Therefore, a relationship must be found between some characteristics of the population and the length of line installed on it, making use of some known data, that may be employed to predict the real length of line in any other village. We will try to solve this problem by generating different models which can determine the unknown relationship, provided with the measured line length (y), the number of inhabitants (x_1) and the mean distance from the center of the town to the three furthest clients (x_2), considered as the radius of population i in the sample, in a sample of 495 rural nuclei [11].

The results obtained with the said methods are shown in Table 4, where *WM(r)* stands for the LRG-method considered with r granularity level linguistic partitions, *HSLR(LRG-method,n(1),n(2))* for the Two-level method with initial and final granularity levels partitions [6] and *I-HSLR(LRG-method,n(1),n(p), k)* as the Iterative method with initial, final granularity levels partitions, and number of iterations [7]. The global methods are described with the same parameters as the former methods but with a prefix (G) indicating their global condition. Additionally, $\#R$ stands for the number of rules of the corresponding HRB, MSE_{tra} and MSE_{tst} for the values obtained in the MSE measure computed over the training and test data sets, respectively. The other parameters used in these experiments are listed in the appendix.

Table 4: Results obtained in the low voltage electrical application considering $\alpha = 1.1$.

Method	MSE_{tra}	MSE_{tst}	$\#R$
WM(3)	594276	626566	**7**
WM(5)	298446	282058	13
WM(9)	197613	283645	29
HSLR(WM,3,5)	178950	167318	12
I-HSLR(WM,3,9,2)	**153976**	**165458**	35
G-HSLR(WM,3,5)	177735	180721	15
G-I-HSLR(WM,3,9,2)	159851	189119	31

In view of the results obtained in the experiments, we should remark some important conclusions:

- *From the accuracy point of view:*

  The different models which make use of the HKB clearly outperform the WM-method ones in all granularity level linguistic partitions and in both data sets, training and test. Now comparing the hierarchical approaches, it can be seen that the linguistic model generated from Two-level HSLR-LM is a little bit less accurate than the G-HSLR(WM,3,5) one in the approximation of the training set, but it has significantly better values for the resulting test errors. Otherwise, the local Iterative methodology outperforms the global and the Two-level ones in both kinds of errors.

- *From the complexity point of view:*

  The hierarchical methods have obtained relatively simple models if we consider the accuracy improvements achieved over the initial models generated by the WM-method. The most clear examples are performed by the comparison of WM(5) or WM(9) with HSLR(WM,3,5). This simpler model become more accurate than the other results in MSE_{tra} and MSE_{tst}, with a lesser number of rules than the most accurate WM-method experiment.

  In view of these results, we should note that it is not always true that a linguistic model whose linguistic variables have terms defined over partitions with higher granularity levels, and consequently with more rules, models better a problem than a simpler one [6]. This is also corroborated in Table 4, where WM(9) does not improve WM(5) in MSE_{tst}. All of this, remarks the importance of the use of local based methods which only improve those difficult subspaces of a problem as a gradual model refinement.

5 Concluding Remarks

In this paper, a HKB has been proposed which is a new approach to design linguistic models accurate to a high degree and suitably interpretable by human beings. Some HKB learning processes capable of automatically generating linguistic models following the said approach have been introduced as well, and their behavior has been compared in solving a real-world problem. The proposed process has obtained very good results.

6 Appendix: Parameters used in the Experiments

The initial DB used for the HSLR-LM is constituted by three primary linguistic partitions formed by *three, four,* and *five linguistic terms* with triangular-shaped fuzzy sets giving meaning to them:

$$DB(1,3) = \{S^3, M^3, L^3\}$$
$$DB(1,4) = \{VS^4, S^4, L^4, VL^4\}$$
$$DB(1,5) = \{VS^5, S^5, M^5, L^5, VL^5\}$$

where S, M, L, VS and VL stand for Small, Medium, Large, Very Small, and Very Large, respectively. The parameters used in all of the experiments are listed in Table 5:

Table 5: Parameters

Generation Parameters	
δ -(2·n-1)-linguistic partition terms selector-	0.1
τ -used to calculate E_i-	0.5
α -used to decide the expansion of rule-	1.1
GA Selection Parameters	
Number of generations	500
Population size	61
Mutation probability	0.1
Crossover probability	0.6

References

[1] A. Bardossy, L. Duckstein, Fuzzy Rule-Based Modeling with Application to Geophysical, Biological and Engineering Systems, CRC Press (1995).

[2] A. Bastian, How to Handle the Flexibility of Linguistic Variables with Applications, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 2:4 (1994) 463-484.

[3] O. Cordón, F. Herrera, and A. Peregrín, Applicability of the Fuzzy Operators in the Design of Fuzzy Logic Controllers, Fuzzy Sets and Systems 86 (1997) 15-41.

[4] O. Cordón, F. Herrera, A Three-stage Evolutionary Process for Learning Descriptive and Approximative Fuzzy Logic Controller Knowledge Bases from Examples, International Journal of Approximate Reasoning 17: 4 (1997) 369-407.

[5] O. Cordón, F. Herrera, L. Sánchez, Solving Electrical Distribution Problems Using Hybrid Evolutionary Data Analyis Techniques, Applied Intelligence 10 (1999) 5-24.

[6] O. Cordón, F. Herrera, Villar P., Analysis and Guidlines to Obtain a Good Uniform Fuzzy Partition Granularity for FRBSs using Simulated Annealing, International Journal of Approximate Reasoning, 2000. To appear.

[7] O. Cordón, F. Herrera, I. Zwir, Linguistic Modeling by Hierarchical Systems of Linguistic Rules, Technical Report #DECSAI-99114, Dept. of Computer Science and Artificial Intelligence, University of Granada, Spain, July 1999.

[8] O. Cordón, F. Herrera, I. Zwir, Hierarchical Systems of Linguistic Rules Learning Methodology: Part I: An Iterative Approach, Technical Report; Part II. Introducing Rule Reinforcement, #DECSAI-000106, Dept. of Computer Science and Artificial Intelligence, University of Granada, Spain, March 2000.

[9] H. Ishibuchi, K. Nozaki, N. Yamamoto, H. Tanaka, Selecting Fuzzy If-Then Rules for Classification Problems Using Genetic Algorithms, IEEE Transactions on Fuzzy Systems 3:3 (1995) 260-270.

[10] W. Pedrycz (Ed.), Fuzzy Modelling: Paradigms and Practice, Kluwer Academic Press (1996).

[11] L. Sánchez, Interval-valued GA-P Algorithms, IEEE Transactions on Evolutionary Computation, 2000. To appear.

[12] M. Sugeno, T. Yasukawa, A Fuzzy-logic-based Approach to Qualitative Modeling, IEEE Transactions on Fuzzy Systems 1:1 (1993) 7-31.

[13] P. Thrift, Fuzzy logic synthesis with genetic algorithms. Proceedings of Fourth International Conference on Genetic Algorithms (ICGA'91), Morgan Kauffman Pub. (1991) 509-513.

[14] L.X. Wang, J.M. Mendel, Generating Fuzzy Rules by Learning from Examples, IEEE Transactions on Systems, Man, and Cybernetics 22 (1992) 1414-1427.

[15] L.A. Zadeh, The Concept of a Linguistic Variable and its Application to Aproximate Reasoning, Information Science, Part I: vol. 8 (1975) 199-249; Part II: vol. 8 (1975) 301-357; Part III: vol. 9 (1975) 43-80.