Linguistic Modeling with Weighted Double-Consequent Fuzzy Rules Based on Cooperative Coevolution *

R. Alcalá
Dept. Computer Science
University of Jaén, E-23071 Jaén, Spain
alcala@ujaen.es

J. Casillas, O. Cordón, F. Herrera
Dept. Computer Science and A.I.
University of Granada, E-18071 Granada, Spain
{casillas,ocordon,herrera}@decsai.ugr.es

Abstract

This paper presents the use of weighted double-consequent rules to improve the accuracy of the fuzzy linguistic models. However, the use of these kinds of rules makes significantly more complex the modeling and learning process. To solve this problem, we use an advanced evolutionary technique, the cooperative coevolutionary model.

Keywords: fuzzy linguistic modeling, double-consequent fuzzy rules, weighted fuzzy rules, cooperative coevolution.

1 Introduction

A way to improve the linguistic modeling (LM) performance with a slight description loss is to extend the usual linguistic model structure to be more flexible. Two possibilities to relax the model structure are:

- *Use of double-consequent rules*, where each combination of antecedents may have two consequents [3, 6].
- *Consideration of weighted rules*, where an importance degree is associated to each rule in the fuzzy reasoning process [2, 7, 12].

In the same way, even more flexible fuzzy rules may be obtained combining both approaches to design linguistic models based on *weighted double-consequent fuzzy rules*, thus involving a potential improvement of the accuracy. However, the use of weights and double-consequents makes significantly more complex the modeling process as it increases the solution search space since new parameters are considered in addition to the traditional approach.

Recently, an advanced evolutionary technique, cooperative coevolution [8, 10, 9], has been proposed to solve problems with a large search space by independently evolving two or more species which together comprise solution structures.

This contribution propose weighted double-consequent rules to design fuzzy linguistic models by means of a cooperative coevolutionary algorithm coevolving two species, the subset of rules best cooperating and the weights associated to them.

The paper is organized as follows. In Section 2, the said ways to relax the model structure are presented. In Section 3, the weighted double-consequent fuzzy rule structure is proposed, as well as a cooperative coevolutionary method to derive these kinds of rules. Experimental results are shown in Section 4, whilst some concluding remarks are pointed out in Section 5.

2 Preliminaries

2.1 Double-Consequent Fuzzy Linguistic Rules

More flexible linguistic models may be obtained by allowing them to present fuzzy rules where each combination of antecedents may have two consequents associated [3, 6]:

IF X_1 is A_1 and ... and X_n is A_n
THEN Y is $\{B_1,B_2\}$,

with X_i (Y) being the linguistic input (output) variables, A_i being the linguistic label used in the i-th input variable, and B_1 and B_2 the two linguistic terms

* Supported by the Spanish CICYT, project PB98-1319

associated to the output variable.

Since each double-consequent fuzzy rule can be decomposed into two different rules with a single consequent, the usual plain fuzzy inference system can be applied. The only restriction imposed is that the defuzzification method must consider the matching degree of the rules fired, for example, the *center of gravity weighted by the matching degree* defuzzification strategy [4] may be used.

The consideration of this structure to generate advanced linguistic models is initially proposed in [6]. Another approach, according to the Accurate Linguistic Modeling (ALM) methodology, is introduced in [3]. This methodology consists of two steps:

1. Firstly, two rules, the primary and secondary in importance, are obtained in each fuzzy input subspace considering a specific generation process. In this contribution, the generation process proposed by Wang and Mendel [11] is considered.
2. Then, after decomposing each double-consequent rule into two independent simple ones, the selection process explained in [3] is employed to select the subset of rules best cooperating. It is based on a binary-coded GA where each gene indicates if the corresponding rule is considered or not in the final rule base.

2.2 Weighted Fuzzy Linguistic Rules

Another possibility to extend the classical model structure making it more flexible is to use weighted rules [2, 7, 12]:

IF X_1 is A_1 and ... and X_n is A_n
THEN Y is B *with* $[w]$,

where w is the real-valued rule weight, and *with* is the operator modeling the weighting of a rule.

With this structure, the fuzzy reasoning must be extended. A possibility is to infer with the FITA (First Infer, Then Aggregate) scheme [4] and compute the defuzzified output, y_0, as the following *weighted sum*:

$$y_0 = \frac{\sum_i m_i \cdot w_i \cdot P_i}{\sum_i m_i \cdot w_i},$$

with m_i being the matching degree of the rule i, w_i being the weight associated to the i-th rule, and P_i being the characteristic value of the output fuzzy set corresponding to the i-th rule. In this paper, the center of gravity will be considered as characteristic value [4].

These weights are usually considered to handle inconsistencies with advanced inference methods [12] or neural networks [2]. Moreover, some proposals make use of them to improve the model accuracy with an automatic learning of weights using different techniques such as gradient descent processes [7].

3 Evolutionary Learning of Weighted Double-Consequent Rules

This section proposes the use of a more flexible linguistic model structure that combines the two said approaches, thus having weighted double-consequent rules with the following structure:

IF X_1 is A_1 and ... and X_n is A_n
THEN Y is $\{B_1, B_2\}$ *with* $[w_1, w_2]$,

with w_1 and w_2 being the weights associated to the rules composed using the consequents B_1 and B_2, respectively. Hence, a weighted double-consequent rule can be seen as two weighted single-consequent rules with the same antecedent and different consequents.

To generate linguistic models with this new structure, we may follow an operation mode similar to the ALM methodology [3] introduced in the previous section, but including the weight learning. Therefore, after the first step of the ALM methodology, where an initial set of numerous double-consequents rules is generated, the two following tasks must be performed:

- Selection of a subset of cooperative rules.
- Derivation of the weights for these rules.

These interdependent tasks significantly increase the search space with respect to the original methodology making the choice of the search technique considered crucial.

3.1 The Cooperative Coevolutionary Algorithm

A coevolutionary algorithm [8] involves two or more species (populations) that permanently interact among them by a coupled fitness. Thereby, in spite of each species has its own coding scheme and reproduction operators, when an individual must be evaluated, its goodness will be calculated considering some individuals of the other species. This coevolution makes easier to find solutions to complex problems.

As we have seen, the problem that concerns us can be easily decomposed into two subtasks, the rule selection and the weight derivation. Therefore, it can be solved by coevolving two species cooperating to form the complete solution by learning a set of weighted rules. In the following subsections, the main characteristics of the proposed cooperative coevolutionary algorithm are presented.

3.1.1 Interaction Scheme Between Species

The objective will be to minimize the well known *mean square error* (MSE):

$$\mathrm{MSE}_{ij} = \frac{1}{2 \cdot N} \sum_{l=1}^{N} (F_{ij}(x^l) - y^l)^2 \ ,$$

with N being the number of training data, $F_{ij}(x^l)$ being the output inferred from the model obtained by combining the individuals i and j of the species 1 and 2 when the input x^l is presented, and y^l being the known desired output.

Thus, individuals in the species 1 and 2, are respectively evaluated with the fitness functions f_1 and f_2, defined as follows:

$$f_1(i) = \min_{j \in R_2 \cup P_2} \mathrm{MSE}_{ij} \ and \ f_2(j) = \min_{i \in R_1 \cup P_1} \mathrm{MSE}_{ij} \ ,$$

with i and j being individuals of species 1 and 2 respectively, R_1 and R_2 being the set of the fittest individuals in the previous generation of the species 1 and 2 respectively, and P_1 and P_2 being individual sets selected at random from the previous generation of the species 1 and 2 respectively. The combined use of these kinds of sets make the algorithm have a trade-off between exploitation ($R_{1|2}$) and exploration ($P_{1|2}$). The cardinalities of the sets $R_{1|2}$ and $P_{1|2}$ are previously defined by the designer.

3.1.2 Species 1: Fuzzy rule selection

For the species 1, we will use the genetic rule selection method proposed in [3]. The **coding scheme** generates binary-coded strings of length m (number of single-consequent rules in the previously derived rule set). Depending on whether a rule is selected or not, the alleles "1" or "0" will be respectively assigned to the corresponding gene. Thus, a chromosome c_1^p will be a binary vector representing the subset of rules finally obtained.

The whole **initial pool** is generated at random but one individual, which represents the complete previously obtained rule set. For this species, the standard two-point **crossover** operator is used. As regards the **mutation** operator, it flips the value of the gene.

3.1.3 Species 2: Weights derivation

The **coding scheme** generates real-coded strings of length m. The value of each gene indicates the weight used in the corresponding rule. They may take any value in the interval $[0, 1]$. Now, a chromosome c_2^p will be a real-valued vector representing the weights associated to the fuzzy rules considered.

The **initial pool** for this species is generated with a chromosome having all the genes with the value "1", and the remaining individuals taking values randomly generated within the interval $[0, 1]$. The max-min-arithmetical **crossover** operator is considered. As regards the **mutation** operator, it simply involves changing the value of the selected gene by other value obtained at random within the interval $[0, 1]$.

4 Experimental Results in the Electrical Maintenance Cost Estimating Problem

We will analyze the accuracy of the linguistic models generated from the proposed process (WALM-CC) compared to the four following methods: the ad hoc data-driven method proposed by Wang and Mendel (WM) [11], the said ALM method [3], a simple GA that learns the weights of the rule set derived by WM (WRL), and a simple GA that learns weighted double-consequent rules as a first approximation to the problem (WALM). With this aim, we have chosen the problem of estimating the maintenance costs of the medium voltage electrical network in a town [5]. This problem has been already presented and used in [1].

The *minimum t-norm* have been selected as implication and conjunctive operator, and the *center of gravity weighted by the matching* strategy as defuzzification operator [4].

4.1 Experimental Results and Analysis

The following values have been considered for the parameters of each method: 61 individuals (except for WALM-CC with 31 for each species), 1000 generations, 0.6 as crossover probability, 0.2 as mutation probability, 0.35 for the weight factor in the max-min-arithmetical crossover (when weights are considered). The three fittest individuals ($|R_{1|2}| = 3$) and two random individuals ($|P_{1|2}| = 2$) of each species are considered for the coupled fitness in WALM-CC.

The results obtained by the five methods analyzed are shown in Table 1, where *#R* stands for the number of rules, and MSE_{tra} and MSE_{tst} respectively for the error obtained over the training and test data. The best results are shown in boldface.

Table 1: Results obtained in the electrical problem

Method	#R	MSE_{tra}	MSE_{tst}
WM	66	71,294	80,934
ALM	**50**	51,714	58,806
WRL	66	33,639	33,319
WALM	61	37,356	42,477
WALM-CC	59	**24,961**	**28,225**

In view of the obtained results, we can see that the ALM and WRL methods achieve significant improvements over the WM method. As regards the WALM method, the simple GA-based method to learn weighted double-consequent fuzzy rules, the results obtained were worse than the ones from WRL, when theoretically the former should obtain better results. This confirms the need of a more sophisticated optimization technique. Indeed, analyzing the model obtained by the WALM-CC method, we can conclude that it presents the best performance in both approximation (MSE_{tra}) and generalization (MSE_{tst}).

5 Concluding Remarks and Further Work

The very accurate results of the proposed coevolutive learning method, in conparison with other related approaches, has been contrasted when solving a real-world problem. They are a consequence of the coevolutionary approach ability to tackle with decomposable complex problems.

Notice that the proposed method does not perform membership function learning, which gives more freedom degrees to the model structure involving more accurate results [1]. As further works, we propose to add other mechanisms —like the one proposed in [1]— that make the linguistic models more flexible.

References

[1] J. Casillas, O. Cordón, F. Herrera, A cooperative coevolutionary algorithm for jointly learning fuzzy rule bases and membership functions, Proc. 2nd EUSFLAT Conf., Leicester, England, 2001.

[2] J.-S. Cho, D.-J. Park, Novel fuzzy logic control based on weighting of partially inconsistent rules using neural network, J. Intelligent and Fuzzy Syst. 8 (2000) 99–110.

[3] O. Cordón, F. Herrera, A proposal for improving the accuracy of linguistic modeling, IEEE Trans. Fuzzy Syst. 8:3 (2000) 335–344.

[4] O. Cordón, F. Herrera, A. Peregrín, Applicability of the fuzzy operators in the design of fuzzy logic controllers, Fuzzy Sets Syst. 86:1 (1997) 15–41.

[5] O. Cordón, F. Herrera, L. Sánchez, Solving electrical distribution problems using hybrid evolutionary data analysis techniques, Applied Intelligence 10 (1999) 5–24.

[6] K. Nozaki, H. Ishibuchi, H. Tanaka, A simple but powerful heuristic method for generating fuzzy rules from numerical data, Fuzzy Sets Syst. 86 (1997) 251-270.

[7] N.R. Pal, K. Pal, Handling of inconsistent rules with an extended model of fuzzy reasoning, J. Intelligent and Fuzzy Syst. 7 (1999) 55–73.

[8] J. Paredis, Coevolutionary computation, Artificial Life 2 (1995) 355–375.

[9] C.A. Peña-Reyes, M. Sipper, Fuzzy CoCo: a cooperative coevolutionary approach to fuzzy modeling, IEEE Trans. Fuzzy Syst. (to appear). Draft available at http://lslwww.epfl.ch/~penha.

[10] M.A. Potter, K.A. De Jong, Cooperative coevolution: an architecture for evolving coadapted subcomponents, Evolutionary Computation 8:1 (2000) 1–29.

[11] L.X. Wang, J.M. Mendel, Generating fuzzy rules by learning from examples, IEEE Trans. Syst. Man. Cybern. 22 (1992) 1414–1427.

[12] W. Yu, Z. Bien, Design of fuzzy logic controller with inconsistent rule base, J. Intelligent and Fuzzy Syst. 2 (1994) 147–159.