Learning Accurate TSK Models Based on Local Fuzzy Prototyping

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

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

Abstract

This work presents the use of local fuzzy prototypes as a first approximation to obtain accurate local semantics-based Takagi-Sugeno-Kang rules. A two-stage evolutionary algorithm considering the interaction between input and output variables has been developed. Firstly, it performs a local identification of prototypes, and then, a post-processing stage is considered to refine them. The proposal has been tested with a real-world problem achieving good results.

Keywords: TSK fuzzy models, fuzzy prototypes, genetic fuzzy rule-based systems.

1 Introduction

In Takagi-Sugeno-Kang (TSK) Fuzzy Rule-Based Systems (FRBSs) [10], the learning of the premises and consequents is usually performed separately, obtaining the optimum consequents for a previously learned premise set without considering the interaction between input and output variables.

Fuzzy clustering is one of the most useful techniques, detecting the possible groupings and establishing some hypothesis about the structure present in the data. However, some authors state several drawbacks: they are very sensitive to the presence of outliers and the cluster centers or estimates for the parameters are poor [8].

Recently, these kinds of learning techniques have been taken into account as prototype-identification algorithms, summarizing a dataset by a number of representative prototypes. However, they still comes with the same drawbacks. In this way, fuzzy rule generation methods also can be seen as identification algorithms of fuzzy rule prototypes, i.e., fuzzy model builders whose main purpose is to extract the most suitable set of fuzzy rules from an object (input-output data) according to an optimization measure. Additionally, they organize results and summarize them by interest criteria, in order to provide a more compact and useful representation of the resulting structures.

Two main approaches can be considered to obtain FRBSs. A *global semantics-based approach* where a global collection of fuzzy sets is considered by all the fuzzy rules and, a *local semantics-based approach* where each fuzzy rule has associated its own local fuzzy sets.

Our main objective in this paper is to obtain highly accurate fuzzy models although it involves to loss interpretability to some degree. Therefore, we propose the use of local semantics-based Mamdani fuzzy rules as local fuzzy prototypes to obtain accurate local semantics-based TSK rules, considering the interaction between input and output variables and taking into account the fuzzy nature of these kinds of rules. To do so, we present a two-stage Genetic FRBS (GFRBS) following the MOGUL paradigm [4], a methodology to obtain GFRBSs under the Iterative Rule Learning (IRL) approach.

The paper is organized as follows. Next section describes the general TSK fuzzy model structure. Section 3 discusses the main differences between the global and the local fuzzy prototype identification. Section 4 presents the structure of the proposed GFRBS. Experimental results are shown in Section 5. Finally, some concluding remarks are pointed out in Section 6.

2 Preliminaries: TSK FRBS Architecture and Main Features

In [10], Takagi, Sugeno and Kang presented a mathematical tool to procure a fuzzy model of a system. The fuzzy model is based on rules usually presenting the following structure:

$$\begin{array}{l} R_i : If\ X_1\ is\ A_{i1}\ and\ \ldots\ and\ X_n\ is\ A_{in} \\ Then\ Y = p_{i1} \cdot X_1 + \ldots + p_{in} \cdot X_n + p_{i0} \end{array},$$

where X_i are the system input variables and Y is the system output variable that determines a local linear input-output relation by means of the real-valued coefficients p_{ij}. The output of an FRBS considering a knowledge base composed of m TSK rules is computed as the weighted average of the individual rule outputs, y_i —$i = 1, \ldots, m$—:

$$\frac{\sum_{i=1}^{m} h_i \cdot y_i}{\sum_{i=1}^{m} h_i},$$

with $h_i = T(A_1(x_1), \ldots, A_n(x_n))$ being the matching degree between the antecedent part of the i-th rule and the current system inputs —$x_1, \ldots, x_n$—, and with T being a t-norm.

As seen, these partial relations are combined by aggregation, taking into account their dominance in their respective area of application and the conflict existing in the overlapped areas [10]. In this way, TSK fuzzy systems present the following interesting features: *locality*, *smooth switching* and *existence of mathematical tools for system design*.

3 Local vs Global Fuzzy Rule Prototype Identification

In this work, we propose the use of Mamdani fuzzy rules as fuzzy prototypes to identify a set of fuzzy subspaces grouping data with similar behavior. As we have seen, two different approaches can be considered to obtain these kinds of rules, those based on global and local semantics, obtaining global or local fuzzy prototypes respectively. Each of them presents different advantages and drawbacks in terms of accuracy and interpretability.

The approaches based on global semantics, as a consequence of the inflexibility of the concept of linguistic variable, present the following drawbacks: there is a lack of flexibility because of the rigid partitioning of the spaces; the fuzzy partition input spaces is hard when the input variables are dependent themselves; the partitioning of the spaces is inefficient and does not scale when the input-output mapping varies in complexity within the space; the size of the fuzzy rule base directly depends on the number of variables and linguistic terms in the system.

These drawbacks are solved by local semantics-based approaches presenting interesting advantages: *the expressive power of the rules* that present their own specificity in terms of the fuzzy sets involved in them and, *the number of rules is adapted to the complexity of the problem*. However, an important drawback of this local approach is that these kinds of systems are less interpretable than the global ones.

In this paper we will focus on developing more accurate fuzzy models, so we propose the use of local semantics-based Mamdani rules as local fuzzy prototypes for local identification of TSK fuzzy rules. The local fuzzy prototypes are based on rules presenting the following structure, where A_i and B are fuzzy sets specific to each fuzzy rule:

$$R_i : If\ X_1\ is\ A_{i1}\ and\ \ldots\ and\ X_n\ is\ A_{in}\ Then\ Y\ is\ B$$

4 Structure of the Proposed GFRBS

In this section, we present a two-stage GFRBS to generate local semantics-based TSK FRBSs. It is based on the existence of a set of input-output training data $E_N = \{e_1, \ldots, e_l, \ldots, e_N\}$ with $e_l = (ex_1^l, \ldots, ex_n^l, ey^l)$, N being the data set size, and n being the number of input variables.

The local identification of prototypes induces competition among rules by only considering the quality of the approximation performed by each rule. To do so, the proposed method has been integrated in MOGUL [4] by using an IRL-based approach. However, the global cooperation among rules should be considered in order to increase the generalization power of the system modeled. Following the MOGUL approach, a post-processing stage is considered for this purpose. In this way, the learning method substantially reduces the search space size by dividing the genetic learning process into two stages.

4.1 Local Process for Identifying Prototypes

To obtain the set of local semantics-based Mamdani fuzzy rules (fuzzy prototypes) we have used a method described in [2]. Since this method is based on local covering measures to induce competition among rules, considering the *completeness* and *consistency* properties [4] is recommendable to improve the behavior of the generated fuzzy rule bases. In our case, completeness is verified demanding that each example is covered to a degree $\varepsilon \in \Re$. On the other hand, to verify the consistency, the *positive* and *negative example* concepts [4] are considered. Thereby, the accuracy of a simple fuzzy rule[1], R_i, on the set of examples, E_N, is measured by using a multicriteria fitness function:

$$F(R_i) = \Psi_{E_N}(R_i) \cdot G_\omega(R_i) \cdot g_n(R_i^-),$$

designed to take into account three different criteria [4]: *high frequency value* $(\Psi_{E_N}(R_i))$, *high average covering degree over positive examples* $(G_\omega(R_i))$ and *small negative example set* $(g_n(R_i^-))$. This method may be briefly summarized in the following steps:

a) Perform a strong fuzzy partition for each variable (uniform triangular-shaped membership functions).

b) Generate for each e_l the global fuzzy rule best covering it. Then, evaluate all the global fuzzy rules and select the rule with the highest value in the fitness function $(F(R_i))$.

c) The most promising global fuzzy rule is locally tuned to identify the local fuzzy prototype best grouping the data located in the corresponding subspace. This process is computed by means of the (1+1)-Evolution Strategy ((1+1)-ES) described in [2] considering as fitness function $F'(R_i) = F(R_i) \cdot LNIR(R_i)$ where $LNIR(R_i)$ is a penalty function to avoid excessive proximity among prototypes.

d) Finally, the obtained prototype is added to the final set of fuzzy prototypes. The data covered by this set to a certain degree are removed and not considered for future iterations. The iterative process ends up when no more uncovered training data remains.

To obtain the TSK consequents, once the set of local fuzzy prototypes is obtained and considering the same antecedents, the existing partial linear input-output relation is computed using the data located in each input subspace by means of the (μ,λ)-ES presented in [1] to minimize the Mean Square Error (MSE).

[1] With global or local nature.

4.2 Post-Processing Stage

Two different processes are considered at this stage to minimize the final MSE, the *genetic simplification process* and the *genetic tuning process*:

a) *Genetic Simplification Process.*

This process, described in [2], is based on a standard binary-coded Genetic Algorithm (GA). It has the aim of selecting the subset of rules best cooperating among the rules generated in the previous stage.

b) *Genetic Tuning Process.*

This method is an adaptation of that proposed in [1] to tune TSK FRBSs based on global semantics. It is based on a hybrid GA-ES algorithm in which each individual represents a complete knowledge base. An (1+1)-ES is considered as a genetic operator to locally tune a percentage δ of the best individuals in each generation. Deal with local semantics, the variation interval for each fuzzy set is himself for each individual.

5 Experimental Results in the Estimating the Length of Low Voltage Lines

To analyze the behavior of the proposed TSK GFRBS (M→TSK) we have chosen the problem of estimating the length of low voltage lines for an electric company [3]. Several methods are considered for comparison, two local and five global semantics based methods. The local semantics-based methods are: the Gustafson-Kessel fuzzy clustering method (FMID) [5] and a method to perform function approximation with a generalized regression neural network (GRNN) [11]. On other hand, the global semantics-based methods are: the first approach to learn global semantics-based TSK FRBSs (TS) [10], a single-stage GFRBS (LT) [7], a neural FRBS (ANFIS) [6], a two-stage GFRBS (M-TSK) [1] and Fuzzy C-Means combined with a validation index (SY-FCM) [9].

The initial linguistic partitions are comprised by *five linguistic terms* with triangular-shaped fuzzy sets. The same values for all the related parameters have been considered (i.e., 0.6 as crossover probability in all the genetic approaches). The values of the parameters considered for M→TSK are the following: *Local* *process for identifying prototypes*: $\varepsilon = 1.5$, $\omega = 0.05$ and $k = 0.1$ in the fitness function; $c = 0.9$ and 100 iterations for the (1+1)-ES; $\mu = 15$, $\lambda = 100$, $\gamma = 0.2 \cdot \mu = 3$, $\theta = 0.7$, $q = 5$, $\vec{r} = (r_{\vec{x}}, r_{\vec{\sigma}}, r_{\vec{\alpha}}) = (2,0,0)$, $\vec{\zeta} = (\zeta_{\vec{x}}, \zeta_{\vec{\sigma}}, \zeta_{\vec{\alpha}}) = (\mu, \mu, 1)$, $(n_\sigma, n_\alpha) = (0,0)$ and 500 iterations for the (μ,λ)-ES; *Genetic simplification process*: $N = 61$, $P_c = 0.6$, $P_m = 0.1$ and 500 generations; *Genetic tuning process*: $N = 61$, $P_c = 0.6$, $P_m = 0.1$, $a = 0.35$, $b = 5$, $d = 0.001$, 1000 generations, 25 (1+1)-ES iterations, $\alpha = 0$ and $c = 0.9$ (the updating amount of the Rechenberg's $1/5$-success rule in the (1+1)-ES [**?**]). The results obtained by the methods analyzed are shown in Table 1.

Table 1: Results obtained in the electrical problem

Method	#R/Complex.	$\mathbf{MSE_{tra}}$	$\mathbf{MSE_{tst}}$
TS	4	170644	158949
LT	4	169761	160110
M-TSK	20	132917	167826
ANFIS	20	108279	923650
SY-FCM	5	508426	464130
FMID	5	181040	164670
GRNN	(396 × 2) *neur.*	170460	198140
M→TSK	26	**124022**	**137752**

In view of these results, the proposed method obtained the best results, with improvements of about a 13.34% in test and a 27.32% in training over TS. Moreover, M→TSK present a good balance between approximation and generalization, even considering a higher number of rules. This fact denotes the good fuzzy partitioning that this method achieves.

6 Concluding Remarks

Our proposal shown the most accurate results in approximation and specially in generalization when solving a real-world problem. The obtained model presented a moderate number of rules and can be interpreted from a local point of view. Moreover, this method presents an appropriate convergence.

References

[1] O. Cordón, F. Herrera. A two-stage evolutionary process for designing TSK fuzzy rule-based systems. *IEEE Transactions on Systems, Man, and Cybernetics. Part B: Cybernetics*, volume 29:6, pages 703-715, 1999.

[2] O. Cordón, F. Herrera. Hybridizing genetic algorithms with sharing scheme and evolution strategies for designing approximate fuzzy rule-based systems. *Fuzzy Sets and Systems*, volume 118:2, pages 235-255, 2001.

[3] O. Cordón, F. Herrera, L. Sánchez. Solving electrical distribution problems using hybrid evolutionary data analysis techniques. *Applied Intelligence*, volume 10, pages 5-24, 1999.

[4] O. Cordón, M.J. del Jesus, F. Herrera, M. Lozano. MOGUL: A methodology to obtain genetic fuzzy rule-based systems under the iterative rule learning approach. *International Journal of Intelligent Systems*, volume 14:9, pages 1123-1153, 1999.

[5] D. Gustafson, W. Kessel. Fuzzy clustering with a fuzzy covariance matrix. *Proc. of the IEEE CDC*, pages 761-766, San Diego (USA), 1979.

[6] J.R. Jang. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, volume 23:3, pages 665-684, 1993.

[7] M.A. Lee, H. Takagi. Integrating design stages of fuzzy systems using genetic algorithms. *Proc. of the 2nd IEEE International Conference on Fuzzy Systems (FUZZ-IEEE'93)*, pages 613-617, San Francisco (USA), 1993.

[8] M. Ménard, Fuzzy clustering and switching regression models using ambiguity and distance rejects. *Fuzzy Sets and Systems*, volume 122, pages 363-399, 2001.

[9] M. Sugeno, T. Yasukawa. A fuzzy-logic-based approach to qualitative modeling. *IEEE Transactions on Fuzzy Systems*, volume 1:1, pages 7-31, 1993.

[10] T. Takagi, M. Sugeno. Fuzzy identification of systems and its application to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, volume 15:1, pages 116-132, 1985.

[11] P.D. Wasserman. Advanced Methods in Neural Computing. Pages 155-161, Van Nostrand Reinhold (NY), 1993.