1994 INTERNATIONAL SYMPOSIUM ON ARTIFICIAL NEURAL NETWORKS

PROCEEDINGS

ISANN'94

National Cheng Kung University
Tainan, Taiwan, Republic of China
A Similarity Measure for Fuzzy Rules in a Fuzzy Neural Network

Chun-Tang Chao, Young-Jeng Chen, Ching-Cheng Teng
Institute of Control Engineering, National Chiao-Tung University, Taiwan, R.O.C.

ABSTRACT:
This paper presents a fuzzy neural network system (FNNS) for implementing fuzzy inference systems. In the FNNS, a fuzzy similarity measure for fuzzy rules is proposed to eliminate redundant fuzzy logical rules. Moreover, a fuzzy similarity measure for fuzzy sets is applied to combine similar input linguistic term nodes. Thus we obtain a method for reducing the complexity of a fuzzy neural network. We also propose a new and effective on-line initialization method for choosing the initial parameters of the FNNS. A computer simulation is presented to illustrate the performance and applicability of the proposed FNNS.

I. INTRODUCTION

The benefits of combining fuzzy logic and neural networks have been explored extensively in the literature, e.g., the fuzzy neural network in [2, 6], the adaptive-network-based fuzzy inference system in [3], and the fuzzy logical system in [14]. However, no efficient process for reducing the complexity of a fuzzy neural network has been presented. Furthermore, the concept of a measure of similarity between fuzzy sets has been applied in many ways [1, 9]. Lin and Lee [7] presented an algebraic and geometric derivation to provide their proposed FSM (Fuzzy Similarity Measure) with a clear mathematical and physical meaning. However, a fuzzy similarity measure has yet to be applied to reduce the complexity of a fuzzy neural network.

The FNNS, with a proposed on-line initialization method, can easily construct the initial state of the fuzzy logical rules. We apply the fuzzy similarity measure to combine similar linguistic terms into a single linguistic term to reduce the complexity of the FNNS. We also present a fuzzy similarity measure for fuzzy rules to eliminate redundant fuzzy logical rules.

II. FUZZY NEURAL NETWORK SYSTEM (FNNS)

The initial network structure adopted in the proposed FNNS is shown in Fig. 1. The structure is distinguished by its direct construction of fuzzy rules without any other adjustment. For example, suppose we encounter the jth fuzzy rule described as follows:

\[
\text{IF } z_1 \text{ is } A_1^j \text{ and } z_2 \text{ is } A_2^j \text{ and } \ldots \text{ and } z_n \text{ is } A_n^j \text{ THEN } y \text{ is } \beta^j. \tag{1}
\]

A connectionist structure based on this fuzzy rule is illustrated in Fig. 2.

On the other hand, the disadvantage of the network structure is that it requires a large number of term nodes. As we shown in Fig. 1, we require \( m \times n \) term nodes in layer two for \( n \) inputs and \( m \) fuzzy rules at the initial time. We will apply the fuzzy similarity measure in Section 3 to combine similar term nodes corresponding to a fixed input linguistic variable \( z_i \) and overcome this problem. The numerical output of the fuzzy inference system with center average defuzzifier, product inference rule, and singleton fuzzifier is of the following form:

\[
y = \frac{\sum_{j=1}^{m} \beta^j (\prod_{i=1}^{n} \mu_{A_i^j}(z_i))}{\sum_{j=1}^{m} \prod_{i=1}^{n} \mu_{A_i^j}(z_i)}, \tag{2}
\]

where \( \mu_{A_i^j} \) denotes the membership function of fuzzy set \( A_i^j \).

Once the FNNS has been initialized, a gradient descent-based back-propagation algorithm (BP) [11, 13] is employed to adjust the parameters of the fuzzy neural network by using the training patterns.
III. SIMILARITY MEASURE FOR FUZZY SETS AND FUZZY RULES

Once supervised learning using the BP algorithm is finished, we may find that some of the fuzzy sets in layer 2 are almost the same. We can use the following fuzzy similarity measure [7] to check the degree of similarity of two fuzzy sets:

\[
\text{Degree}(A_1 = A_2) = E(A_1, A_2) = \frac{M(A_1 \cap A_2)}{M(A_1 \cup A_2)},
\]

For any two fuzzy sets \(A_1\) and \(A_2\), \(M(A_1 \cup A_2)\) can be easily derived as follows:

\[
M(A_1 \cup A_2) = M(A_1) + M(A_2) - M(A_1 \cap A_2).
\]
To make the computation of Eq. (3) feasible, we can use a tent function (triangular function) to approximate a Gaussian function, that is

\[ \exp\left[-\frac{(x - m)^2}{\sigma^2}\right] \rightarrow \max[0, \frac{\sigma\sqrt{\pi} - |x - m|}{\sigma\sqrt{\pi}}]. \]  

(5)

**A. Approximation Equations for the Similarity Measure**

We consider the similarity measure in four different cases based on the triangular membership functions (Figures 3(i)—3(iv)).

Case (i): \( m_1 = m_2 \) and \( \sigma_1 \geq \sigma_2 \).

\[
E(A_1, A_2) = \frac{M(A_2)}{M(A_2)} = \frac{\sigma_2}{\sigma_1}.
\]

(6)

Case (ii): \( |\sigma_1 - \sigma_2|\sqrt{\pi} \leq m_1 - m_2 \leq (\sigma_1 + \sigma_2)\sqrt{\pi} \) and \( m_1 > m_2 \).

\[
E(A_1, A_2) = \frac{(c_1 + c_2)h_1}{2(\sigma_1 + \sigma_2)\sqrt{\pi} - (c_1 + c_2)h_1},
\]

where

\[
\begin{align*}
c_1 &= \frac{\sigma_1(m_2 - m_1) + \sigma_1(\sigma_1 + \sigma_2)\sqrt{\pi}}{\sigma_1 + \sigma_2}, \\
c_2 &= \frac{\sigma_2(m_2 - m_1) + \sigma_2(\sigma_1 + \sigma_2)\sqrt{\pi}}{\sigma_1 + \sigma_2}, \\
and \quad h_1 &= \frac{(m_2 - m_1) + (\sigma_1 + \sigma_2)\sqrt{\pi}}{(\sigma_1 + \sigma_2)\sqrt{\pi}}.
\end{align*}
\]

Case (iii): \( m_1 - m_2 \leq |\sigma_2 - \sigma_1|\sqrt{\pi} \) and \( m_1 > m_2 \).

\[
E(A_1, A_2) = \frac{c_1h_1 + c_2h_2 + c_3h_3}{2(\sigma_1 + \sigma_2)\sqrt{\pi} - (c_1h_1 + c_2h_2 + c_3h_3)},
\]

where

\[
\begin{align*}
c_1 &= \frac{\sigma_1(m_2 - m_1) + \sigma_1(\sigma_2 + \sigma_1)\sqrt{\pi}}{\sigma_1 + \sigma_2}, \\
c_2 &= \frac{\sigma_1(m_2 - m_1) + \sigma_1(\sigma_2 - \sigma_1)\sqrt{\pi}}{\sigma_2 - \sigma_1}, \\
c_3 &= \frac{2\sigma_1\sqrt{\pi} - (c_1 + c_2)}{\sigma_2 - \sigma_1}, \\
h_1 &= \frac{(m_2 - m_1) + (\sigma_2 + \sigma_1)\sqrt{\pi}}{(\sigma_2 + \sigma_1)\sqrt{\pi}}, \\
h_2 &= \frac{(m_2 - m_1) + (\sigma_2 - \sigma_1)\sqrt{\pi}}{(\sigma_2 - \sigma_1)\sqrt{\pi}}, \\
and \quad h_3 &= h_1 + h_2.
\end{align*}
\]

Case (iv): \( m_1 - m_2 > (\sigma_1 + \sigma_2)\sqrt{\pi} \) and \( m_1 > m_2 \).

\[ E(A_1, A_2) = 0. \]

(9)
Figure 3: Similarity measure of two triangular fuzzy sets, $A_1$ and $A_2$: (i) $A_2 \subseteq A_1$; (ii) two membership functions have an intersection point, $(s_1, h_1)$; (iii) two membership functions have two intersection points, $(s_1, h_1)$ and $(s_2, h_2)$; (iv) $A_1 \cap A_2 = \emptyset$.

With a given reference value $\gamma_s$, $0 < \gamma_s \leq 1$, if $E(A_1, A_2) \geq \gamma_s$, then we can combine $A_1$ and $A_2$ into a new fuzzy set $A_{\text{new}}$. We try to determine $A_{\text{new}}$ in the following.

Case (i):

$$\sigma_{\text{new}} = \frac{\sigma_1 + \sigma_2}{2}$$

and

$$m_{\text{new}} = m_1 \text{ or } m_2.$$  \hspace{1cm} (10)

Case (ii):

$$\sigma_{\text{new}} = \frac{(m_1 - m_2) + (\sigma_1 + \sigma_2)/\sqrt{\pi}}{2\sqrt{\pi}}$$

and

$$m_{\text{new}} = \frac{(m_1 + m_2) + (\sigma_1 - \sigma_2)/\sqrt{\pi}}{2}.$$  \hspace{1cm} (12)

Case (iii):

$$\sigma_{\text{new}} = \frac{\sigma_1 + \sigma_2}{2}$$

and

$$m_{\text{new}} = \frac{m_1 + m_2}{2}.$$  \hspace{1cm} (14)

B. Similarity Measure of Fuzzy Rules

To determine whether two fuzzy rules are similar, we must evaluate the degree of similarity of the fuzzy rules. With the proposed FNNS, more specifically, we need to calculate the degree of similarity of both the consequences and preconditions. We will describe the similarity measure for the fuzzy rules described below:

$$R^k: \text{IF } x_1 \text{ is } A_1^k \text{ and } x_2 \text{ is } A_2^k \text{ and } \ldots \text{ and } x_n \text{ is } A_n^k \text{ THEN } y \text{ is } \beta^k$$

-164-
$R^i$: IF $x_1$ is $A^i_1$ and $x_2$ is $A^i_2$ and ... and $x_n$ is $A^i_n$ THEN $y$ is $\beta^i$,

where $R^k$ and $R^l$ represent the $k$th and the $l$th fuzzy rules, respectively.

(A) Similarity measure for consequences

To calculate the degree of similarity of two consequence weights $\beta^k$ and $\beta^l$, we should define a fuzzy set $A_c$. Then the similarity measure for $\beta^k$ and $\beta^l$ can be characterized as follows:

$$E_c(\beta^k, \beta^l) = \left\{ \begin{array}{ll} 1, & \text{if } \mu_{A_c}(\beta^k - \beta^l) \geq \gamma_c \\ 0, & \text{if } \mu_{A_c}(\beta^k - \beta^l) < \gamma_c \end{array} \right. \quad (16)$$

where $E_c(\beta^k, \beta^l)$ is the degree of similarity of $\beta^k$ and $\beta^l$ and $\gamma_c$, $0 < \gamma_c \leq 1$, is a reference value determined by the user. In the FNNS, the membership function of $A_c$ is a triangular function as shown below:

$$\mu_{A_c}(x) = \max[0, \frac{\beta_{\text{max}} - \beta_{\text{min}} - |x|}{\beta_{\text{max}} - \beta_{\text{min}}}] \quad (17)$$

where $\beta_{\text{max}}$ and $\beta_{\text{min}}$ are the maximum and minimum consequence weights in the training result of the FNNS.

(B) Similarity measure for preconditions

The similarity measure of the preconditions ($E_p$) can be characterized as follows:

$$E_p(A^k, A^l) = \min\{E(A^k_1, A^l_1), E(A^k_2, A^l_2), \cdots, E(A^k_n, A^l_n)\} \quad (18)$$

where

$$A^k = [A^k_1, A^k_2, \cdots, A^k_n]$$

and

$$A^l = [A^l_1, A^l_2, \cdots, A^l_n].$$

Once $E_p(A^k, A^l)$ reaches a reference value $\gamma_p$, $0 < \gamma_p \leq 1$, then all of these fuzzy set pairs are considered to be very similar. This is the reason why we employ the min operation on Eq.(18).

Based on the discussion in (A) and (B), the similarity measure of the fuzzy rules is defined as

$$E_r(R^k, R^l) = E_c(\beta^k, \beta^l) \cdot E_p(A^k, A^l) \quad (19)$$

The user is asked to set a reference value $\gamma_r$, $0 < \gamma_r \leq 1$, in the FNNS. Then any two fuzzy rules $R^k$ and $R^l$ with $E_r(R^k, R^l) \geq \gamma_r$ can be combined into a new fuzzy rule $R^{\text{new}}$. Once $R^{\text{new}}$ is taken to replace both $R^k$ and $R^l$, then the corresponding term nodes $R^k$ and $R^l$ will be eliminated. The term nodes of $R^{\text{new}}$ can be directly obtained by using the combination method of fuzzy sets presented in Sec. 3.1, i.e., $A^{\text{new}}_i$ is the fuzzy set combination of $A^k_i$ and $A^l_i$ for $i = 1, \cdots, n$. On the other hand, If $\gamma_r$ is high enough, then the consequence weight $\beta^{\text{new}}$ of $R^{\text{new}}$ can be simply chosen as $\beta^{\text{new}} = \frac{\beta^k + \beta^l}{2}$. This above similarity measure of fuzzy rules can be easily extended to MIMO systems.

IV. A NEW ON-LINE INITIALIZATION METHOD

Since the initial structure of FNNS does not take an ordinary fuzzy partition of the input space, how to choose the initial parameters of the FNNS becomes an important problem. Suppose, at instant $j$, $1 \leq j \leq m$, a training pattern $(x_1(j), \cdots, x_n(j); y(j))$ is presented. We can directly set the parameters

$$\beta^i = y(j) \quad (20)$$

and

$$m_{ij} = x_i(j), \quad 1 \leq i \leq n. \quad (21)$$
The remaining problem is how to determine the corresponding width \( (\sigma_{ij}) \) for \( A_i^k \), this is also the main problem in the on-line initialization method. In the fuzzy neural network systems [2, 3, 14], the initial value of parameters can be easily set in such a way that the membership functions are equally spaced along the operating range of each input variable. Then these membership functions will satisfy \( \epsilon \)-completeness [4, 5], which means that given a value \( x \) of one of the inputs in the operating range, we can always find a linguistic label \( A \) such that \( \mu_A(x) \geq \epsilon \). In this manner, the fuzzy inference system can provide smooth transition and sufficient overlapping from one linguistic label to another. It is especially mentioned that if the \( \epsilon \)-completeness condition is not satisfied, there may be no fuzzy rules fired when the input data is fed into the fuzzy neural network.

Before going further to show the choice and characteristic of \( \sigma_{ij} \), we introduce the following notation. We note that the following notation are based on a fixed \( k \) or \( A_i^k \), \( 1 \leq k \leq m \).

- \( A_i^R \): the fuzzy set which is on the right side of \( A_i^k \) and is closest to \( A_i^k \).
- \( A_i^L \): the fuzzy set which is on the left side of \( A_i^k \) and is closest to \( A_i^k \).
- \( m_{iR} \): the corresponding center of \( A_i^R \).
- \( m_{iL} \): the corresponding center of \( A_i^L \).
- \( A_{i,r} \): the rightest fuzzy set in \( A_i^k \), \( j=1, \cdots, m \).
- \( A_{i,l} \): the leftest fuzzy set in \( A_i^k \), \( j=1, \cdots, m \).
- \( m_{i,r} \): the corresponding center of \( A_{i,r} \).
- \( m_{i,l} \): the corresponding center of \( A_{i,l} \).

Let \( X_i \) denote the universe of discourse of the input \( x_i \), we can also treat fuzzy sets \( A_i^k \), for \( j = 1 \) to \( m \), as fuzzy numbers defined in \( X_i \). In the FNNS, we let \( A_i = (A_1^i, A_2^i, \cdots, A_m^i) \) be a semi-closed fuzzy set [8], i.e., a fuzzy set with

\[
\mu_{A_{i,l}}(x_i) = 1, \quad x_i \in X_i \text{ and } x_i \geq m_{i,r} \tag{22}
\]

\[
\mu_{A_{i,l}}(x_i) = 1, \quad x_i \in X_i \text{ and } x_i \leq m_{i,l} \tag{23}
\]

The special choice for \( \sigma_{ij} \) is

\[
\sigma_{ij} = \frac{\max\{|m_{ij} - m_{iR}|, |m_{ij} - m_{iL}|\}}{\sqrt{\ln \lambda_i}} \tag{24}
\]

where \( \lambda_i \) is the overlapping factor, \( 0 < \lambda_i < 1 \).

We can prove that the semi-closed fuzzy set \( A_i = (A_1^i, A_2^i, \cdots, A_m^i) \), where each linguistic label \( A_i^j \) has a Gaussian membership function constructed by the preceding initial \( m_{ij} \) (Eq. (21)) and \( \sigma_{ij} \) (Eq. (24)), will satisfy \( \epsilon \)-completeness. That is,

\[
\text{for all } x_i \in X_i, \quad \text{there exists } k \in 1, 2, \cdots, m, \text{ such that } \mu_{A_i^k}(x_i) \geq \epsilon = \lambda_i,
\]

where \( \lambda_i, 0 < \lambda_i < 1 \), is the overlapping factor.

V. AN ILLUSTRATIVE EXAMPLE

In this example [10], the plant to be identified is described by the second-order difference equation

\[
y(k + 1) = f[y(k), y(k - 1)] + u(k),
\]

-166-
where the unknown function \( f \) has the form
\[
f[y(k), y(k - 1)] = \frac{y(k)y(k - 1)[y(k) + 2.5]}{1 + y^2(k) + y^2(k - 1)}.
\]

A series-parallel type of identifier [10], implemented by the FNNS, is described by the equation
\[
\dot{y}(k + 1) = \hat{f}[y(k), y(k - 1)] + u(k).
\]

We initially set 40 as the number of rules and the overlapping factors, \( \lambda_i = 0.7 \) for \( i = 1 \) to 2, are chosen in on-line initialization method. Suppose one epoch of learning takes 200 time points, the supervised learning is continued for 200 epochs of training. The similarity measure is applied to the input fuzzy sets obtained in the preceding process to combine the similar fuzzy rules and term nodes. Table 1 shows the number of rules after rule combination under reference values \( \gamma_c = 0.9 \) and \( \gamma_r = 0.9, \ldots, 0.1 \). To show the feasibility of doing rule combination by proposed similarity measure for fuzzy rules, we take \( \gamma_r = 0.1 \) as an example in the following process. Furthermore, the reference value of the similarity degree \( \gamma_c \) was set for term node combination. Table 2 shows the number of term nodes after term node combination under different values of \( \gamma_c \). For brevity, let us only consider \( \gamma_c = 0.4 \) in the following simulation and comparison.

The final rules are obtained after supervised learning is applied again for just 30 epochs of learning. The outputs of the system and the FNNS with the final fuzzy rules are shown in Fig. 4. As shown in Fig. 4, the result is still desirable. Some comparisons between the proposed FNNS and other identifiers are summarized these comparisons.

<table>
<thead>
<tr>
<th>( \gamma_r )</th>
<th>0.9</th>
<th>0.8</th>
<th>0.7</th>
<th>0.6</th>
<th>0.5</th>
<th>0.4</th>
<th>0.3</th>
<th>0.2</th>
<th>0.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of rules</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>39</td>
<td>39</td>
<td>38</td>
<td>34</td>
<td>29</td>
<td>22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \gamma_c )</th>
<th>0.9</th>
<th>0.8</th>
<th>0.7</th>
<th>0.6</th>
<th>0.5</th>
<th>0.4</th>
<th>0.3</th>
<th>0.2</th>
<th>0.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of term nodes for ( y(k) )</td>
<td>20</td>
<td>20</td>
<td>19</td>
<td>18</td>
<td>17</td>
<td>15</td>
<td>13</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>Number of term nodes for ( y(k - 1) )</td>
<td>21</td>
<td>21</td>
<td>20</td>
<td>18</td>
<td>18</td>
<td>16</td>
<td>13</td>
<td>12</td>
<td>11</td>
</tr>
</tbody>
</table>

--- output of the original system -.-. : output of the FNNS

Figure 4: Outputs of the original system and the FNNS (under final fuzzy rules).
### Table 3: Comparisons of the proposed system with other identifiers.

<table>
<thead>
<tr>
<th>Model Name</th>
<th>Final Number of Rules</th>
<th>Final Number of Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narendra et al.[10]</td>
<td>$N_{20,10,1}$</td>
<td>250</td>
</tr>
<tr>
<td>L.X. Wang[12]</td>
<td>40</td>
<td>200</td>
</tr>
<tr>
<td>proposed FNNS</td>
<td>22</td>
<td>88</td>
</tr>
</tbody>
</table>

### VI. CONCLUSIONS

In this paper, a fuzzy neural network system called the FNNS has been presented for implementing fuzzy inference systems. The main purpose of the FNNS is to produce a simpler fuzzy inference system with fewer fuzzy logical rules and adjustable parameters, which will be more efficient and useful in practical applications. In order to accomplish this purpose, we propose a fuzzy similarity measure for fuzzy rules to eliminate redundant fuzzy logical rules and apply a fuzzy similarity measure for fuzzy sets to combine similar input linguistic term nodes. We also derive a new and effective on-line initialization method for choosing the initial parameters of the FNNS. The simulation shows that the FNNS indeed yields simpler and more efficient results.

### References


