Tuning of Fuzzy Inference Systems Through Unconstrained Optimization Techniques

ROGERIO ANDRADE FLAUZINO, IVAN NUNES DA SILVA
Department of Electrical Engineering
State University of São Paulo – UNESP
CP 473, CEP 17033-360, Bauru, São Paulo
BRAZIL
raflauzino@ladaps.feb.unesp.br, ivan@feb.unesp.br

Abstract: - This paper presents a new methodology for the adjustment of fuzzy inference systems. A novel approach, which uses unconstrained optimization techniques, is developed in order to adjust the free parameters of the fuzzy inference system, such as its own parameters of the membership function, and the weight of the inference rules. This methodology is interesting, not only for the results presented and obtained through computer simulations, but also for its generality concerning to the kind of fuzzy inference system used. Therefore, this methodology is expandable either to the Mandani architecture or also to that suggested by Takagi-Sugeno. The validation of the presented methodology is accomplished through estimation of time series. More specifically, the Mackey-Glass chaotic time series is used for the validation of the proposed methodology.

Keywords: - Fuzzy systems, optimization problems, parameters estimation, time series analysis, neural systems.

1 Introduction

The fuzzy inference systems design comes along with several decisions taken by the designers since it is necessary to determine, in a coherent way, the number of membership functions for the outputs and inputs, and also the specification of the fuzzy rules sets of the system, besides defining the strategies of rules aggregation and defuzzification of output sets. The need to develop systematic procedures to help the designers has been wide, since very often the trial and error technique is the only one available [1].

At present time, there are several researchers engaged in studies related to the design techniques involving fuzzy inference systems. A brief resume about the different approaches for tuning of fuzzy inference systems may be found in [2].

This paper presents a methodology of tuning fuzzy inference systems based on unconstrained optimization techniques. This methodology has the objective of minimizing an energy function associated to the fuzzy inference system.

The definition of the energy function must obey the performance requirements of the fuzzy system. This way, the energy function may be explicited as the mean squared error between the output of the fuzzy system and the desired results, which are provided by the tuning set, similar to the artificial neural networks with supervised training. The energy function may be also defined through the performance parameters desired to the fuzzy system behavior, as it happens in the adjustment of the process controllers acting in a determined plant.

It may be observed that the correct definition of energy function is fundamental to the success of the desired adjustment [3].

Generally, in applications involving the identification and fuzzy modeling, it is convenient to use energy functions that express the error between the desired results and the results provided by the fuzzy system. An example is the use of the mean squared error and normalized mean squared error as energy functions. In the context of systems identification, besides the mean squared error, data regularization indicators may be added to the energy function in order to improve the system response in presence of noises (from training data) or when the tuning set has a constrained data quantity.

In the absence of a tuning set, as it happens in parameters adjustment of a process controller, the energy function can be defined through a function that considers the desired requirements of a design [4]. The example of tuning a fuzzy controller, such requirements may be, for instance, maximum overshoot signal, setting time, rise time, or less usual ones, like the undamped natural frequency of the system among others.

Therefore, the definition of the energy function in the context of tuning fuzzy inference systems becomes part of their specifications, as it occurs with the inputs number, the membership functions, and the choice of the used techniques in fuzzification processes, rules inference and defuzzification.

Besides adding a new task to the process of fuzzy inference system creation, the definition of the energy function associated to the system, jointly with the techniques underlined in this paper, saves the designers efforts in steps of tuning the membership functions, including the creation of the fuzzy inference rules.

Using this new approach from the definition point of view, the fuzzy system becomes defined as a three layers model. Each one of these layers represents the tasks performed by the fuzzy inference system, such as fuzzification tasks, fuzzy rules inference and defuzzification.

The fuzzy inference system adjustment proposed in this paper is performed through the adaptation of the free parameters from each one of these layers, with the objective of minimizing the energy function previously defined.

In principle, the adjustment can be made layer by layer separately, nevertheless not preventing it to be made in all layers at each iteration. The operational differentiation of each layer, where the parameters adjustment of a layer doesn't influences the performance of the others, allows the individual adjustment of the layers. Thus, the routine of fuzzy inference system tuning acquires a larger flexibility when compared to the training process used in artificial neural networks.

To validate the proposed methodology, it is developed a fuzzy inference system to predict the Mackey-Glass time series. Due to its chaotic nature, this estimation problem offers an adequate application to validation of the hereby-underlined approach [5].

This article is organized as follows. In Section 2, a review about fuzzy inference systems is presented to elaborate all needed considerations to the tuning methodology coming on Section 3. In Section 4 is made a brief resume about the Glass-Glass chaotic time series. In Section 5, the simulation results are presented. Finally, conclusions are described in Section 6.

2 Fuzzy Inference Systems

The fuzzy inference systems may be treated as systems that use the concepts and operations defined by the fuzzy set theory, since they use the fuzzy inference process to perform their operational functions. Basically, these operational functions include the inputs fuzzification of the system, the inference rules associated to it, the aggregation of rules and the later defuzzification of the aggregation results, which represent the outputs of the fuzzy system [6].

This way, it may be observed that the fuzzy inference systems have different functions clearly defined, allowing those systems the functions interpretation through the representation of a multilayer model.

Considering the operational functions performed by the fuzzy inference systems, it is convenient to represent them by a three layers model. Thus, a fuzzy inference system may be given by the sequential composition of the input layer, by the inference layer of the fuzzy rules and by the output layer.

The input layer has functionalities of connecting the input variables (coming in from outside) with the fuzzy inference system and also their fuzzifications through respective membership functions.

In the inference layer of the fuzzy rules or just inference layer, the input fuzzified variables are combined among them, according to defined rules, using as support the operations defined in the theory of fuzzy sets. The set resulting of the aggregation process is then defuzzified, resulting in the fuzzy inference system output. The aggregation process and the defuzzification process of the fuzzy set output are made by the output layer. It is important to observe, concerning to the output layer, that although it performs the two processes above described, it is also responsible for storing the membership functions of the output variables.

In the following sub-sections further details will be presented about how fuzzy inference systems can be represented through a three layers model.

2.1 Input Layer

As previously presented, the fuzzy inference system input has the purpose of connecting the inputs coming from the environment with the fuzzy system, as well as the fuzzification of those according to the membership functions associated to the fuzzy system.

The system inputs fuzzification has the purpose of determining the degree of each input related to the fuzzy sets associated to each input variable. To each input variable of the fuzzy system can be associated as many fuzzy sets as necessary. This way, given a fuzzy system with one only input and, to this input, associated with N functions, that is N fuzzy sets which define that input, the output of the input layer will be a column vector with N elements representing the degrees of the input membership in relation to those fuzzy sets.

If we define the input of this fuzzy system with one only input, by the scaling of *x*, then the input layer output of the fuzzy system will be the vector ***y1***, that is:

$$\boldsymbol{y1}(x) = \left[p_1(x) \quad p_2(x) \quad \cdots \quad p_N(x) \right]^T \tag{1}$$

where $p_k(.)$ is the defined membership function for the x input, referring to the k-th fuzzy set associated to this input.

The generalization of the input layer concept for a fuzzy system owning m input variables is achieved if each input of this fuzzy system is modeled as a sub-layer of the input layer.

Thus, in equation (2) x is the i-th input of the fuzzy system, $\boldsymbol{p_k}(.)$ is the k-th vector of the membership functions associated to the x_k input and to the $\boldsymbol{y}_k$ vector. Each sub-layer has its own fuzzy sets defined by the fuzzy inference functions vector $p_k(.)$.

The output vector of the input layer $\boldsymbol{Y}(\boldsymbol{x})$ may be given as presented in (2), that is:

$$\boldsymbol{Y}(\boldsymbol{x}) = \begin{bmatrix} \boldsymbol{y1} \\ \boldsymbol{y2} \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} \boldsymbol{p1}(x_1) \\ \boldsymbol{p2}(x_2) \\ \vdots \\ p_m(x_m) \end{bmatrix} \tag{2}$$

There are several membership functions that can be defined. One of the necessary requisites for those functions is that they must be normalized in closed domain [0,1]. However, it is convenient that the membership functions are defined in a simple and convenient way, aiming their computational implementation, with objective of a higher processing speed and rational use the memory.

Besides providing benefits from the computational point of view, it is convenient that the membership functions may have a reduced number of free parameters in order that the tuning algorithm performs the tuning task in an adequate way.

2.2 Inference Layer

The inference layer of a fuzzy system has the functionality of processing the fuzzy inference rules defined for it. Another functionality of the inference layer is to provide a knowledge base for the process. The inference rules are processed in parallel, the same way as the sub-layers of the input layers are.

Inside this context, this set of rules has fundamental importance to the correct functioning of the fuzzy inference system. There are several methods for the extraction of fuzzy rules from the tuning set.

In this article, initially, the fuzzy inference system has all the possible inferred rules. Therefore, the tuning algorithm has the task of weighting the inference rules.

Weighting the inference rules is an adequate way to represent the most important rules in the fuzzy system, or even to allow that conflicting rules are related to each other without any verbal completeness loss.

Thus, it is possible to express the i-th fuzzy rule as in (3), that is:

$$R_i(\boldsymbol{Y}(\boldsymbol{x})) = w_i r_i(\boldsymbol{Y}(\boldsymbol{x})) \tag{3}$$

where $R_i(.)$ is the function representing the fuzzy weight value of the i-th fuzzy rule, w_i is the weight factor of the i-th fuzzy rule and $r_i(.)$ represents the fuzzy value of the i-th fuzzy rule.

2.3 Output layer

The output layer of the fuzzy inference system aims to aggregate the inference rules, as well as the defuzzification of the fuzzy set generated by the aggregation of inference rules.

In the fuzzy inference systems design, the choice of not only the aggregation method but also the defuzzification method constitutes a very important decision. The aggregation method of the fuzzy inference rules must be in such a way that the fuzzy set resulting from aggregation is capable of adequately representing the knowledge explicited by this set of fuzzy rules. By analogy, the method chosen for the defuzzification must be capable of expressing, in a crisp value, the fuzzy set resulting from the fuzzy aggregation.

Besides the operational aspects, the aggregation and defuzzification methods must attend the requisites of computational performance in order to reduce the computational effort needed in the fuzzy system processing.

In this paper, the output layer of the inference system is also adjusted. The adjustment of this layer occurs in a similar way to what occurs with the input layer of the fuzzy system.

3 Adjustment of the Fuzzy Inference System

The formalization of a fuzzy inference system in the form of a multilayer system, as presented in Section 2, can be justified not only by the different operational division of each one of these layers, but also by the presence in each of the different free parameters.

For example, let a fuzzy system with two inputs, each one with three gaussian membership functions, with a total of five inference rules, and having an output defined with two gaussian membership functions. It is known that, for each gaussian membership function, two free parameters exist: the mean and the standard deviation. This way, the number of free parameters of the input layer will be 12. For each inference rule has been associated a weighting factor, so, there are five free parameters in the inference layer. In relation the to output layer, the same considerations made for the input layer are valid. Therefore, four free parameters are in the output layer.

This way, the mapping f between the input space $\boldsymbol{x}$ and the output space $\boldsymbol{y}$ may be defined as in (4)

$$\boldsymbol{y} = f(\boldsymbol{x}, \boldsymbol{w1}, \boldsymbol{w2}, \boldsymbol{w3}) \tag{4}$$

where $\boldsymbol{w1}$, $\boldsymbol{w2}$ and $\boldsymbol{w3}$ respectively represent the vectors of the input membership functions parameters, the weight of the inference rules and the output membership functions parameters.

The definition of the energy function to be minimized remains in function of the fuzzy mapping. Considering that the tuning set $\{\boldsymbol{x}, \boldsymbol{d}\}$ is fixed during the whole adjustment process, it may be written:

$$\xi_{(x,y)} = \xi_{(x,y)}(\boldsymbol{w1}, \boldsymbol{w2}, \boldsymbol{w3}) \tag{5}$$

where ξ represents the energy function associated to the fuzzy inference system f.

$$\boldsymbol{y} = f(\boldsymbol{x}, \boldsymbol{w1^*}, \boldsymbol{w2^*}, \boldsymbol{w3^*}) \tag{6}$$

where $\boldsymbol{w1}^*$, $\boldsymbol{w2}^*$ and $\boldsymbol{w3}^*$ are the free parameters values of the fuzzy inference systems after the adjustment process.

3.1 Unconstrained Optimization Techniques

Taken an energy function $\xi_{(x,y)}(\boldsymbol{w1}, \boldsymbol{w2}, \boldsymbol{w3})$ in the form previously defined for a fuzzy inference system, in the assumption that a function might be differentiable in relation to a determined interval by the $\boldsymbol{w1}$, $\boldsymbol{w2}$ and $\boldsymbol{w3}$ vectors, that is, differentiable in relation to free parameters of the fuzzy inference system. Therefore, it is desired to find an optimum solution that may fulfill the following conditions:

$$\xi(\boldsymbol{w}_1^*, \boldsymbol{w}_2^*, \boldsymbol{w}_3^*) \leq \xi(\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3) \tag{7}$$

With the finality of simplifying the notation, these three vectors may be represented by one unique vector, resulting from the vector concatenation of $\boldsymbol{w1}$, $\boldsymbol{w2}$ and $\boldsymbol{w3}$, that is:

$$\boldsymbol{w} = \begin{bmatrix} \boldsymbol{w1}^T & \boldsymbol{w2}^T & \boldsymbol{w3}^T \end{bmatrix}^T \tag{8}$$

Thus, the expression (7) may be rewritten by:

$$\xi(\boldsymbol{w}^*) \leq \xi(\boldsymbol{w}) \tag{9}$$

Therefore, it may be observed that, to attend the condition expressed in (9), it is necessary to solve an unconstrained optimization problem. Then, the $\boldsymbol{w}$ vector may be obtained by the following equation:

$$\boldsymbol{w}^* \stackrel{\Delta}{=} \arg \min_{w} \xi(\boldsymbol{w}) \tag{10}$$

The condition that expresses the optimum solution in (10) must attend to the following solution:

$$\nabla \xi(\boldsymbol{w}^*) = \boldsymbol{0} \tag{11}$$

where ∇ is the gradient operator, that is:

$$\nabla \xi(\boldsymbol{w}) = \left[\frac{\partial \xi}{\partial w_1}, \frac{\partial \xi}{\partial w_2}, \cdots, \frac{\partial \xi}{\partial w_m} \right]^T \tag{12}$$

In problems like this, involving the minimization of energy functions, it is desired that to each iteration, the energy function value would be less than the energy function value of the previous iteration.

There are several techniques used in solving the unconstrained optimization problems. A detailed description of the unconstrained optimization techniques may be found in [7].

The choice of the most adequate technique to be used is conditioned to the form by which the energy function is defined. For example, the Gauss-Newton method for the unconstrained optimization may be more applicable in problems where the energy function is defined as:

$$\xi(\boldsymbol{w}) = \frac{1}{2} \sum_{i=1}^{n} e^2(i) \tag{13}$$

where $e(i)$ is the absolute error in relation to the i-th tuning pattern.

In this paper, a derivation of the Gauss-Newton method is used for the fuzzy inference system tuning. The optimization algorithm used was the Levenberg-Marquardt method [8]. The calculation of differential equations was performed with the help of the finite differences method.

4 Mackey-Glass Chaotic Time Series

The research around chaotic series created new paradigms about the existent modeling techniques. In this way, the present research appeals to new fundaments for the series prediction. On the other hand, the determinism inherent to chaos implies that many phenomenons, formerly seen as random, may be treated in a predictive way.

The prediction of the Markey-Glass chaotic time series is a classic estimation problem. Generally, this problem is used to test the generalization capacity of such as systems coming from the computational intelligence, like neural networks and fuzzy inference systems. The dynamic properties of Mackey-Glass time series are rich in complexity. The Mackey-Glass differential equation may be expressed by:

$$\frac{dx(t)}{dt} = -bx(t) + \frac{ax(t-\tau)}{1 + x(t-\tau)^c} \qquad (14)$$

The Mackey-Glass time series was one of the first models for the time quantization of producing white cells in the human organism [9]. In general, and in this work too, the values of the constants in (14) are adopted as being $a = 0.2$, $b = 0.1$ and $c = 10$. The value for the delay constant τ is chosen as being 17.

The Mackey-Glass time series may be obtained by the integration of the equation in (14). More specifically, it has been used the second order Runge-Kutta method with integration step equal to 0.1. The result of this integration is shown in Figure 1.

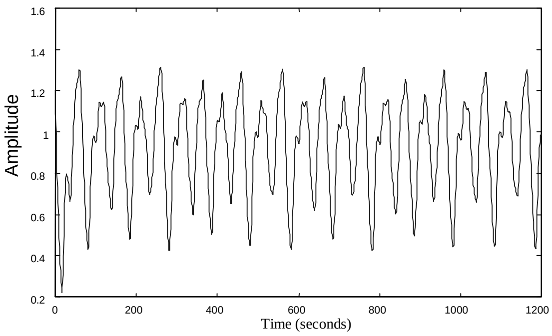

Fig. 1- Mackey-Glass time series.

5 Methodology and Results

Using the methodology presented in this paper for the fuzzy inference systems tuning and, based on the Mackey-Glass time series, a fuzzy inference system of Mandani type was developed with the objective to predict this series. The tuning set was constituted by 500 patterns. The input variables of the fuzzy inference system were four, corresponding to values $x(t-18)$, $x(t-12)$, $x(t-6)$ and $x(t)$. As an output variable was adopted $x(t+6)$.

The fuzzy inference system was defined having four fuzzy sets attributed to each input variable and also to the output variable. A total of 64 inference rules have been used in the inference process.

The energy function of the system was defined as being the mean squared error between the desired values $x(t+6)$ and the values $\bar{x}(t+6)$

$$\xi_{(\boldsymbol{w1,w2,w3})} = \frac{1}{2}\sum_{i=1}^{L}\left[x_i(t+6) - \bar{x}_i(t+6)\right]^2 \qquad (14)$$

where L is the number of data used in the tuning process (L=500).

After minimization of (14), the membership functions of the fuzzy inference system were adjusted as illustrated in Figures from 2 to 5.

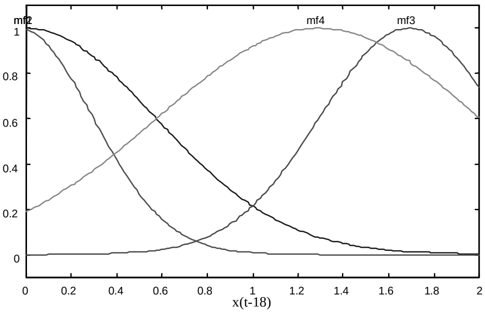

Fig. 2 - Input membership functions to $x(t-18)$.

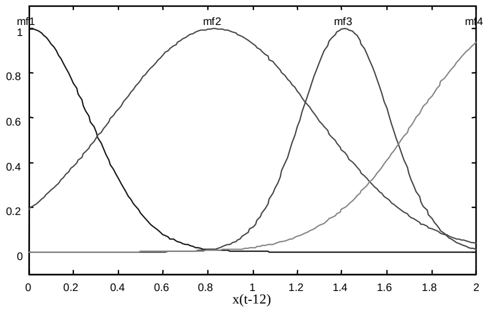

Fig. 3- Input membership functions to $x(t-12)$.

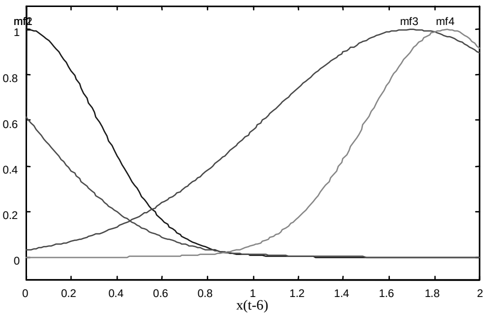

Fig. 4 - Input membership functions to $x(t-6)$.

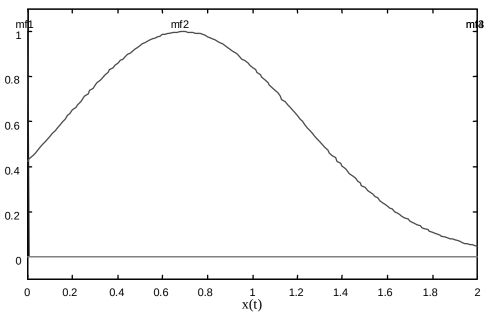

Fig. 5 - Input membership functions to $x(t)$.

In Figure 6 is presented the output membership functions of the fuzzy inference system.

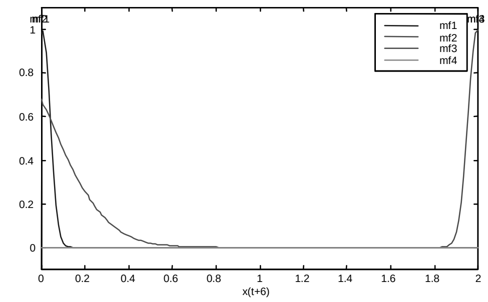

Fig. 6 - Output membership functions $x(t+6)$.

In Figure 7 is presented the result of prediction provided by the fuzzy inference system for 1000 points.

The mean squared error of estimation for the proposed problem was 0.000598 with standard deviation of 0.02448.

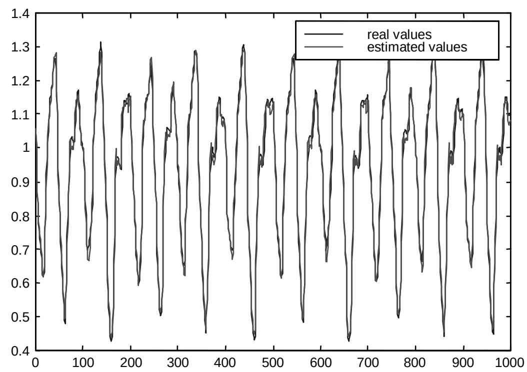

Fig. 7 - Estimation of the fuzzy inference system for the Mackey-Glass series.

For comparison, it was developed a fuzzy inference system tuning with ANFIS (Adaptive Neural-Fuzzy Inference System) method. This fuzzy inference system was made with 10 membership functions for each input, being the knowledge base constituted by 10 rules. The mean squared error of estimation for the proposed problem was 0.000165 with standard deviation of 0.0041.

6 Conclusions

In this paper was underlined the basic foundations around the fuzzy inference systems tuning process, from the unconstrained optimization techniques.

In order that the tuning may be efficient it is necessary that the energy function is perfectly underlined for the adjustment process.

For the validation of the proposed methodology, it was studied the estimation of the Mackey-Glass chaotic time series. The comparison of the results was made with those results provided from the ANFIS methodology. Although the results provided by the ANFIS methodology were better, the amount of fuzzy sets associated to each input was superior to the number of fuzzy sets used in the fuzzy system developed along this work.

This approach offers new perspectives of research related to the fuzzy inference systems, allowing thus that problems previously treated only with the help of artificial neural networks may now be treated through fuzzy inference systems.

As future works, it is intended to develop efficient techniques for the construction of inference rules bases with the objective of optimizing their structures in the whole.

References:

[1] M. Figueiredo and F. Gomide, Adaptive Neuro Fuzzy Modeling, *Proceedings of the Sixth IEEE International Conference on Fuzzy Systems*, Vol. 3, 1997, pp. 1567-1572.

[2] S. Guillaume, Designing Fuzzy Inference Systems from Data: Na Interpretability - Oriented Review, *IEEE Trans. Fuzzy Systems*, Vol. 9, No. 3, 2001.

[3] R. Kamimura, T. Takagi and S. Nakanishi, Improving Generalization Performance by Information Minimization, *IEEE World Congress on Computational Intelligence*, Vol. 1, 1994, pp. 143-148.

[4] S. Becker, Unsupervised Learning Procedures for Neural Networks, *International Jornal of Neural Systems*, Vol. 2, 1991, pp. 17-33.

[5] W. Wan, K. Hirasawa, J. Hu and J. Murata, Relation Between Weight Initialization of Neural Networks and Pruning Algorithms: Case Study on Mackey-Glass Time Series, *International Joint Conference on Neural Networks*, Vol. 3, 2001, pp. 1750-1755.

[6] J. R. Jang, ANFIS: Adaptive-Network-Based Fuzzy Inference System, *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 23, 1993, pp. 665-685.

[7] D. P. Bertsekas, *Nonlinear Programming*, Athenas Cientific, 1995.

[8] D. Marquardt, An Algorithm for Least Squares Estimamation of NonLinear Parameters, *J. Soc. Ind. Appl. Math*, 1963, pp. 431-441.

[9] M. C. Mackey and L. Glass, Oscillations and Chaos in Physiological Control Sciences, *Science*, pp. 287-289, 1997.