A Connectionist Approach to Reorganization of Knowledge Structures in Fuzzy Systems

Ali Arya
Dept. of Automation and Control, Fariné Co.
P.O. Box: 14335-996, Tehran, Iran.

Kambiz Badie
Iran Telecom. Research Center ,and
Intelligent Research Faculty, Tehran, Iran.

Abstract

This paper addresses the problem of reorganizing knowledge structures in fuzzy systems. Reviewing the most accepted approaches to this problem, it is shown that capabilities such as adaptability and determination of linguistic labels and Membership Functions (MFs) are problems which are still in the offing. Within our framework, we propose an approach to modeling issues in fuzzy systems which is extendible to more complex knowledge structures, compared to simple rule bases.

1. Introduction

Adaptive definition, acquisition, and arrangement of information, in terms of numeric data or qualitative concepts, which is called reorganization of knowledge structures here, is the basic requirement of engineering, and even non-engineering, decision-making systems. In case of fuzzy systems, which are widely accepted as powerful decision-making tools, these knowledge structures are, primarily, membership functions (including number and position of fuzzy labels) and rule bases. In this case, development of an efficient framework for reorganization of knowledge structures, leads to an adaptive fuzzy system with no, or minimum, presumption regarding membership functions and rules. Due to its capabilities in handling complex modeling and control problems, such a system is, increasingly, studied in recent literature [2,3,6,10].

Neuro-fuzzy or fuzzy-neural, along with other approaches have been, recently, used to resolve the reorganization (adaptation) problems of fuzzy systems. Although these approaches are successful in some aspects, however their knowledge structures can not, necessarily, be reorganized such that adaptive capabilities which are essential to a fuzzy system can be covered. These capabilities are, mainly, "adaptability of the concepts and parameters (e.g. rules and weights)", "determination of the linguistic variables values (fuzzy labels)", "definition of new membership function", and "selection of the system variables".

In this paper, we have proposed a connectionist approach to reorganize the knowledge structures in a fuzzy system. This proposal is based on applying a self-organizing neural net works as a clustering tool to determine the linguistic labels of each linguistic variable, and establishes the significant associations between the input and output labels which leads to rule generation. Another function of this network is to process the excitatory/inhibitory interactions of the nodes which stand for the different fuzzy labels of a similar variable. In this way, some kind of knowledge is embedded in the network which is essential to the definition of membership functions. On the other hand, the concept of utility weight vector [1] is used as the measure of validity in a regulatory mechanism which enhances the performance of network. The next section reviews some existing approaches to adaptive fuzzy systems. Based on analyzing the advantages and drawbacks of these schemes, our proposed approach is discussed in section III. After some simulation results and evaluation of the capabilities in section IV, a few concluding remarks are presented in section V.

2. Some Existing Approaches to Adaptive Fuzzy Systems

Due to lack of sufficient expert knowledge in most applications, designing automatic mechanisms for rule generation and membership function definition is a topic of interest for fuzzy systems researchers. Mamdani [7] has introduced one of the first approaches to an automatic (or adaptive) system for control purposes. He has used the controller error, plant inverse model and a performance table to change adaptively the rule base of an ordinary fuzzy logic controller. Pedrycz [8] and some other researchers have applied the concept of fuzzy relational equations to modeling and identification problems.

Rapid development of neural networks and evolutionary computing, offers attractive ideas to be used

0-7803-3796-4/97/$10.00©1997IEEE

for adaptation and learning capabilities in fuzzy systems. One of the most notable works in this regard, by Kong and Kosko [6], is based on the concept of Fuzzy Associative Memory (FAM). Each FAM rule defines a patch in the input-output state space. The unsupervised Differential Competitive Learning (DCL) and product-space clustering adaptively generate these rules using the training data. Although widely affective and applicable, this approach proposes no way for defining membership functions and selecting fuzzy labels of each variable. In addition, its adaptation is limited to a training interval. On the other hand, the basic idea of using a clustering network is the key element of many other methods, suggested later.

Sugeno and Yasukawa [9] proposed a method for qualitative modeling, which uses off-line clustering of input-output data. This method, efficiently, determines the number and distribution of fuzzy labels and approximates the trapezoidal membership functions. The results are finally translated to a linguistic form to be used as the qualitative model of the system. The main disadvantage of this method is that all processes are off-line and there is no formalism for adaptation and on-line learning. Also, the algorithms proposed for structure identification and parameter estimation are relatively complex and time-consuming.

Among many researchers who worked on neural network implementation of fuzzy systems, Jang [3] proposed Adaptive-Network-Based Fuzzy Inference System (ANFIS), which may be considered as a typical scheme in this regard. The fuzzy inference system is constructed through a multi-layer network, with each layer corresponding to one of the functions of fuzzy system, i.e. fuzzification (membership function calculations), decision-making (inference operations on the rules), weighted summation, and defuzzification. Some algorithms are provided for learning the parameters in each layer to guaranty the adaptability of the system. ANFIS supports most of the requirements of an adaptive fuzzy system and may be extended to include fuzzy modeling tasks. But the fuzzification process is too constrained (regarding the number of fuzzy labels and membership functions) and the learning algorithms are complex, time-consuming, and unreliable to generate an optimized inference system, regarding the rules and the MFs. Many other researchers [10] have proposed neural network driven fuzzy systems, aimed to improve the above approach. Although successful in some regards, none of them may be considered as a comprehensive framework for adaptive fuzzy modeling.

Genetic Algorithms establish another group of approaches to adaptive fuzzy systems. In these works [2,5], the parameters to be modified (like the MFs geometrical parameters and rules weights) construct bit arrays. Applying genetic operators on these arrays, and using some fitness functions, result in optimization of the parameters. The need for a fitness function and the plant model to test the generations, are the main drawbacks that make these approaches unsuitable for on-line applications.

3. The Proposed Approach

3.1. Basic Scheme

For the sake of simplicity, without loss of generality, we have focused our attention to a single-input/single-output fuzzy modeling system. Fuzzy modeling is an effective approach for modeling and identification purposes, in the cases where the systems under study are uncertain and preferably of a fuzzy nature [11]. According to this approach, system behavior, unlike the conventional models, is to be modeled by means of qualitative /linguistic rules (Figure 1).

Needless to say, for other types of fuzzy systems, like

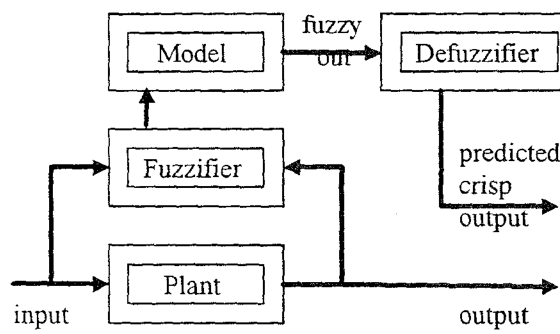

Figure 1. Fuzzy Modeling System

fuzzy controllers, basic structure of the system is the same. The only necessary changes may be, for instance, replacing the plant input/output with the plant command/action, so the system generates the fuzzy and crisp commands instead of predictions.

The existing methods in fuzzy modeling are mostly based on clustering the input/output data, and assigning the derived clusters to some appropriate fuzzy sets [6,9]. To suggest an adaptation mechanism for fuzzy systems, including fuzzy modeling, we note the following considerations:

1. An intrinsically adaptive neural net is the best choice to implement the inference system.
2. No presumption, regarding the fuzzy labels or MFs should be defined.

3. System overall performance is to be considered as an index in learning, in addition to instantaneous matching of the model parameters with plant state.
4. The system should be extendible to fuzzy knowledge structures other than simple fuzzy rule-based systems

With the above considerations, we present a fuzzy modeling system, as shown in Figure 2. The system is essentially constructed by a self-organizing neural net, which primarily serves to cluster the input-output data. It clusters the data to a sufficient number of classes and assign a membership degree for each variable with respect to each class. The input class nodes (X-group) are associated to the output class nodes (Y-group) via a rule definition matrix. Each association, activated in this way, stands for a model rule. The output of X-group is the current fuzzy state of the plant, and the output of Y-group is its predicted fuzzy state for the next sampling time. The defuzzifier unit (or layer) may translate this fuzzy state to a crisp value.

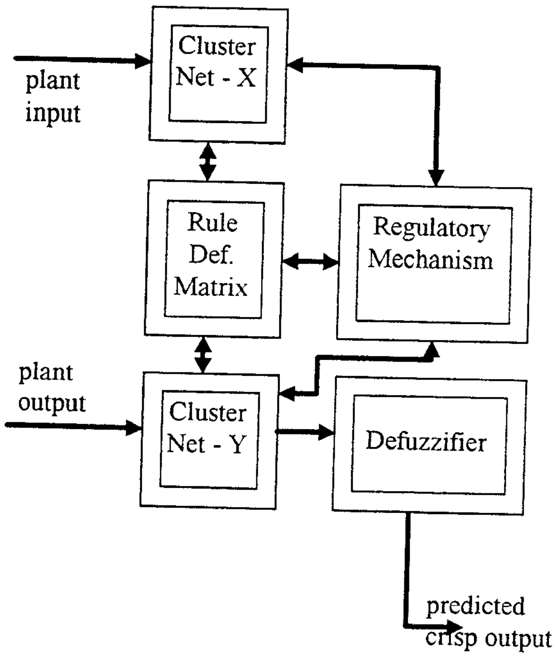

Figure 2. Proposed Adaptive Fuzzy Modeling System

A regulatory mechanism verifies the overall performance of the system and modifies or disactivates unnecessary or malfunctioning classes and associations. Activation of these inference elements is controlled by its Utility Weight Vector (UWV). UWV has been previously [1] introduced as an adaptation tool for fuzzy controllers. There, UWV has been defined as a two-component vector consisted of matching and performance degrees of a rule. Matching degree is related to the memberships of the IF part and the performance degree related to the effect of THEN part in controllery action and error. Here, matching degree is divided to X-group and Y-group matching degrees (d_{mx} and d_{my}) and the performance degree (d_p) represents the effect in the modeling error.

3.2. Neural Processing of Knowledge in Fuzzy System

Clustering SubNet

As mentioned above, the primary purpose of neural network, used in this proposal, is to cluster the input/output data. This clustering yields the required fuzzy labels (linguistic values) for each variable. Self organizing characteristic of the net, makes it possible to create sufficient number of classes, with appropriate distances. The clustering net, for both input and output data, is shown in Figure 3.

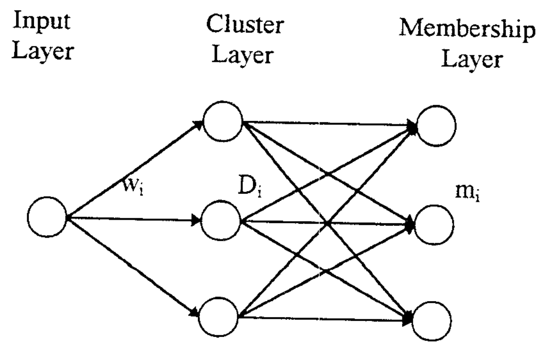

Figure 3. Clustering Subnet

This network is consisted of three layers: input, cluster, and membership layer. Data is passed through the input layer to the cluster layer, at which the classes are set. Then for each class, a membership value is calculated in the third layer. The Algorithm-1 describes the basic function of this net. This algorithm utilizes the ideas proposed by Huntsberger and Ajjimarangsee [4] to compute the membership degrees in a self-organizing network.

<u>Algorithm-1</u>

Step 1: Randomly initialize the weights w_i and set the weight updates Δw_i to zero. Select a value between 0 and 1 for the learning rate r_l.

Step 2: Get data in the input layer.

Step 3: Calculate distance.

$$D_i = (x - w_i)^T (x - w_i) \tag{1}$$

Step 4: Select the winner in cluster layer. If the distance value for this node is lower than a threshold, D_T, create a new cluster (node).

Step 5: Update the weights.

$$w_i = w_i + r_l \Delta w_i \tag{2}$$
$$\Delta w_i = m_i(x - w_i) \tag{3}$$

Step 6: Calculate the membership degrees.

if $D_i = 0$
then $m_i = 1$

if $D_l = 0$ (l≠i)
or $x < w_{j-1}$
or $x > w_{j+1}$
then $m_i = 0$

else

$$m_i = \left(\sum_{l=0}^{c-1}\left(\frac{D_i}{D_l}\right)\right)^{-1} \tag{4}$$

where c is the number of clusters (nodes in cluster and membership layer).

Rule Definition Matrix

The clustering subnet, described above, sets the necessary classes required for input and output data. Next step in the modeling procedure is to associate the input classes to the output classes in order to establish the behavioral rules of the system. We see that this approach differs from the existing approaches to neural-network-driven fuzzy systems in the way it defines the rules. In those implementations, there is an inference layer, with each neuron operating as a rule by achieving inference functions such as MIN operation. This may be considered as the Neuron-As-Rule strategy. Here, because of the nature of fuzzy modeling systems, and its need for qualitative descriptions in both sides of the rules, it is more efficient to have independent clustering layers for input and output variables, and the association between them as a rule. We call this, the Association-As-Rule strategy.

	Low	**Med.**	**High**
Low	W_{00}	W_{01}	W_{02}
Med.	W_{10}	W_{11}	W_{12}
High	W_{20}	W_{21}	W_{22}

Figure 4. Sample Rule Definition Matrix

To implement the Association-As-Rule strategy, we use a Rule Definition Matrix which connects the X-group fuzzy labels to the Y-group. Each element of this matrix is the dependency weight of the related Y-group label to a X-group label. Activation of a Y-group label in the prediction procedure is the MIN of its dependency weights. Figure 4 shows a typical Rule-Definition-Matrix for a single-input/single-output system with three fuzzy labels for each variable.

The elements of rule definition matrix are learned using UWVs for each of them. UWV is a three-component vector, defined and learned as follows (or by more complex functions if necessary):

$$d_{mxi}(k) = \frac{1}{k+1}\sum_{l=0}^{k} m_{xi}(l) = \frac{k}{k+1}d_{mxi}(k-1) + \frac{1}{k+1}m_{xi}(k) \tag{5}$$

$$d_{myj}(k) = \frac{1}{k+1}\sum_{l=0}^{k} m_{yj}(l) = \frac{k}{k+1}d_{myj}(k-1) + \frac{1}{k+1}m_{yj}(k) \tag{6}$$

$$d_{pij}(k) = \frac{1}{k+1}\sum_{l=0}^{k} m_{xi}(l)m_{yj}(l) = \frac{k}{k+1}d_{pij}(k-1) + \frac{1}{k+1}m_{xi}(k)\,m_{yj}(k) \tag{7}$$

$$W_{ij} = \alpha . d_{pij} . d_{mxi} \tag{8}$$

where:

d_{mxi}	ith class X-group matching degree
d_{myj}	jth class Y-group matching degree
d_{pij}	ijth rule performance degree
m_{xi}	ith class X-group membership
m_{yj}	jth class Y-group membership
k	sampling time index
W_{ij}	ijth rule weight

Equation-8 guaranties that establishing an association as a modeling rule, is based on its validity and truth (performance part) and also its efficiency and rate of occurrence (matching part).

Regulatory Mechanism

Due to noisy or rarely-occurring data, the inference network may create classes and establish associations which are not correct or effective. To avoid such problems, a regulatory mechanism is required to inspect the overall performance of the system and enhance it. The UWVs, developed for each rule, are used in this mechanism to verify the validity of fuzzy labels and their corresponding associations.

This mechanism checks the matching and performance degrees of each rule, iteratively. If these degrees decrease lower than some predefined thresholds, the related node or association (corresponding to membership or performance) will be removed or set to a zero weight.

Modes of Operation

The proposed modeling system operates in two independent modes, learning and prediction. In the learning mode, it learns the fuzzy labels for each variable and establishes the valid associations between them. This mode is based on the algorithms and formulas described earlier. In the prediction mode, the systems uses the learned parameters to predict the plant state in the next sampling time, based on current state and inputs. The defuzzifier unit, connected to the output of Y-group layer, is used to convert the fuzzy description of next plant state to a crisp value.

All the rules, based on their associative weights, may be considered in output prediction. Also we may select more important and valid rules to take part in prediction process.

To convert the fuzzy output of the model to a crisp value, a simple defuzzification may be defined as follows:

$$out = \frac{\sum_{j=0}^{cy-1}\sum_{i=0}^{cx-1} m_{xi} W_{ij} w_j}{\sum_{j=0}^{cy-1}\sum_{i=0}^{cx-1} m_{xi} W_{ij}} \qquad (9)$$

where w_j stands for the nominal (center) value of class j in the Y-group, and cx and cy are the number of clusters in X and Y-group.

Obviously, the associations with higher weights may be considered as the rule base of the fuzzy model, if we are to model the plant qualitatively.

4. Evaluation

4.1. Simulation Results

The proposed approach has been evaluated using a series of computer simulations. As an example, the results of modeling a single-input/ single-output system is presented here. The system is characterized by following simple input/output relation:

$$y(k) = x^2(k) + 4\,x(k) + 1 \qquad (10)$$

where k is the sampling index, and x / y are the input / output of the system in the kth sampling time.

With a uniform distribution of input values in intervals [0,1] and [2,3], which leads to output values in [1,6] and [13,22], the classes generated for X and Y-group, and their respective UWVs, are shown in Figure 5 and Table 1. We see that fuzzy labels are defined only at those areas of universe of discourse where the variable has been observed enough. It means that, without any direct presumption regarding the fuzzy labels and MFs, the system can, adaptively, handle the fuzzification process.

TABLE.1
Performance Degrees for labels in Fig.5.a,b

Y-Group / X-Group	Low w = 1.52 md = 0.23	M.Low w = 2.41 md = 0.26	M.High w = 3.78 md = 0.21	High w = 5.12 md = 0.26
Low w = 0.17 md = 0.38	0.229	0.143	0.003	0.0
Medium w = 0.47 md = 0.26	0.003	0.096	0.134	0.014
High w = 0.83 md = 0.32	0.0	0.001	0.066	0.246

A comparison of predicted and real output of the plant is also presented in Figure 6. As we may expect, for the values greater than the highest class, prediction error will raise. This is due to use of nominal value of fuzzy labels (w_j) in defuzzification, which prevents the system from generating predictions greater than w_j values.

4.2. Extension to Knowledge Structures with more Complexity

Although we limit our discussion to the case of fuzzy rule-based modeling, the proposed approach may be extended to other types of fuzzy systems or other types of knowledge structures with more complexity. For instance, as we have noted, simple modifications make this approach proper to be used in fuzzy controllers or, in general, fuzzy rule-based decision-making systems.

One of the most advantages of this approach, however, is its extendibility to knowledge structures other than rule bases. A good candidate in this regard, is the Fuzzy Cognitive Map (FCM) [6]. Association-As-Rule strategy may be simply considered as the Association-As-Causality, which is conceptually, equivalent to the interactive relations of nodes in an FCM. Here, instead of edge values in {-1, 0, 1}, as in simple FCMs, we may consider fuzzy labels for activation of each event in the map. X and Y-group nodes in our proposed approach,

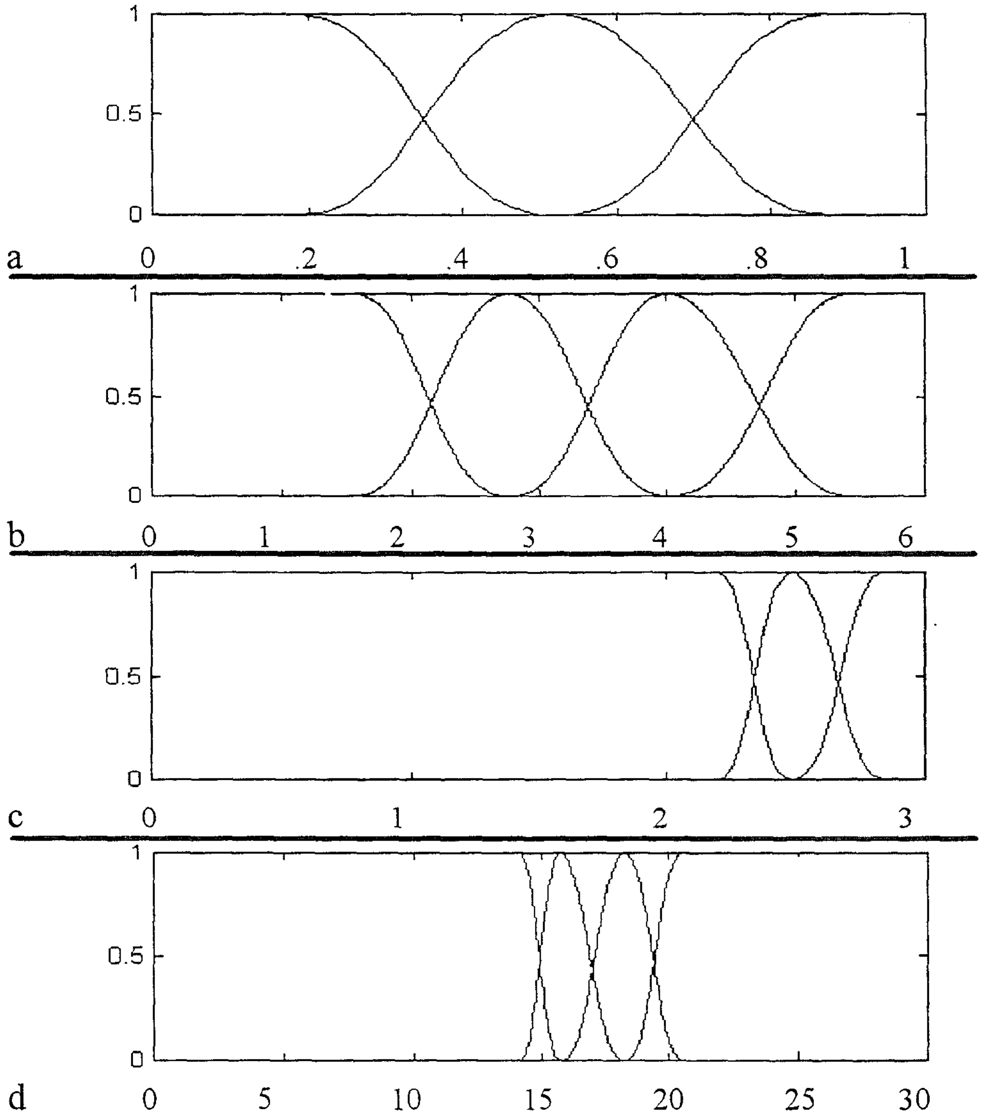

Figure 5. a) X-group fuzzy labels and MFs vs. data in [0,1] b) Y-group labels for (a) c) X-group fuzzy labels and MFs vs. data in [2,3] d) Y-group labels for (c)

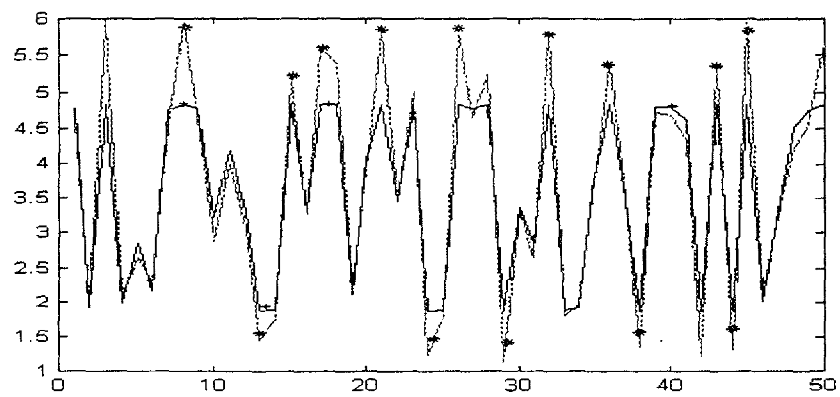

Figure 6. Actual (*) and predicted (+) plant output

take the role of cause and effect nodes of FCM, respectively. In this case, two groups may have common nodes, because of mutual causality of events in an FCM. We see that Neuron-As-Rule strategies can not be extended to such cases as FCM, so easily. It should be noted that considering the rule-based framework as the back-bone of every fuzzy system, is a very hard restriction that prevents the concept of fuzziness from being applicable to many information processing scenarios. When "the role model for *Computing with Words* (fuzzy logic and soft computing in general) is the human mind" [12], why limit ourselves to just one of its functionalities, i.e. rule-based inference.

5. Concluding Remarks

This work is, primarily, aimed at making an elementary step toward a comprehensive framework for adaptive fuzzy modeling. Regarding this, we have considered the capabilities such as "adaptability of the concepts and parameters (e.g. rules and weights)", "determination of the linguistic variables values (fuzzy labels)", "definition of new membership functions", and "selection of the system variables and the concepts essential to modeling". In this way, we have introduced an inference network, consisted of intrinsically adaptive units, which extracts the necessary fuzzy labels and rules out of the input/output data.

The main advantages of our approach are as follows:

1. The set of clustering network and regulatory mechanism operates in two modes of learning and prediction, simultaneously. Therefore it may be used effectively, in on-line applications.
2. There is no presumption regarding the number and location of fuzzy labels in the variables universe of discourse, and these are set, adaptively, by the modeling system. It should be noted that the shapes of MFs are still constrained by the distance function, used in the clustering subnet.
3. Simultaneous learning of input by more than one cluster neuron, results in inhibitory/ excitatory interaction of fuzzy labels and refinement of MFs. More complicated structures and nodal interaction may be considered to, adaptively and more effectively, define and modify the shape of MFs according to the semantic relations of fuzzy labels and the input/output data.
4. The system is capable of being implemented by parallel and distributed processing elements.
5. The proposed approach may be extended to include more complex knowledge structures, other than fuzzy rule bases.

References

[1] A. Arya, K. Badie, H. Zohour and M. Khazane-dari, "An adaptive fuzzy controller for the NC machine of a CAD/CAM system," in *Proc First Iranian Conf. on Elec Eng, ICEE-93*, Tehran Polytechnic, 1993.

[2] H. Ishibuchi, et al., "Selecting fuzzy IF-THEN rules for classification problems using genetic algorithms," *IEEE Trans Fuzzy Systems*, vol. 3, no. 3, 1995.

[3] J.R. Jang, "ANFIS: Adaptive-network-based fuzzy inference system," *IEEE Trans. Sys., Man, Cybern.*, vol. 23, no. 3, 1993.

[4] T.L. Huntsberger and P. Ajjimarangsee, "Parallel self-organizing feature maps for unsupervised pattern recognition," *Int'l J. General Systems*, vol. 16, no. 4, 1990.

[5] C.L. Karr and E.J. Gentry, "Fuzzy control of pH using genetic algorithms," *IEEE Trans Fuzzy Systems*, vol. 1, no. 1, 1993.

[6] S.G. Kong and B. Kosko, "Adaptive fuzzy system for backing up a truck-and-trailer," *IEEE Trans. Neural Networks*, vol. 3, no. 2, 1992.

[7] E.H. Mamdani, et al., "Use of fuzzy logic for implementing rule-based control of industrial processes," *Fuzzy Sets and Decision Analysis*, Elsevier, 1984.

[8] W. Pedrycz, "An identification algorithm in fuzzy relational systems," *Fuzzy Sets and Systems*, vol. 13, pp. 153-167, 1984.

[9] M. Sugeno and T. Yasukawa, "A fuzzy-logic-based approach to qualitative modeling," *IEEE Trans. Fuzzy Sys.*, vol. 1, no. 1, 1993.

[10] C.T. Sun, "Rule-base structure identification in an adaptive-network-based fuzzy inference system," *IEEE Trans Fuzzy Systems*, vol. 2, no. 1, 1994.

[11] L.A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Trans Sys., Man, Cybern.*, vol. 3, no. 1, 1973.

[12] L.A. Zadeh, "Fuzzy logic = computing with words," *IEEE Trans Fuzzy Systems*, vol. 4, no. 2, 1996.