Adaptive Membership Function Fusion and Annihilation in Fuzzy If–Then Rules

B.G. Song, R.J. Marks II, S. Oh, P. Arabshahi *, T.P. Caudell †, J.J. Choi‡

Abstract— The parameters of the input and output fuzzy membership functions for fuzzy If-Then min-max inferencing can be adapted using supervised learning applied to training data. Under the assumption that the inference surface is in some sense smooth, the process of adaptation can reveal overdetermination of the fuzzy system in two ways. First, if two membership functions come sufficiently close to each other, they can be fused into a single membership function. Second, if a membership function becomes too narrow, it can be deleted. In both cases, the number of If-Then rules is reduced. In certain cases, the overall performance of the fuzzy system can be improved by this adaptive pruning.

INTRODUCTION

The parameters of the input and output fuzzy membership functions for fuzzy If-Then inferencing can be adapted using supervised learning applied to training data [1-5]. The specific case of adaptation of min-max inferencing using steepest descent [6] has the advantage of adapting only those membership functions used in the fuzzy decision process for each training data input-output pair.

In the process of adapting, two membership functions may drift close together. If the underlying target surface which we wish to estimate is smooth, then the membership functions can be fused into a single membership function. Alternately, if a membership function becomes too narrow, it can be totally deleted. In either case, the fuzzy decision process is pruned. In artificial neural networks, pruning neurons from hidden layers can improve the performance of the neural network [7]. Likewise, the performance of fuzzy inference can be improved through the adaptation and pruning of membership functions. The number of If-Then rules is also correspondingly reduced.

ADAPTIVE TRAINING

Considered are fuzzy If-Then rules of the type

If x is X_i and y is Y_j, then z is Z_k

*The authors are with Dept. of Electrical Engineering, University of Washington, FT-10, Seattle, WA 98195.

†The author is with Boeing Computer Services, P.O. Box 24346, 7L-22, Seattle, WA 98124.

‡The author is with Boeing Computer Services, P.O.Box 24346, 6C-04, Seattle, WA 98124.

0-7803-0614-7/93$03.00 ©1993IEEE

where X_i, Y_j and Z_k are the linguistic variables corresponding to x, y and z. The possibility of the kth consequent is:

$$\alpha_k = \max_{S_k} \min \left[\mu_{X_i}(x), \mu_{Y_j}(y)\right].$$

Here μ_{X_i} and μ_{Y_j} are the input membership functions, and the set S_k is defined as:

$$S_k = \{i, j \mid X_i \text{ and } Y_j \text{ are antecedents of a rule with consequent } Z_k\}.$$

Assume that the center of mass of $\mu_{X_i}(x)$ is m_{X_i} and the dispersion (spread) of $\mu_{X_i}(x)$ is parameterized by σ_{X_i}. The parameter σ_{X_i} is also proportional to the area of $\mu_{X_i}(x)$. The membership functions $\mu_{Y_j}(y)$ and $\mu_{Z_k}(z)$ are likewise parameterized.

If the output membership functions are $\mu_{Z_k}(z)$, then the defuzzified output using the center of mass of the sum of weighted output membership functions is

$$o = \frac{\sum_k \alpha_k m_{Z_k} \sigma_{Z_k}}{\sum_k \alpha_k \sigma_{Z_k}}. \tag{1}$$

In [6], we outline a procedure whereby the parameters of the input and output membership functions can be adapted under the condition that the true output, t, is known for the the input (x, y). The error is

$$E = \frac{1}{2}(o-t)^2,$$

and the parameters are adjusted using steepest descent. For the parameter σ_{X_i}, for example, the adjustment is

$$\sigma_{X_i} \Leftarrow \sigma_{X_i} - \eta \frac{\partial E}{\partial \sigma_{X_i}},$$

where η is the step size. The partial derivative can be computed using error back-propagation which, for the problem under consideration, is presented in detail in [6].

Although we will use min-max inferencing, the procedure of membership function fusion and annihilation can be applied to other fuzzy inference methods, wherein, for example, alternate forms of defuzzification are used or intersections and unions other than min and max are employed [8, 9].

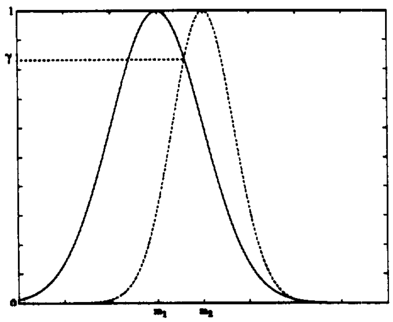


Figure 1: Illustration of the criterion for fusion. When two membership functions become sufficiently close so that the maximum of their intersection exceeds γ, then the two membership functions are fused into a single membership function.

Herein, we will assume all linguistic variables are scaled to the universe of discourse on the interval $[-1, 1]$. Gaussian membership functions of the form

$$\mu(x) = \exp\left[-\left(\frac{x-m}{\sqrt{2}\sigma}\right)^2\right],$$

will be used throughout.

Membership Function Fusion

Fusion of two membership functions occurs when they become sufficiently close to each other. Annihilation occurs when a membership function becomes sufficiently narrow. As illustrated in Fig. 1, two membership functions are fused when the supremum of their intersection exceeds a threshold, γ. If the means of the membership functions prior to fusion are m_1 and m_2, then the mean of the fused membership is set equal to the center of mass of the sum of the membership functions:

$$m_{\text{fusion}} = \frac{m_1\sigma_1 + m_2\sigma_2}{\sigma_1 + \sigma_2},$$

where σ_1 and σ_2 are the spread parameters of the two membership functions. Similarly, the spread of the fused function is obtained from

$$\sigma^2_{\text{fusion}} = \frac{\sigma_1^3 + \sigma_2^3}{\sigma_1 + \sigma_2}.$$

Membership fusion has a direct impact on the fuzzy decision process. To illustrate, consider Table 1. Here, N = negative, Z = near zero and P = positive. Assume that the membership functions for x corresponding to N and Z fuse. The two left most columns of the rule table are combined into one. A new linguistic variable, called NZ labels this column. It remains to specify the corresponding rules. When two adjacent rules are the same prior to fusing, the answer is simple. For example, since $X_i = N$ and Z both have Z as a consequent for $Y_j = Z$, the clear choice for the fused rule table for $X_i = NZ$ and $Y_j = Z$ is the consequent Z. For $Y_j = N$, however, there are different consequents when $X_i = N$ and $X_i = Z$. To determine the consequent for $X_i = NZ$ and $Y_j = N$ (marked '?' in Table 1), we chose to query the training data base. Specifically, training data was found where $(x, y) \approx (m_{NZ}, m_N)$. The value of the target, t, for this input pair is compared to the means of the existing output membership functions. The membership function having the closest mean is assigned as the consequent.

y \ x	N	Z	P
N	Z	P	P
Z	Z	Z	P
P	N	N	Z

$\Longrightarrow$

y \ x	NZ	P
N	?	P
Z	Z	P
P	N	Z

Table 1: Rule tables before (left) and after (right) fusion of two fuzzy membership functions of the variable x.

Output membership functions can also fuse. If, for example, the output Z fuses with N in the left hand rule table in Table 1, the resulting fused rule table will place NZs in the six boxes currently occupied with Zs or Ns.

Once fusion occurs, the membership functions are further adapted to the training data. Additional fusion or annihilation can follow.

Membership Function Annihilation

If the contribution of a fuzzy membership function becomes insignificant, then it can be annihilated. To illustrate, consider Fig. 2. The membership function $\mu_2(x)$ becomes insignificant with respect to the membership function, $\mu_1(x)$, when, for all x,

$$\sigma_1\mu_1(x) \geq \beta\sigma_2\mu_2(x)$$

where $\beta \geq 1$ parameterizes the degree of insignificance. High β corresponds to a severe criterion for annihilation. It is sufficient for the above criterion to hold only for $x = m_2$:

$$\sigma_1\mu_1(m_2) \geq \beta\sigma_2\mu_2(m_2) = \beta\sigma_2$$

The process is valid when the underlying target surface is smooth.

When an input membership function is annihilated, all rules using it are deleted from the fuzzy rule base. For

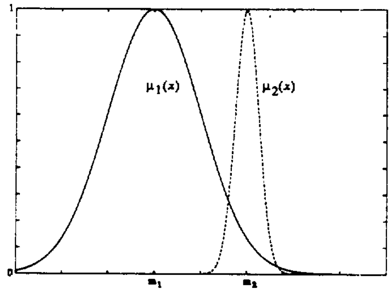


Figure 2: Illustration of the process of membership function annihilation. When the membership function, $\mu_2(x)$, becomes narrow with respect to an adjacent membership function, it can be annihilated.

y \ x	N	Z	P
N	Z	P	P
P	N	N	Z

Table 2: When the membership function for $Y_j = Z$ in the left table in Fig. 2 is annihilated, the rule table shown here results.

x \ y	1	2	3
1	1	2	1
2	2	3	2
3	1	2	1

Table 3: Target Rule Table for Example 1.

x \ y	1	2	3	4	5	6	7	8	9	10	11
1	1	1	1	2	3	3	3	2	1	1	1
2	1	1	2	3	3	3	3	3	2	1	1
3	1	2	2	3	3	4	3	3	2	2	1
4	2	3	3	4	4	5	4	4	3	3	2
5	3	3	3	4	5	5	5	4	3	3	3
6	3	3	4	5	5	5	5	5	4	3	3
7	3	3	3	4	5	5	5	4	3	3	3
8	2	3	3	4	4	5	4	4	3	3	2
9	1	1	1	2	3	3	3	2	1	1	1
10	1	1	2	3	3	3	3	3	2	1	1
11	1	2	2	3	3	4	3	3	2	2	1

Table 4: Rule Table for Example 1.

example, if the membership function corresponding to $Y_j = Z$ in the left hand rule table in Table 1 is annihilated, then the rule table after annihilation would be as shown in Table 2.

An output membership function can likewise be annihilated. In such a case, one of the remaining membership functions must take its place in the rule table. The choice, again, is made by a query to the training data base as was done for input membership function fusion.

After annihilation, the membership parameters can be further adapted using the training data. Additional annihilation and/or fusion might subsequently result.

EXAMPLES

We illustrate the process of membership function fusion and annihilation with two examples. The first is a proof of principle wherein convergence is to a solution known to be optimal. The second uses adaptation to fit a given target surface. We used the parameters $\beta = 2$ and $\gamma = 0.9$ for input membership functions and $\gamma = 0.95$ for the output. Iteration was performed until $\Delta E/E \approx 10^{-3}$. In cases where a membership function could either be fused or annihilated, annihilation was given priority.

1) Convergence to a Known Solution

In this example, the target membership functions shown in Fig. 3(a) were used. The target rule table is shown in Table 3. Using a universe of discourse on $[-1, 1]$, the membership functions are indexed from 1 for large negative numbers upward. The largest index corresponds to large positive numbers.

A total of 500 training data points were randomly generated from these target functions.

Overdetermined initialization is shown in Fig. 3(b) with a rule table shown in Table 4. Input membership functions are spaced evenly. Spacing of output membership functions is determined from a histogram of the training data target values. The histogram is divided into intervals of equal area. The number of intervals is chosen to be equal to the number of output membership functions. The means of the output membership functions are places at the boundaries of these intervals.

The result of the first steepest descent adaptation is shown in Fig. 3(c). Compare this to Fig. 3(d). The two left most membership functions for x (top plot) fuse. The third membership function for x is annihilated, etc. For the output, two membership functions are annihilated. The rule table becomes that shown in Table 5.

The membership functions in Fig. 3(d) are further trained. The result is shown in Fig. 3(e). Compare this to Fig. 3(f), where four input membership functions are annihilated. The results of Fig. 3(f) are adapted and converge to the result shown in Fig. 3(g). As can be seen in Fig. 3(h), two more input membership functions are annihilated. Further iteration yields Fig. 3(i). For y (middle plot), three membership functions fuse to two membership functions (see Fig. 3(j)). The fuzzy rule table corresponding toFig. 3(j) is as shown in Table 6.

The results in Fig. 3(j) are adapted to those shown in Fig. 3(k). Fusion occurs as shown in Fig. 3(l). Ad-

y / x	1	2	3	4	5	6	7	8	9	10
1	1	1	2	2	2	2	2	1	1	1
2	2	2	2	3	3	3	2	2	2	2
3	2	2	3	3	3	3	3	2	2	2
4	2	2	3	3	3	3	3	2	2	2
5	2	2	2	3	3	3	2	2	2	2
6	1	2	2	2	2	2	2	2	1	1
7	1	1	2	2	2	2	2	1	1	1
8	1	1	2	2	2	2	2	1	1	1

Table 5: Modified Table 4 after first steepest descent adaptation followed by fusion and annihilation.

y / x	1	2	3	4
1	1	2	2	1
2	2	3	3	2
3	2	3	3	2
4	1	2	2	1

Table 6: Table 4 after further adaptation, fusion and annihilation.

ditional adaptation results in the middle two membership functions for y (middle plot) shown in Fig. 3(m) to be graphically indistinguishable. They are fused in Fig. 3(n). The rule table is now exactly the target table in Table 3. The input membership functions are the same as in Fig. 3(a). The output membership functions are not the same; all defuzzifications from these membership functions though, are. Output membership functions $\{\mu_{Z_k}(x)\}$ will yield the same defuzzification as the membership functions $\{\mu_{Z_k}(x/\sigma)\}$ when defuzzification is performed as in Eq. 1.

2) Regression Fitting of a Surface

In this example, we assume a target surface of

$$\sin[\pi(x_1 + x_2)] \cos[\pi(x_1 - x_2)] .$$

The initial membership functions are shown in Fig. 4(a). A contour plot of the target is shown in Fig. 5(a). The first initialization is shown in Fig. 5(b). A total of ten steps of iteration followed by fusion and annihilation were required prior to convergence. The results are shown in Figs. 4(b) and 5(c). Convergence mean square error is shown in Fig. 6. Between odd and even steps (e.g., 3 and 4), error is reduced by steepest descent. Between the even and odd steps (e.g., 4 and 5) fusion and annihilation are applied, generally resulting in an increase in error.

The final rule table is shown in Table 7. The number of rules has been reduced from 441 (21^2) to 169 (13^2). The cardinality of the set of consequents has been reduced from 8 to 5.

y / x	1	2	3	4	5	6	7	8	9	10	11	12	13
1	3	4	4	3	2	2	3	4	4	3	2	2	3
2	4	5	5	4	3	3	4	5	5	4	3	3	4
3	4	5	5	4	3	3	4	5	5	4	3	3	4
4	3	4	4	3	2	2	3	4	4	3	2	2	3
5	2	3	3	2	1	1	2	3	3	2	1	1	2
6	2	3	3	2	1	1	2	3	3	2	1	1	2
7	3	4	4	3	2	2	3	4	4	3	2	2	3
8	4	5	5	4	3	3	4	5	5	4	3	3	4
9	4	5	5	4	3	3	4	5	5	4	3	3	4
10	3	4	4	3	2	2	3	4	4	3	2	2	3
11	2	3	3	2	1	1	2	3	3	2	1	1	2
12	2	3	3	2	1	1	2	3	3	2	1	1	2
13	3	4	4	3	2	2	3	4	4	3	2	2	3

Table 7: Final rule table for Example 2.

Acknowledgments

This work was supported in part by Boeing Computer Services and the Washington Technology Center.

References

[1] C.T. Lin and C.S.G. Lee, "Neural network based fuzzy logic control and decision system", *IEEE Trans. Computers*, vol. 40, pp.1320-1336, 1991.

[2] H. Nomura, I. Hayashi and N. Wakami, "A self-tuning method of fuzzy control by descent

method", *Proc. IEEE Int. Conf. Fuzzy Systems, FUZZ-IEEE '92*, San Diego, CA, March 8-12, 1992 (IEEE Press).

[3] Y. Hayashi, E. Czogala and J.J. Buckley, "Fuzzy neural controller", *Proc. IEEE Int. Conf. Fuzzy Systems, FUZZ-IEEE '92*, San Diego, CA, March 8-12, 1992 (IEEE Press).

[4] J.S.R. Jang, "Fuzzy controller without domain experts", *Proc. IEEE Int. Conf. Fuzzy Systems, FUZZ-IEEE '92*, San Diego, CA, March 8-12, 1992 (IEEE Press).

[5] L.X. Wang and J.M. Mendel, "Back propagation fuzzy system as nonlinear dynamic system identifiers", *Proc. IEEE Int. Conf. Fuzzy Systems, FUZZ-IEEE '92*, San Diego, CA, March 8-12, 1992 (IEEE Press).

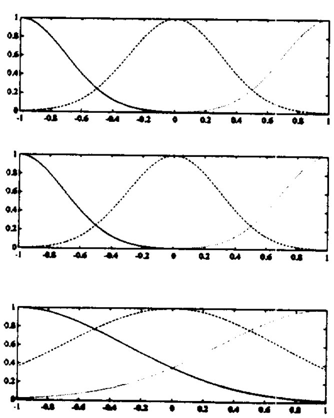
[6] R.J. Marks II, S. Oh, P. Arabshahi, T.P. Caudell, J.J. Choi, B.G. Song, "Steepest Descent Adaptation of Min-Max Fuzzy If-Then Rules", *Proc. Int. Joint Conf. Neural Networks*, Beijing, China, November 3-6, 1992 (IEEE Press).

[7] R. Reed, "Pruning Algorithms- A Survey", *IEEE Trans. Neural Networks*, - in press.

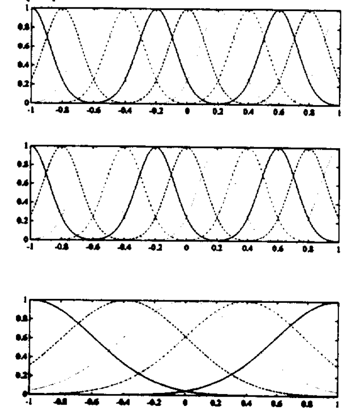
[8] G.J. Klir and T.A. Folger, **Fuzzy Sets, Uncertainty, and Information**, Prentice-Hall, 1988.

[9] T. Terano, K. Asai and M. Sugeno, **Fuzzy Systems Theory and its Applications**, Academic Press, 1992.

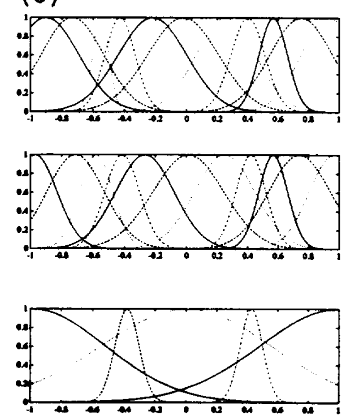
(a)



(b)



(c)



(d)

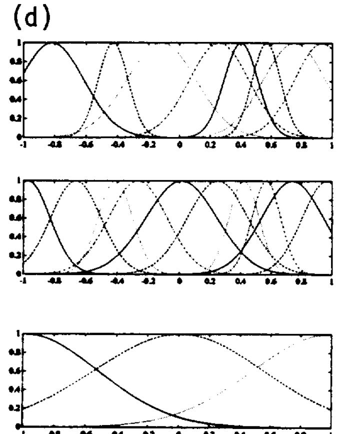


Figure 3: (a) Target membership functions for Example 1. The top, middle, and bottom plots are for μ_{X_i}, μ_{Y_j}, and μ_{Z_k} respectively. (b) Initial membership functions. (c-n) Evolution of the adaptation, fusion, and annihilation process.

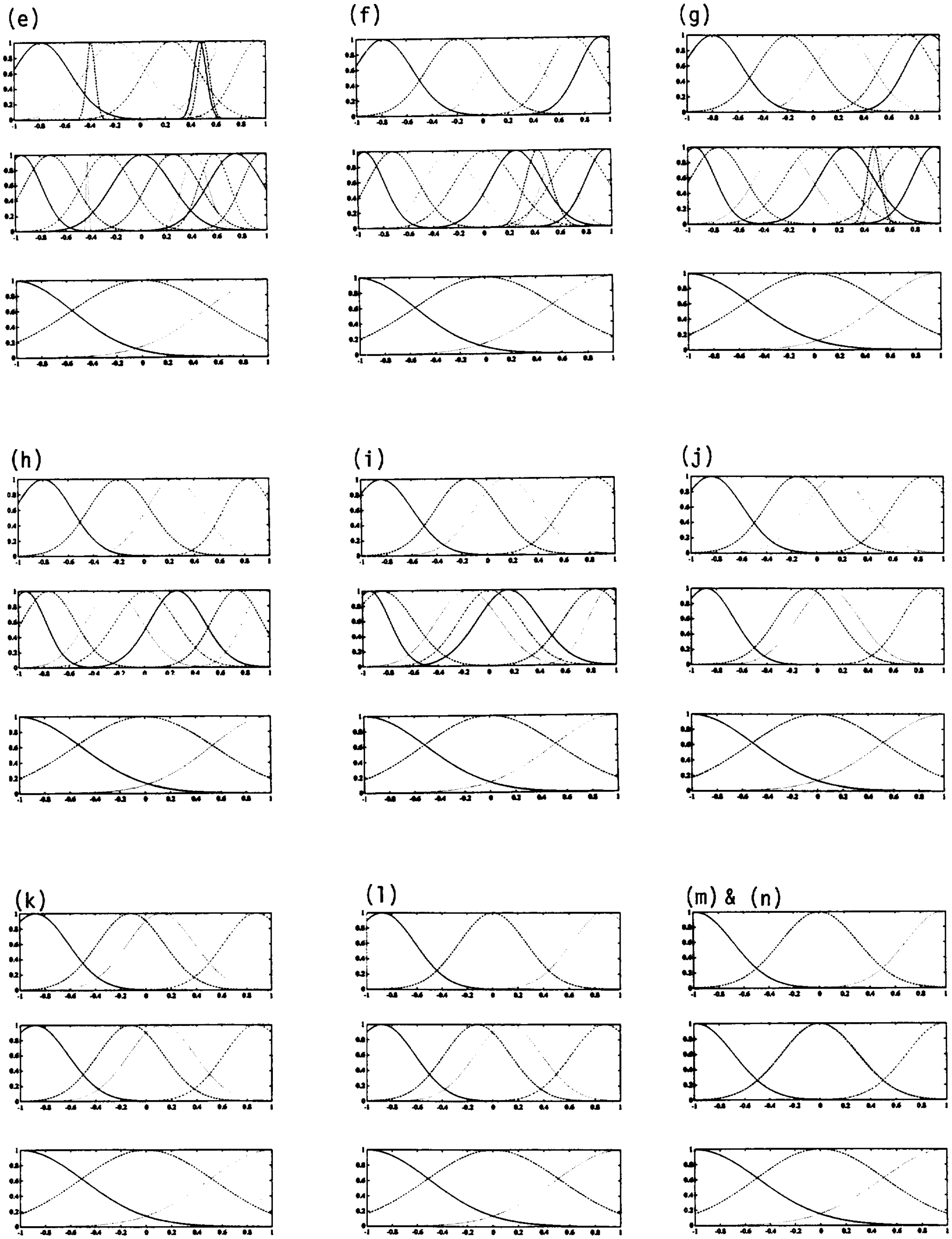


Figure 3: (continued).

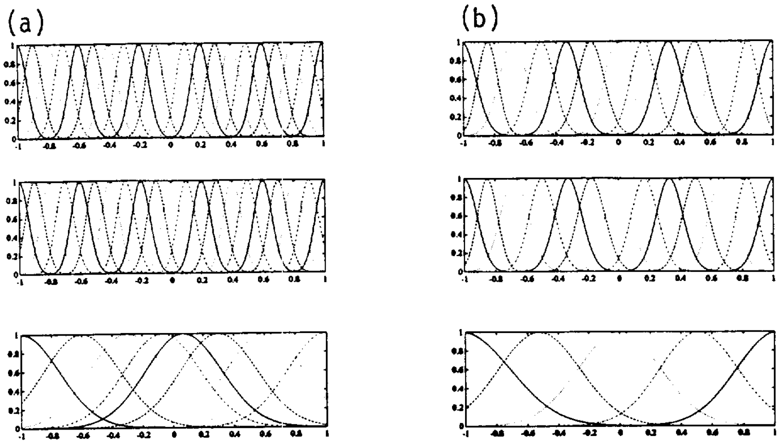


Figure 4: (a) Initial membership functions for Example 2. (b) Final membership functions for Example 2.

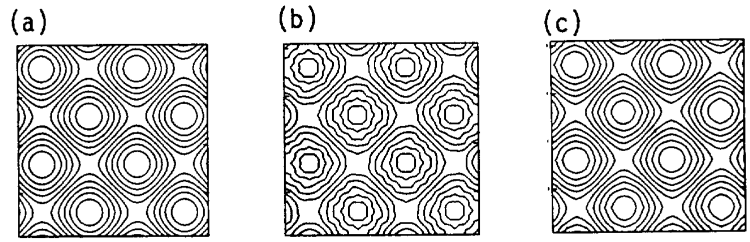


Figure 5: Contour plots of the (a) target (b) initialization, and (c) final result for Example 2.

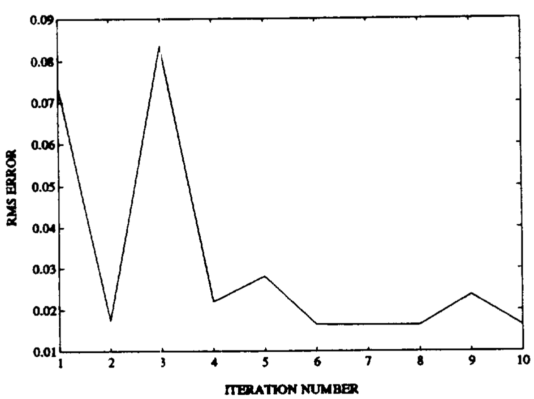


Figure 6: Convergence of the of the root mean square error for Example 2.