Abstract—A neurofuzzy approach for a given set of input-output training data is proposed in two phases. Firstly, the data set is partitioned automatically into a set of clusters. Then a fuzzy if-then rule is extracted from each cluster to form a fuzzy rule base. Secondly, a fuzzy neural network is constructed accordingly and parameters are tuned to increase the precision of the fuzzy rule base. This network is able to learn and optimize the rule base of a Sugeno like Fuzzy inference system using Hybrid learning algorithm, which combines gradient descent, and least square mean algorithm. This proposed neurofuzzy system has the advantage of determining the number of rules automatically and also reduce the number of rules, decrease computational time, learns faster and consumes less memory. The authors also investigate that how neurofuzzy techniques can be applied in the area of control theory to design a fuzzy controller for linear and nonlinear dynamic systems modelling from a set of input/output data. The simulation analysis on a wide range of processes, to identify nonlinear components on-line in a control system and a benchmark problem involving the prediction of a chaotic time series is carried out. Furthermore, the well-known examples of linear and nonlinear systems are also simulated under the Matlab/Simulink environment. The above combination is also illustrated in modeling the relationship between automobile trips and demographic factors.

Keywords—Fuzzy control, Neuro-Fuzzy Techniques, Fuzzy Subtractive Clustering, Extraction of Rules, and Optimization of Membership Functions.

I. INTRODUCTION

In a conventional fuzzy approach the membership functions and the consequent models are fixed by the model designer according to a priori knowledge. If this set is not available but a set of input-output data is observed from the process, the components of a fuzzy system (membership and consequent models) can be represented in a parametric form and the parameters are tuned by neural networks. In that case the fuzzy systems turn into neurofuzzy system. A fuzzy system can explain the knowledge it encodes but can’t learn or adapt its knowledge from training examples, while a neural network can learn from training examples but can not explain what it has learned. Fuzzy systems and neural networks have complementary strengths and weaknesses. As a result, many researchers are trying to integrate these two schemes to generate hybrid models that can take advantage of strong points of both. This is also the motivation of our research.

Recently, neurofuzzy system modeling has attracted a lot of attention [6, 10, 14, 15, 23, 24, 27-32]. In general, this approach involves two major phases, structure identification and parameter identification. The structure identification amounts to determine the proper number of rules needed i.e. finding how many rules are necessary and sufficient to properly model the available data and the number of membership functions for input and output variables. Parameter learning phase is used to tune the coefficients of each rule (like the shape and positions of membership functions). Fast computation speed is attained by requiring much less tunable parameters. There is a need for effective methods for tuning the membership functions so as to minimize the output error measure or maximize performance index.

For structure identification, different researchers [16-20, 25, 30-32] use different methods to extract initial fuzzy rules from given input-output data. Clustering techniques [2-4, 22, 25] have been recognized as a powerful alternative approach to develop fuzzy systems. Clustering of numerical data forms the basis of many classification and system-modeling algorithms. The purpose of clustering is to identify natural grouping of data from a large data set to produce a concise representation of a system’s behavior. Clustering algorithms typically require the user to prespecify the number of cluster centers and their initial locations. The preceding discussion shows that different researchers have used different clustering algorithms and different cluster validity indices to decide on the number of rules. A clustering method called subtractive clustering forms the basis of the present work.

For parameter identification, most systems [2, 14-15, 19] use backpropagation to refine parameters of the system. However backpropagation suffers from the problem of local minima and low convergence rate. To alleviate these difficulties, different methods of least square estimation (LSE) are proposed [17]. However these methods suffer from the necessity of initializing a certain parameter. The hybrid algorithm means combination of gradient descent and least square estimate (LSE) solves the above problem as discussed in [10]. Hence in this paper backpropagation learning rule is used to tune the parameters in the hidden layer and parameters in the output layer will be identified by the least square method.

This paper contributes a neurofuzzy system for the given set of input-output data, which is obtained in two steps. First, the data set is partitioned into a set of clusters based on the similarity of data. Then using subtractive clustering algorithm a fuzzy if-then rule is extracted from each cluster to form a fuzzy rule base. Secondly, a fuzzy neural network is designed...
accordingly to optimize the parameters of the fuzzy system. To decrease the size of the search space and speed up the convergence we are using a hybrid learning algorithm which combines the gradient descent and least square estimator (LSE) method. The proposed combination has the advantage of determining the number of rules automatically, learns faster, consumes less memory and produces lower root mean square error than other methods. A benchmark problem in chaotic time series prediction and automobile trip generation based on its demographics are simulated to compare the performance of the above combination with the published results of other algorithms.

II. NEUROFUZZY LEARNING

The learning scheme is mainly composed of two steps. In the first step, the number of rules nodes (hence the structure of the network) and initial rule parameters (weights) are determined using structure identification; in the latter all parameters are adjusted using parameter identification as shown in Fig. 1.

To initiate the structure tuning, a training set composed of input-output data, which contains n inputs and one output must be provided. Without loss of generality, we assume that the data points have been normalized in each dimension so that they are bounded by a unit hypercube. We consider each data point as a possible cluster center and define a measure of their potential as discussed in [2, 3]. To extract the fuzzy rules obtained in first phase. To realize the described fuzzy inference mechanism, the operation of a neural network is shown in Fig. 2 and described below:

1. Layer 1: Units in this layer receives the input value \((X_1, X_2, \ldots, X_n)\) and acts as the fuzzy sets representing the corresponding input variable. Nodes in this layer are arranged into \(j\) groups; each group representing the IF-part of a fuzzy rule. Node \((i, j)\) of this layer produces its output \(O^{(1)}_{ij}\) by computing the corresponding Gaussian membership function:

\[
O^{(1)}_{ij} = A_{ij}(X_j)
\]

2. Layer 2: The number of nodes in this layer is equal to the number of fuzzy rules. A node in this layer represents a fuzzy rule; for each node, there are \(n\) fixed links from the input term nodes representing the IF-part of the fuzzy rule. Node \(O^{(2)}_{ij}\) of this performs the AND operation by product of all its inputs from layer 1. For instance,

\[
O^{(2)}_{ij} = \prod_{i=1}^{n} O^{(1)}_{ij}
\]

3. Layer 3: This layer contains only one node whose output \(O^{(3)}\) represents the result of centroid defuzzification, i.e.,

\[
O^{(3)} = \frac{\sum_{j=1}^{J} O^{(2)}_{ij} c_j}{\sum_{j=1}^{J} O^{(2)}_{ij}}
\]

Here \(c_j\) is the class of data as discussed above and it is also called the fuzzy singletons defined on output variables. Apparently, \(m_{ij}\) and \(\sigma_{ij}\) are the parameters that can be tuned to improve the performance of the system. After that a hybrid learning algorithm which combines gradient descent and least square estimator method is used to refine these parameters. Each epoch of the hybrid learning procedure is composed of a forward pass and backward pass. In the forward pass, input data is supplied and functional signals go forward to calculate each node output. The consequent parameters are identified by least square estimator method. After identifying the parameters, the functional signals keep going forward till the error measure is calculated. In the backward pass, the error rates (derivative of the error measure w.r.t. each node output) propagate from the output end towards the input and the...
premise parameters are updated by gradient method. The details of Hybrid learning algorithm is given by Jang in [10] and we are using the same procedure.

Fig. 2 Architecture of Fuzzy neural network

III. SIMULATION RESULTS

The simulations results are further divided into two parts. In the first part, Neurofuzzy learning is used to identify the Fuzzy PI and PD type controllers. In second phase, we shall use the proposed learning algorithm to illustrate some examples such as a benchmark problem of chaotic time series prediction. A comparison between our systems and the published results of other two systems proposed by Jang [10] in 1993 and Chiu [2] in 1994 is presented.

A. Identification of Fuzzy Controller

The neurofuzzy learning is used for identification of PD type FLC (FPDC) and PI type FLC (FPIC). The Block diagram of FPIC is shown in Fig. 3.

The change in error is defined as

\[
\Delta e(k) = e(k) - e(k - 1)
\]

where \( e(k) \) is the error at the \( k \)th sample.

All membership functions (MFs) for controller inputs (i.e., \( e \) and \( \Delta e \)) and incremental change in controller output (i.e., \( \Delta u \)) are defined on the common normalized domain \([-1, 1]\). The membership functions are shown in Fig. 4. The operation of PI type FLC can be described by

\[
u(k) = u(k - 1) + \Delta u(k)
\]

In (6), \( k \) is the sampling instant and \( \Delta u \) is the incremental change in controller output, which is determined by the rules of the form If \( e \) is \( E \) and \( \Delta e \) is \( \Delta E \), then \( \Delta u \) is \( \Delta U \). The rule base for computing is a standard [21] one as shown in Table I.

![Fig. 4 MFs for e, Δe and Δu](image)

<table>
<thead>
<tr>
<th>( \Delta e/e )</th>
<th>NB</th>
<th>NM</th>
<th>NS</th>
<th>ZE</th>
<th>PS</th>
<th>PM</th>
<th>PB</th>
</tr>
</thead>
<tbody>
<tr>
<td>NB</td>
<td>NB</td>
<td>NB</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>ZE</td>
</tr>
<tr>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NS</td>
<td>NS</td>
<td>ZE</td>
<td>PS</td>
</tr>
<tr>
<td>NS</td>
<td>NB</td>
<td>NM</td>
<td>NS</td>
<td>NS</td>
<td>ZE</td>
<td>PS</td>
<td>PM</td>
</tr>
<tr>
<td>ZE</td>
<td>NB</td>
<td>NM</td>
<td>NS</td>
<td>ZE</td>
<td>PS</td>
<td>PM</td>
<td>PB</td>
</tr>
<tr>
<td>PS</td>
<td>NM</td>
<td>NS</td>
<td>ZE</td>
<td>PS</td>
<td>PS</td>
<td>PM</td>
<td>PB</td>
</tr>
<tr>
<td>PM</td>
<td>NS</td>
<td>ZE</td>
<td>PS</td>
<td>PM</td>
<td>PM</td>
<td>PM</td>
<td>PB</td>
</tr>
<tr>
<td>PB</td>
<td>ZE</td>
<td>PS</td>
<td>PS</td>
<td>PM</td>
<td>PM</td>
<td>PB</td>
<td>PB</td>
</tr>
</tbody>
</table>

The FPIC in Fig. 3 uses 49 rules and 7 membership functions in each variable to compute output and exhibits good performance [21]. Next, we investigate the following – Given some data describing the output (\( \Delta u \)) as a function of Inputs (i.e., \( e \) and \( \Delta e \)), now main aim is to extract a smaller set of rules using neurofuzzy learning to do the same. Then, the performance of the simple controller (identified system) compare with the original one. Now the following steps are followed:

1. Data Generation

To identify the FPIC and FPDC, some data is needed, i.e., a set of two-dimensional input vectors \( X = \{X_1, X_2, \ldots, X_n\} \) and the associated set of one-dimensional output vectors \( Y = \{Y_1, \ldots, Y_n\} \) where \( X = \{e \) and \( \Delta e \} \) and \( Y = \{u\} \) is required. Here, the training data has been generated by sampling input variables \( e \) and \( \Delta e \) uniformly at the step size of 0.1 from the existing FLC [21] and computing the value of \( \{u\} \) for each sampled point. The number of data points generated is 441.

2. Rule Extraction and Membership Functions

After generating the data, the next step is to estimate the initial rules. Then after applying Subtractive Clustering algorithm, four clusters (rules) are extracted. The unit step

Fig. 3 Block Diagram of FPIC
response using these four rules is not so close to the identified system [7, 9, 25]. Hence, there is a need of optimization of these rules. Parameter optimization is used for tuning of membership functions to minimize the output error measure or maximize performance index using neural networks. Hybrid learning algorithm is used for training to modify the above parameters after obtaining the Fuzzy inference system from subtracting clustering. This algorithm iteratively learns the parameter of the premise membership functions via back propagation and optimizes the parameters of the consequent equations via linear least-squares estimation. The training is continued until the error measure becomes constant. As the value of these parameters change, the Gaussian membership function varies accordingly. The membership functions after optimization for \( e \) are \( m_{f1}, m_{f2}, m_{f3} \) and \( m_{f4} \), and for \( \Delta e \) are \( m_{1}, m_{2}, m_{3} \) and \( m_{4} \) shown in Fig. 5. Finally the rules are written in the form of: Rule \( i \): If \( e \) is \( m_{fi} \) & \( \Delta e \) is \( m_{i} \) then class is \( c_{i} \) when indices \( i=1 \) to 4.

![Fig. 5 MFs for e, \( \Delta e \)](image)

### B. Results

The neurofuzzy learning has been tested on a variety of linear and nonlinear processes. The objective here is to justify whether the Fuzzy controller with less number of rules and membership functions can provide the same level of performance as that of the original one (system with 49 rules).

To demonstrate the effectiveness of the proposed combination, the results are reported for system with 49 rules and system with optimized rule base. After reducing the rules the computation become fast and it also consumes less memory. The computational time is calculated using the Process Explorer –Sysinternals software [27]. It has been shown clearly in the Table and the values of computational time and memory using same simulation time with and without clustering for each system is given in tabular forms in Table II.

In case of Fuzzy PI type and PD type controller, the system with 49 rules (original system) is denoted by FPIC and FPDC and system with 4 (optimized rules) rules is denoted by HFPIC and HFPDC. In this paper, it is emphasize that an identified system is called satisfactory only with respect to its closeness to the target system, here FPDC and FPIC. The proposed FLC is applied to an inverted pendulum (PT2 system) [13] which is given by the following differential equation:

\[
\frac{1}{w_o} \ddot{x} + \frac{2D}{w_o} \dot{x} + x = Vy
\]  

(7)

where \( w_o = \sqrt{50}, D = \frac{2}{5} \sqrt{2}, V = 1 \)  

(8)

The PT2 system models the behaviour of a two mass system, for example, spring damper combinations or revolution controls for electric motors. This system is controlled by FPIC and FPDC in both cases and results are compared in Fig. 6 and Fig. 7.

![Fig. 6 Unit Step Response of inverted pendulum with FPDC and HFPDC](image)

Secondly, the proposed fuzzy controller is applied on coupled tank for controlling the level of fluid. The control input is the pump drive voltage. The sensed output is the water depth in tank 2. The process of coupled tank [13] is given by:

\[
G(s) = \frac{0.4219}{1149.893s^2 + 111.834s + 1}
\]  

(9)

Then unit step response is observed in case of the system with 49 rules and system with reduced rule set using neurofuzzy learning. Response characteristics for the coupled tank with 49 rules and 4 rules are shown in Fig. 8.

The overall performance of the Fuzzy Controllers with 4 rules is compared with those of Fuzzy Controller with 49 rules. Response characteristics of the identified system in both cases (FPDC and FPIC) are very close to the original one. The computational time is decreased using 4 rules which are clearly shown in Table II.
### C. Application Examples

#### Example 1: Prediction of Chaotic Time Series
In this simulation, the proposed combination is used to predict the future values of chaotic time series. The time series used in our simulation is generated by the chaotic Mackey-Glass differential delay equation \[ \frac{dx(t)}{dt} = \frac{0.2x(t-\tau) - 0.1x(t)}{1 + x^{10}(t-\tau)} \] defined below:

The task is to use past values of \( x \) up to the time \( t \) to predict the value of \( x \) at some time \( t+\Delta t \) in the future. The standard input for this type of prediction is \( N \) points in the time series spaced \( s \) apart, i.e., the input vector is \( y = \{x(t-(N-1)s), \ldots, x(t-2s), x(t-s), x(t)\} \). For comparison with the published results of other methods, we use the same parameters \( \tau = 17 \), \( s=6 \), \( N=4 \), \( t = 118 \) to 1117. The first 500 pairs (training data set) were used for training while the remaining 500 pairs (checking data set) were used for validating the identified model. The number of membership functions and rules are initially determined by clustering. Then the Hybrid learning algorithm is used for training to update the parameters. The training is continued until the error measure becomes constant. After 20 epochs error becomes constant and \( \text{RMSE}_{\text{trn}} = 0.0015 \) and \( \text{RMSE}_{\text{chk}} = 0.0014 \), which is same as in [2] but when compared to the published result of Jang [10] the \( \text{RMSE}_{\text{trn}} = 0.0016 \) and \( \text{RMSE}_{\text{chk}} = 0.0015 \) after 499.5 epochs. The predicted values of error for both training data and checking data are comparatively less. The RMSE curves are shown in Fig. 9. Curve with ‘*’ is for training data and with ‘-’ is for checking data.

#### Example 2: Automobile Trip Prediction
In this example, the neurofuzzy learning is used to estimate the number of automobile trips generated from an area based on demographics of the area. Five demographics factors as taken by S. L. Chiu in [2] were considered: population, number of dwelling units, vehicle ownership, median household income and total employment. Hence there are 5 input variables and 1 output variable. Out of the original 100 data points, we will use 75 data points as training data and 25 data points as checking data, (as well as for test data to validate the model). Using the proposed combination, the average modeling error with respect to the training data was \( \text{RMSE}_{\text{trn}} = 0.26 \) and with respect to the checking data was \( \text{RMSE}_{\text{chk}} = 0.33 \) as shown in Fig 9, if compared by the published results of SL Chiu with same \( r_a = 0.5 \) and 3 rules the \( \text{RMSE}_{\text{trn}} = 0.34 \) and \( \text{RMSE}_{\text{chk}} = 0.37 \). Comparison between the both shows that neurofuzzy learning proposed in this paper gives us better results.
This paper has described a neurofuzzy learning which incorporates the architecture of Neural Network based Fuzzy inference system. A given training data set is partitioned into a set of clusters based on subtractive clustering method. This is fast and robust method to generate the suitable initial set of clusters based on subtractive clustering method. This is fast and robust method to generate the suitable initial set of clusters based on subtractive clustering method. This is fast and robust method to generate the suitable initial set of clusters based on subtractive clustering method. This is fast and robust method to generate the suitable initial set of clusters based on subtractive clustering method. This is fast and robust method to generate the suitable initial set of clusters based on subtractive clustering method. Then a hybrid learning algorithm is used to refine the parameters of fuzzy rule base. The advantages of the discussed neurofuzzy learning are that it determines the number of rules automatically, reduces computational time, learns faster and produces lower RMS errors than other method. Furthermore, the proposed method is able to reduce 49 rules to 4 rules maintaining almost the same level of performance. In addition, well known examples of inverted pendulum and a coupled tank are simulated and identification performance. In addition, well known examples of inverted pendulum and a coupled tank are simulated and identification performance. In addition, well known examples of inverted pendulum and a coupled tank are simulated and identification performance. In addition, well known examples of inverted pendulum and a coupled tank are simulated and identification performance. In addition, well known examples of inverted pendulum and a coupled tank are simulated and identification performance. In addition, well known examples of inverted pendulum and a coupled tank are simulated and identification performance.

IV. CONCLUSION

Fig. 9(b) Training and Checking RMSE curves for Example 2.

REFERENCES

[18] G. Castellano and A.M. Fanelli, “Fuzzy inference and rule extraction and using neural network.” URL:
Seema Chopra was born in Punjab, India in 1976. She received B. Tech. degree in Instrumentation and Control Engineering from Kurukshetra University in 2000, M. Tech. degree in Control Engineering from NIT Kurukshetra in 2002. Currently she is doing Ph.D. in Control and Guidance from the department of Electronics and Computer Engineering, Indian Institute of Technology, Roorkee, India under MHRD scheme. She is working as Lecturer in Instrument and Control Department in Haryana Engineering College Jagadhri, India and Centre of Advance Technology, Research Centre, Jagadhri, India since 2001. Her current area of research includes Intelligent Control, Fuzzy Logic, Neural Networks, Fuzzy Clustering and SCADA System.

Ranajit Mitra was born in Allahabad, India in 1945. He received B.Sc. degree from Lucknow University, M. Sc. (Tech.) from University of Allahabad, M.Tech. from Indian Institute of Technology, Kharagpur and Ph.D. degree in Control and Guidance from University of Roorkee, India in 1975. Besides Post Doctoral work at DFVRL laboratory in West Germany, he had been in the faculty of department of System and Control University of Technology Baghdad and Thapar Engineering College, Patiala. For the Past 36 years Dr. Mitra is with department of Electronics and Computer Engineering, Indian Institute of Technology, Roorkee, India (formerly University of Roorkee). His current area of research includes Fuzzy Control, development of new experimental systems in the area of Control and Instrumentation.

Vijay Kumar was born in Biharshariff, India in 1954. He received B.E. degree from M.M.M College of Engineering Gorakhpur, M.E. and Ph.D. degree in Control and Guidance from University of Roorkee, India in 1983 and 2001 respectively. He was lecturer in department of Electronics and Computer Engineering, Indian Institute of Technology, Roorkee, India (formerly University of Roorkee) during 1983 to 2002, where currently he is working as Assistant Professor. His current area of research includes Fuzzy Control, Model Order Reductions.