2003 International Symposium on Advanced Intelligent Systems
September 25-28, 2003, Jeju, Korea

Adaptive Fuzzy Inference System using Pruning Techniques

Chang-Hyun Kim, Byoung-Gi Jang and Ju-Jang Lee
Department of Electrical Engineering and Computer Science
Korea Advanced Institute of Science and Technology
373-1, Guseong-dong, Yuseong-gu, Daejeon, 305-701, Korea
Email: {sunnine, ee185zzang}@odyssey.kaist.ac.kr, jjlee@ee.kaist.ac.kr
Telephone: +82-42-869-8032, Fax: +82-42-869-5432

Abstract— **Fuzzy modelling has the approximation property for the given input-output relationship. Especially, Takagi-Sugeno fuzzy models are widely used because they show very good performance in the nonlinear function approximation problem. But generally there is not the systematic method incorporating the human expert's knowledge or experience in fuzzy rules and it is not easy to £nd the membership function of fuzzy rule to minimize the output error as well. The ANFIS (Adaptive Network-based Fuzzy Inference Systems) is one of the neural network based fuzzy modelling methods that can be used with various type of fuzzy rules. But in this model, it is the problem to £nd the optimum number of fuzzy rules in fuzzy model. In this paper, a new fuzzy modelling method based on the ANFIS and pruning techniques with the measure named impact factor is proposed and the performance of proposed method is evaluated with several simulation results.**

I. Introduction

The main purpose of fuzzy modelling is to achieve a set of local input-output relations that describe a process. As is well known, the problem of system modelling requires two main stages: structure identi£cation and parameter optimization. Structure identi£cation deals with the problem of determining the input-output space partition and, in particular how many rules must be used by the fuzzy system. Parameter optimization is in charge of £nding the optimum values of all the parameters involved in the fuzzy system, i.e., location the membership function (MFs) in the premise of each rule and its consequent.

There have been various approaches to optimize the fuzzy systems. Most of them considers a £xed topology of the fuzzy systems and then optimize the set of parameters under that unchangeable structure. From early works [2], [3] in which the fuzzy controller parameters were determined by trial and error through the works of Wang and Mendel [4] to the more sophisticated approaches [5], [6], [7] based on steepest descent using neuro-fuzzy systems, all of them were based on an a *a priori* £xed fuzzy system topology.

The fact that parameter adjustment has been focused on than system identi£cation is understandable since the latter is a very complex task for which it is very dif£cult to obtain reliable procedures and it is not possible to test system identi£cation without parameter adjustment. The £rst relevant approach to tackle both tasks was TS model [1]. TS model is a particular multiple model de£ned as a set of fuzzy rules of the form "*If* premise *then* af£ne local behavior". The global model is obtained by combining the local models using appropriate weighting functions.

Since then, several different approaches to fuzzy modelling, most of them based on clustering techniques [8], [9], [10] have been proposed. Rules are generated based on the each cluster divided in these methods. There were also many methods using the genetic algorithm [11] or the orthogonal transformation such as SVD, QR decomposition [12] or the proposed criterion [13].

The ANFIS [7] is one of the fuzzy modelling methods focusing on parameter identi£cation and show good performance. But it is still a problem to £nd the optimum number of fuzzy rules in fuzzy model. In this paper, a new pruning technique with appropriate measure named impact factor (IMF) is applied to the ANFIS. The impact factor based on the variance of the MFs and used as the index of the importance of rules. Based on these measure, the pruning technique is applied to remove the useless rules.

The paper is organized as follows: the mathematical formulation of the ANFIS is presented in section II. The IMF and pruning algorithm proposed is described in section III and some simulation studies are reported in section IV. The last section concludes the paper.

II. ANFIS

The ANFIS is one of the methods to organize the fuzzy inference system with given input-output data pairs. It can be applied to general fuzzy models and used widely because of its performance. The ANFIS optimizes the parameters of consequent part using least square method and those of premise part using steepest descent method.

Next, the fuzzy model considered in this paper and the detailed optimization method will be described. We consider the Takagi-Sugeno fuzzy model and assume the case having two input x, y and one output z for the

simplicity. The TS model has the following two rules.
Rule 1: If x is A_1 and y is B_1, then $f_1 = p_1x+q_1y+r_1$
Rule 2: If x is A_2 and y is B_2, then $f_2 = p_2x+q_2y+r_2$

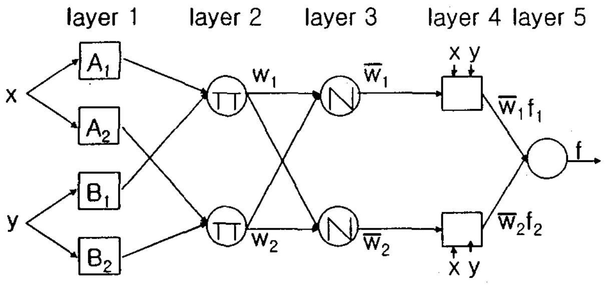

Fig. 1. ANFIS model

The corresponding ANFIS architecture is shown in Fig. 1. A square node (adaptive node) has parameters while a circle node (£xed node) has none. The node functions in the same layer are of the same function family as described below:

Layer 1: Every node i in this layer is an adaptive node with node function

$$O_i^1 = \mu_{A_i}(x) \tag{1}$$

where x is the input node i, and A_i is the linguistic label (*small, large*, etc.). In other words, O_i^1 is the membership function of A_i and it speci£es the degree to which the given x satis£es the quanti£er A_i. We use following bell-shaped function,

$$\mu_{A_i}(x) = exp\{-(\frac{x - c_i}{a_i})^2\} \tag{2}$$

where $\{a_i, c_i\}$ is the parameter set. Parameters in this layer are referred to as *premise parameters.*

Layer 2: Every node in this layer is a £xed node labeled $\prod$ which multiplies the incoming signals and sends the product out. For instance,

$$w_i = \mu_{A_i}(x) \times \mu_{B_i}(y), i = 1, 2 \tag{3}$$

Each node output represents the £ring strength of a rule.

Layer 3: Every node in this layer is a £xed node labeled *N*. The ith node calculates the ratio of the ith rule's £ring strength to the sum of all rules' £ring strengths:

$$\overline{w}_i = \frac{w_i}{w_1 + w_2} \tag{4}$$

For convenience, outputs of this layer will be called *normalized £ring strengths.*

Layer 4: Every node i in this layer is a adaptive node with a node function

$$O_i^4 = \overline{w}_i f_i = \overline{w}_i(p_i x + q_i y + r_i) \tag{5}$$

where $\{p_i, q_i, r_i\}$ is the parameter set. Parameters in this layer will be referred to as *consequent parameters.*

Layer 5: The single node in this layer is a £xed node labeled $\sum$ that computes the overall output as the summation of all incoming signals, i.e.,

$$O_1^5 = overall\ output = \sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \tag{6}$$

The premise and the consequent parameters can be chosen to minimize the following sum of squared error.

$$E = \sum_{m=1}^{N} (T_m - O_m)^2 \tag{7}$$

where T_m is the desired output of mth data and O_m is the output of fuzzy model using mth data and N is the total number of training data set. The steepest descent method as in neural network can applied to £nd the premise parameters and least square estimate can be applied to optimize the consequent parameters. Each step is executed iteratively with another parameters £xed.

III. Pruning Algorithm

To examine the £ring strength of the rule is helpful to determine whether the rule is useful or not. For example, if one rule always has the zero £ring strength to all data, then that rule is useless. Besides, if one rule has almost same £ring strength to all data, that is, the variance of £ring strength is small, then that rule is replaced by bias terms or parameters. Considering these effects, we can remove useless rules based on the measure driven by the variance. The impact factor (IMF) of rules is de£ned as:

$$IMF_i = \frac{1}{N_i} \sum_{m \in R_i} (O_i^4 - \overline{O_i^4})^2 \tag{8}$$

where N_i is the number of data which has an effect on ith rule and $\overline{O_i^4}$ is mean value of ith output in layer 4. This value is used as the index of usefulness of a rule. At each iteration in learning process, the rule with smaller IMF value than the given threshold is removed. And series of these pruning process may cause that removal of MFs itself.

The pruning algorithm is shown in Fig. 2. First, we start with the fuzzy system having many MFs and the initial MFs are separated uniformly over the whole input range. At every iteration of learning procedure, after £nding the premise and consequent parameters, the IMF of each rule is calculated. If the IMF of one rule is smaller than prede£ned threshold τ, that rule is pruned

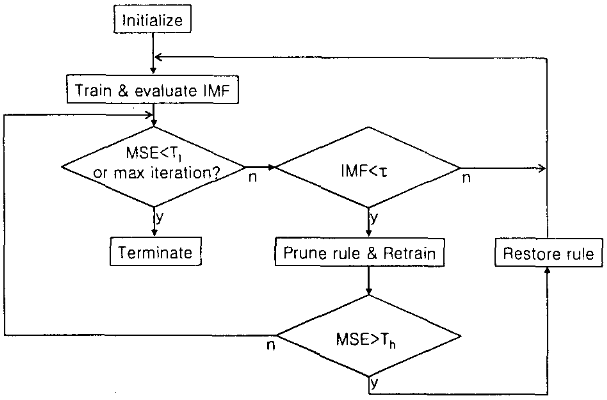

Fig. 2. The pruning algorithm

and the system is retrained. Else the process runs next iteration with no pruned rule. And if the mean squared error of the pruned system is smaller than the high threshold T_h, the structure of system unchanges. Else if the performance is unsatisfactory, the pruned rule is restored and next iteration will be executed. This process is continued until the mean squared error exceed the desired error bound T_l or the iteration reaches the maximum iteration.

IV. Simulation Results

A nonlinear dynamic system modelling is simulated to evaluate the performance of proposed method. The following second order system is given in [12], [14], [15].

$$y(k) = g(y(k-1), y(k-2)) + u(k) \tag{9}$$

where $g(y(k-1), y(k-2)) = \frac{y(k-1)y(k-2)(y(k-1)-0.5)}{1+y^2(k-1)+y^2(k-2)}$.
The objective is to approximate the function $g(y(k-1), y(k-2))$. This problem is 2-input and 1-output case and 400 data set is generated. 200 training data are grouped among them and the other are used in validation. The data set is given as follows: the process starts from initial state $(0, 0)$. The £rst 200 (training) points are generated using the input randomly distributed in the range of $[-1.5, 1.5]$ and the next 200 (validation) points were generated using the input varying with the sinusoidal function, $u(k) = sin(2\pi k/25)$.

At £rst, each input have 5 MFs, that is, total number of rules is 25(=5×5). After pruning process, the fuzzy model having 4(2×2) rules is obtained. The output of the plant and the £nal fuzzy model are shown in Fig. 3. The output error between the two models is plotted in Fig. 4. And the obtained MFs of two input are shown in Fig. 5 and Fig. 6 respectively.

Some comparison results with other methods are shown in Table I. The proposed method £nds the fuzzy model has as small number of rules as the other methods and the performance is comparable. And it is noted that

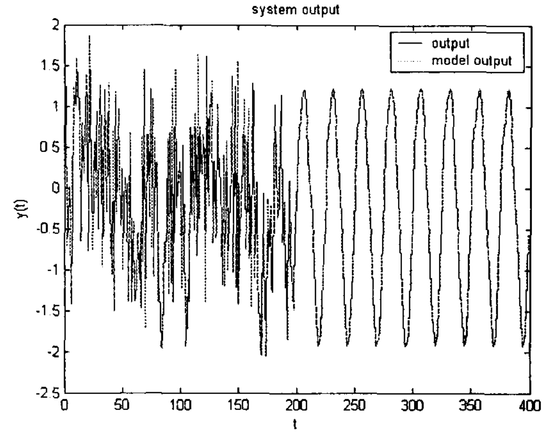

Fig. 3. The output of the plant and the £nal model

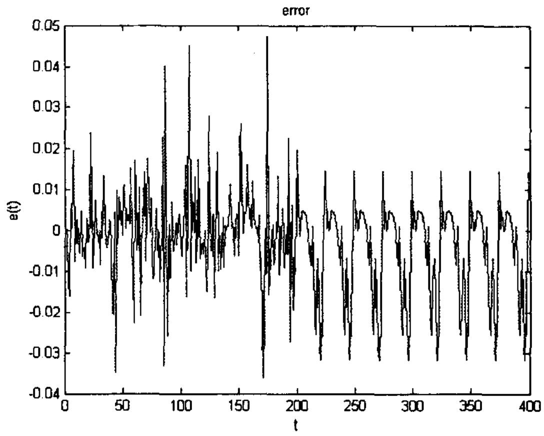

Fig. 4. Output error

the £nal model is dependent on the threshold and error bound. The smaller model is obtained in the case with larger threshold and larger error bound as expected.

V. Conclusion

In this paper, the pruning algorithm based on the IMF is proposed. We use the Takagi-Sugeno fuzzy model and neuro-fuzzy network based learning(ANFIS). Starting from the model has large size, after £nding the premise and consequent parameters using steepest descent and least square method respectively the model is pruned based on the IMF. And these loop is repeated until the MSE is worse than desired threshold. Simulation results show the usefulness of the proposed method. But the £nal structure depends on the simulation parameters and it is still problem to adjust these values. There is a trade-off between accuracy and compactness of model and more analysis is required.

References

[1] T. Takagi and M. Sugeno, Fuzzy identi£cation of systems and its application to modeling and control. *IEEE Trans. Syst., Man, Cybern.*, Vol. 15, pp. 116-132, 1985.

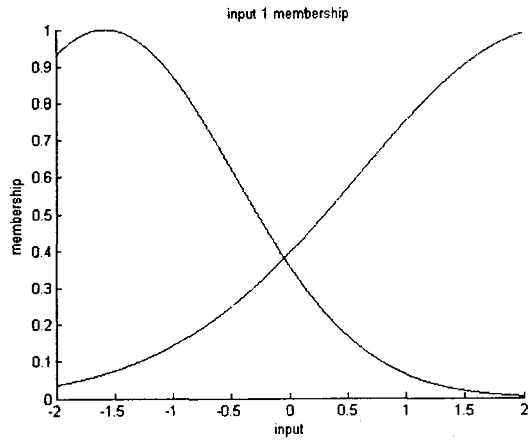

Fig. 5. Final membership functions: x

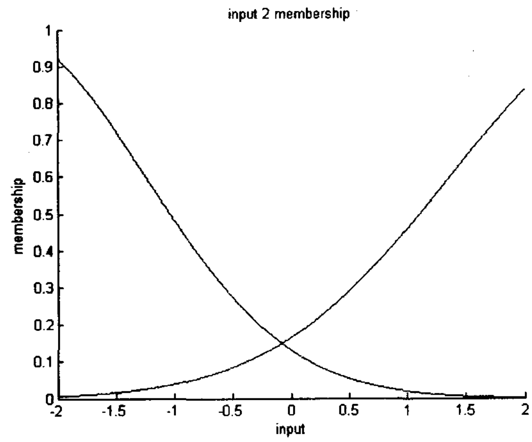

Fig. 6. Final membership functions: y

[2] E. H. Mamdani and S. Assilian, An experiment in linguistic synthesis with a fuzzy logic controller, *Int. J. Man-Machine Studies*, Vol. 7, no.1, pp. 1-12, 1975.

[3] R. M. Tong, The construction and evaluation of fuzzy models, *Advances in Fuzzy Set Theory and Applications*, M. M. Gupta, R. K. Ragade, and R. R. Yager, Eds. Amsterdam, The Netherlands: North-Holland, pp. 559-576, 1979.

[4] L. X. Wang and J. M. Mendel, Generating fuzzy rules by learning from examples, *IEEE Trans. Syst., Man, Cybern.*, Vol. 22, pp. 1414-1427, 1992.

[5] H. Nomura, I. Hayashi, and N. Wakami, A learning method of fuzzy inference rules by descent method,*Proc. IEEE Int. Conf. Fuzzy Systems*, pp. 203-210, 1992.

[6] S. Horikawa, T. Furuhashi, and Y. Uchikawa, On fuzzy modeling using fuzzy neural networks with the back-propagation algorithm, *IEEE Trans. Neural Networks*, Vol. 3, pp. 801-806, 1992.

[7] J. S. R. Jang, ANFIS: Adaptive-network-based fuzzy inference systems,*IEEE Trans. Syst., Man, Cybern.*, Vol. 23, pp. 665-685, 1993.

[8] M. Sugeno and T. Yasukawa, A fuzzy-logic based approach to qualitative modeling, *IEEE Trans. Fuzzy Syst.*, Vol. 1, Feb. 1993.

[9] W. Pedrycz, Conditional fuzzy c-means, *Patt. Recog.*, Vol. 17, pp. 625-631, 1996.

[10] M. L. Hadjili and V. Wertz, Takagi-Sugeno Fuzzy Modeling Incorporating Input Variable Selection, *IEEE Trans. Fuzzy Syst.*, Vol. 10, no.6, pp. 728-742, 2002.

[11] L. Wang and J. Yen, Extracting fuzzy rules for system modelling

TABLE I

THE COMPARISON RESULT

Method	#(rule)	MSE(tr.)	MSE(va.)
Yen *et al.*[14]	24	$2.0e^{-6}$	$6.4e^{-6}$
Yen *et al.*[12]	25	$2.3e^{-4}$	$4.1e^{-4}$
	20	$6.8e^{-4}$	$2.4e^{-4}$
Setnes *et al.*[15]	5	$5.8e^{-3}$	$2.5e^{-3}$
	5	$7.5e^{-4}$	$3.5e^{-4}$
	4	$1.2e^{-3}$	$4.7e^{-4}$
Proposed	4	$1.4e^{-4}$	$1.8e^{-4}$

using a hybrid of genertic algorithms and Kalman £lter, *Fuzzy Sets and Syst.*, Vol. 101, pp 353-362, 1999

[12] J. Yen and L. Wang, Simplifying fuzzy rule-base models using orthogonal transforamtion method, *IEEE Trans. Syst. Man. Cyb.*, Vol 29, No. 1, pp 13-24, 1999.

[13] K. Gasso, G. Mourot and J. Ragot, Fuzzy rule base optimisation: a pruning and merging approach, *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Vol. 1, pp. 67-72, 2000.

[14] J. Yen and L. Wang, Application of statistical information criteria for optimal fuzzy construction, *IEEE Trans. Fuzzy Syst.*, Vol. 6, pp. 362-371, 1998.

[15] M. Setnes and H. Roubos, GA-fuzzy modeling and classi£cation: complexity and performance, *IEEE Trans. Fuzzy Syst.*, Vol.8, pp. 509-522, 2000.