A Lossless Rule Reduction Technique for a Class of Fuzzy Systems

J.BALASUBRAMANIAM, C.JAGAN MOHAN RAO
Department of Mathematics and Computer Sciences,
Sri Sathya Sai Institute of Higher Learning
Prasanthi Nilayam, A.P.,
INDIA – 515134.

Abstract: - In this paper we propose a general framework of Fuzzy Systems. We also propose a novel rule reduction technique for a restricted class of Fuzzy Systems by combining the antecedents of the rules with same consequents. This rule reduction is lossless with respect to inference. A few examples from this class of Fuzzy Systems are given.

Key-words: - Fuzzy Systems, Rule Reduction, Residuated Implications, Fuzzy Inference.

1. Introduction

Following the wide spread usage of Fuzzy Systems, Rule Reduction has emerged as one of the most important areas of research in the field of Fuzzy Control. It is well known that an increase in the number of input variables and/or the number of membership functions in the input domains quickly lead to combinatorial explosion in the number of rules.

The several approaches taken towards rule reduction in Fuzzy Systems can be classified into the following categories: Selection of important rules that contribute to the inference, Elimination of redundant rules based on some criteria or Merger of rules that share some common property. Taking the first approach, rule reduction has been addressed in [1,2,3] using Genetic Algorithms and Evolutionary Methods, in [4,5,6] using Orthogonal Transformations. [15] is a good survey on the above methods. In [7], the author has employed a similar idea as that of ours, i.e., merging rules with similar consequents. In [8], Magne *et al*., use a similarity measure to merge rules. (See §5 for more details). But very little work has been done on rule reduction techniques that preserve the inference, i.e., the outputs of the original and the reduced rule bases are identical. This work proposes a novel rule reduction technique for a restricted class of Fuzzy Systems that preserves the inference.

2. A General Framework for Fuzzy Systems

2.1 Different Stages of a Fuzzy System

A Fuzzy System consists of the following 5 stages:

2.1.1Fuzzification:

In this step, the given crisp input a is fuzzified to get a fuzzy set $\tilde{X}$ on the corresponding input space, i.e., $a \to \tilde{X}$.

2.1.2 Matching:

The input fuzzy sets $(\tilde{X}_{1,} \tilde{X}_2, \ldots, \tilde{X}_n)$ are matched against their corresponding if-part sets of their input spaces in each of the rule antecedents in the Fuzzy System, i.e., $a_i^j = S(A_i^j, \tilde{X})$. (1)

2.1.3 Combining:

In a multi-antecedent fuzzy system, the various matching degrees a_i^j of the n input fuzzy sets to the antecedent of a fuzzy if-then rule is combined to

$$\mu_j = T(a_1^j, \ldots, a_n^j) . \quad (2)$$

2.1.4 Rule Firing:

The combined value μ_j fires the rule consequent or the output fuzzy set Y_j. In many models of fuzzy system, this Y_j is taken as its centroid y_j, i.e., $f_j = f(\mu_j, y_j)$. (3)

2.1.5 Aggregation or combining Inference:

The fired output fuzzy sets (or crisp sets) $f_j,\ j = 1,2,...,m.$; are then aggregated to obtain the final output fuzzy set $y = g(f_1, f_2, ..., f_m)$. (4)

2.2 Different Models of Fuzzy System in the literature:

Let us consider the following system of m fuzzy if-then rules:

R_1 : If X_1 is A_1^1 and … X_n is A_n^1 Then Y is B_1

R_j : If X_1 is A_1^j and … X_n is A_n^j Then Y is B_j

R_m : If X_1 is A_1^m and … X_n is A_n^m Then Y is B_m

where $A_i^j \in F(X_i)$ for $i = 1,2,...,n$ are the input fuzzy sets over the n input domains $X_1, X_2, \ldots, X_n$ and $B_j \in F(Y)$ for $j = 1,2,...,m$ are the output fuzzy sets over the single output domain Y. The following are the two most widely used models of fuzzy systems.

2.2.1 Takagi – Sugeno Fuzzy System:

$$F(x) = \frac{\sum_{j=1}^{m} \mu_j b_j}{\sum_{j=1}^{m} \mu_j} \qquad (5)$$

where we use Singleton Fuzzification. For the input $X = (x_1, x_2, ..., x_n)$, the matching values are given by $a_i^j = A_i^j \wedge \tilde{X}_i$ or $A_i^j \cdot \tilde{X}_i = A_i^j(x_i)$. The combined value of the multi-antecedent if-part is given by $\mu_j = a_1^j \cdot a_2^j \cdot ... \cdot a_n^j = A_1^j(x_1) \cdot A_2^j(x_2) \cdot ... \cdot A_n^j(x_n)$, and b_j is the centroid of the output fuzzy sets $B_j,\ \ j = 1,2,...,m.$

2.2.2 Mamdani Fuzzy System:

The output fuzzy set B given by μ_b is as follows:

$$\mu_b(w) = \bigvee_{j=1}^{m} f_j = \bigvee_{j=1}^{m} \bigwedge_{i=1}^{n} (a_i^j \wedge \mu_{b_j}(w)) \qquad (6)$$

where $a_i^j = A_i^j \wedge \tilde{X}_i$, $\mu_j = \bigwedge_{i=1}^{n} a_i^j$, $f_j = \mu_j \wedge \mu_{b_j}(w)$ and $\mu_{b_j}(w)$ is the output fuzzy set B_j of the j^{th} rule.

2.2.3 Kosko's Standard Additive Model (SAM):

The output is given by

$$F(x) = \frac{\sum_{j=1}^{m} w_j a_j(x) v_j c_j}{\sum_{j=1}^{m} w_j a_j(x) v_j}, \qquad (7)$$

where $a_j(x) = a_1^j(x_1) \cdot ... \cdot a_n^j(x_n)$, v_j and c_j are the volume and centroid of the output fuzzy set B_j of the j^{th} fuzzy if-then rule and w_j are the rule weights. Letting $\mu_j = w_j \cdot a_j(x) \cdot v_j$ we get the Takagi – Sugeno fuzzy system.

2.3 A General Framework for Fuzzy Systems:

From the above two sub-sections 2.1 and 2.2, it appears that the different stages can be mapped to different functions capturing the actions performed at every stage. To this end, we do not consider 'fuzzification' stage since a crisp input to the fuzzy system can be thought of as a singleton – fuzzified input fuzzy set.

Then the different stages and the corresponding mappings capturing their actions can be given by:

Matching: $a_i^j : S(A_i^j, \tilde{X}_i) : F(X_i) \times F(X_i) \to [0,1]$

Combining: $\mu_j : \mu(a_1^j, ..., a_n^j) : [0,1]^n \to [0,1]$

Firing: $f_j : f(m_1, ..., m_m)$

Aggregation: $g = g(f_1, ..., f_m)$

The corresponding functions for S, μ, f and g for the different models of fuzzy systems are tabulated in Table 1.

Table 1. S, μ, f and g for the different models of fuzzy systems

Name / Type	S	μ	f	g	Fuzzification
Takagi – Sugeno	$\wedge$	Product	Product	Weighted Average	Singleton
Mamdani – Type I	$\wedge$	$\wedge$	Product/ Minimum	$\vee$	Any
Mamdani – Type II	$\vee$	$\vee$	Product/ Minimum	$\wedge$	Any
Kosko's SAM	$\wedge$	Product	Product	Weighted Average	Singleton

3. A Novel Rule Reduction for a Restricted class of Fuzzy Systems:

3.1 Inference in MISO – Fuzzy Systems under g, f, m and S

A general Multi Input Single Output (MISO) – fuzzy system is given as follows:

R_1 : If X_1 is A_1^1 and … X_n is A_n^1 Then Y is B_1

R_j : If X_1 is A_1^j and … X_n is A_n^j Then Y is B_j

R_m : If X_1 is A_1^m and … X_n is A_n^m Then Y is B_m

where $A_i^j \in F(X_i)$ for $i = 1,2,...,n$ and $B_j \in F(Y)$ for $j = 1,2,...,m$. Then the general inference in the absence of any input is given by:

$$g\{f[\mu(A_1^1, A_2^1, ..., A_n^1), B_1], f[\mu(A_1^2, A_2^2, ..., A_n^2), B_2], ..., f[\mu(A_1^m, A_2^m, ..., A_n^m), B_m]\} \quad (9)$$

where μ is any antecedent combiner, f is any function representing the rule firing and g is the aggregation.

3.2 Rules with the same consequents:

More often than not, the number of fuzzy sets (membership functions) defined on the single output domain, say r, is typically much less than the number of rules m, i.e., r << m. To eliminate this redundancy, we propose a new type of rule reduction where the rules with same consequents but different antecedents are merged into a single rule. Then we will have only as many rules as there are output membership functions, in fact only those that are part of the original fuzzy system.

The issue involved here is that despite the merging of the above rules, there should be no loss of inference, i.e., the output that would have been obtained for a given input to the original model should be the same as that of the reduced model for the same input.

This necessitates the functions g, f, m and S to possess some properties. These are explored in the next sub section.

3.3 The Restrictions on g, f, m :

Let us consider a MISO – fuzzy system. Without loss of generality, let us take a 2-input 1-output fuzzy system, where X_1 and X_2 are the input domains and Y the output domain. Again, without loss of generality, let us consider the fuzzy system with the following rules:

$$\left.\begin{aligned} R_1 &: A_1, B_1 \to C \\ R_2 &: A_2, B_2 \to C \\ R_3 &: A_3, B_3 \to D \end{aligned}\right\} \quad (10)$$

Then the inference under g, f, and μ in the absence of any input to the fuzzy system is given by

$$g\{f[\mu(A_1, B_1), C], f[\mu(A_2, B_2), C], f[\mu(A_3, B_3), D]\} \quad (11)$$

Since we need to merge the rules R_1 and R_2 having the same consequents, we do the following:

From (11) we have,

$$g\{f[\mu(A_1, B_1) \circ_g \mu(A_2, B_2), C], f[\mu(A_3, B_3), D]\} \quad (12)$$

$$= g\{f[\mu(A_1 \circ_\mu A_2, B_1 \circ_\mu B_2), C], f[\mu(A_3, B_3), D]\} \quad (13)$$

3.3.1 Combining 'Combined' Values:

From (11) we obtain (12) by composing the antecedents $\mu(A_1, B_1)$ and $\mu(A_2, B_2)$ of the rules R_1 and R_2 having the same consequents. This introduces a new operator $\circ_g : I \times I \to I$ such that

$$g[f(A, C), f(B, C)] = f(A \circ_g B, C). \quad (14)$$

Thus the function g should possess a corresponding operator $\circ_g$ such that (14) is satisfied. Also g should be associative.

3.3.2 Combining Fuzzy Sets on the same Domain:

From (12) we obtain (13) by combining fuzzy sets that are defined on the same input domain, i.e., A_1, $A_2 \in F(X_1)$ and B_1, $B_2 \in F(X_2)$. To this end, we introduce another operator $\circ_\mu : I \times I \to I$ such that

$$\mu(A_1, B_1) \circ_g \mu(A_2, B_2) = \mu(A_1 \circ_\mu A_2, B_1 \circ_\mu B_2) \quad (15)$$

In the absence of any input to the fuzzy system, the restrictions applied so far are:

1. g is associative.
2. $g[f(A,C), f(B,C)] = f(A \circ_g B, C)$. (14)
3. $\mu(A_1, B_1) \circ_g \mu(A_2, B_2)$
$= \mu(A_1 \circ_\mu A_2, B_1 \circ_\mu B_2)$ (15)

The above technique is applied when g and μ obey the given equations and possess functions $\circ_g$ and $\circ_\mu$.

3.4 The Reduced Rule Base and Inference in the presence of inputs:

In the above discussion, the function S has not figured. This is because one of the parameters for S is the current input. Let us consider the inference in the above MISO – fuzzy system in the presence of input, say $\tilde{X} = (A, B)$ where $A \in F(X_1)$, $B \in F(X_2)$. The MISO inference can be given as:

$$g\{f[\mu(S(A_1,A),S(B_1,B)),C],\ f[\mu(S(A_2,A),S(B_2,B)),C],\ f[\mu(S(A_3,A),S(B_3,B)),D]\} \quad (16)$$

$$= g\{f[\mu(S(A_1,A),S(B_1,B)) \circ_g \mu(S(A_2,A),S(B_2,B)),C],\ f[\mu(S(A_3,A),S(B_3,B)),D]\} \quad (17)$$

$$= g\{f[\mu\{S(A_1,A) \circ_\mu S(A_2,A), S(B_1,B) \circ_\mu S(B_2,B)\},C],\ f[\mu(S(A_3,A),S(B_3,B)),D]\} \quad (18)$$

$$= g\{f[\mu\{S(A_1 \circ_s A_2, A), S(B_1 \circ_s B_2, B)\},C],\ f[\mu(S(A_3,A),S(B_3,B)),D]\} \quad (19)$$

In the presence of an input $\tilde{X} = (\tilde{A}_1, \ldots, \tilde{A}_n)$, the matching fit values $a_i^j = S(A_i^j, \tilde{A}_i)$ are calculated and thus the function S features in (16) which is otherwise another form of (11).

We obtain (17) from (16) by using (14) and (18) from (17) by applying (15). Since in the reduced rule base A_1 and A_2 are not separately accessible, we go a step further and combine the fuzzy sets on the antecedents on the same domain. To enable us to perform this, we introduce a new operator $\circ_s : F(X) \times F(X) \to I$ i.e., $I \times I \to I$ such that

$$S(A_1, A) \circ_\mu S(A_2, A) = S(A_1 \circ_s A_2, A) \quad (20)$$

Observing (13) and (19) a little closer, we see that for the inference to hold even after the rule reduction, we need

$$\circ_s \equiv \circ_\mu \quad (21)$$

and thus S and μ are related. Thus the only properties g, f, μ and S should posses are :

1. g is associative.
2. $g[f(A,C), f(B,C)] = f(A \circ_g B, C)$. (14)
3. $\mu(A_1, B_1) \circ_g \mu(A_2, B_2)$
$= \mu(A_1 \circ_\mu A_2, B_1 \circ_\mu B_2)$ (15)
4. $S(A_1, A) \circ_\mu S(A_2, A) = S(A_1 \circ_\mu A_2, A)$, (20)

since $\circ_s \equiv \circ_\mu$

The above equations are the well-established Aggregation Equations [10]. [11-14] give a good coverage on general bisymmetry equations.

4. Examples of a few Fuzzy Systems from the above class

4.1 Mamdani-type models with Residuated Implications

For through out this section we will consider, without any loss of generality, the previous 2-input 1-output fuzzy system with 3 rules as given in § 3.3:

$$\begin{aligned} A_1, B_1 &\to C \\ A_2, B_2 &\to C \\ A_3, B_3 &\to D \end{aligned} \quad (10)$$

A Residuated Implication [9] $I : [0.1] \times [0,1] \to [0,1]$ is obtained as the residuation of a binary operator, in our case a t-norm, $t : [0.1] \times [0,1] \to [0,1]$, such that

$$a\ t\ b \le c \Leftrightarrow a \le b\ I\ c,\ \forall a, b, c \in [0,1].$$

The pair (t,I) is called the adjoin couple.

Some of the well-known R-implications and their corresponding t-norms are given in Table 2.

Table 2. Some of the well-known R-implications and their corresponding t-norms

Name	t(a,b)	I(a,b)
Lukasiawicz	max(0,a+b-1)	min(1,1-a+b)
Mamdani	min (a,b)	$a \to b = \begin{cases} 1, \text{ if } a \le b \\ b, \text{ otherwise} \end{cases}$
Larsen	a.b	$a \to b = \begin{cases} 1, \text{ if } a \le b \\ \frac{b}{a}, \text{ otherwise} \end{cases}$

Since $L = ([0,1], \wedge, \vee, t, \to)$ forms a Linearly Ordered Residuated Lattice, we have the following properties of L:

a) $\bigvee_i (a_i \to c) = (\bigwedge_i a_i) \to c$ (22)

b) $\bigwedge_i (a_i \to c) = (\bigvee_i a_i) \to c$ (23)

4.1.1 Mamdani Model – Type I

From Table 1, we know that for the Mamdani model of type 1, we have $g = \vee, \mu = \wedge$ and $S = \wedge$. Taking f as an R – implication, denoted $f = \to$, we have from (9), in the absence of any external input to the fuzzy system in (10)

$$[(A_1 \wedge B_1) \to C] \vee [(A_2 \wedge B_2) \to C] \vee [(A_3 \wedge B_3) \to D]$$

$$= \{[(A_1 \wedge A_2) \wedge (B_1 \wedge B_2)] \to C\} \vee [((A_3 \wedge B_3) \to D] \quad (24)$$

from (22) and grouping A_i's and B_i's. Also (24) is

$$= [(A_1^* \wedge B_1^*) \to C] \vee [((A_3 \wedge B_3) \to D] \quad (25)$$

Thus with $\circ_g = \wedge$ and $\circ_\mu = \wedge$ we have the fuzzy system in (10) with 3 rules reduced to a fuzzy system with 2 rules (25) without any loss of inference. Thus we have $g = \vee, \mu = \wedge, S = \wedge, f = \to, \circ_g = \wedge$ and $\circ_\mu = \wedge$.

In the presence of an input, say $\tilde{X} = (A, B)$, we have:

$$[(S(A_1, A) \wedge S(B_1, B)) \to C] \vee [(S(A_2, A) \wedge S(B_2, B)) \to C] \vee [(S(A_3, A) \wedge S(B_3, B)) \to D] \quad (26)$$

Since in the above model, S = ∧, we have, by substituting for S in (26)

$$[((A_1 \wedge A) \wedge (B_1 \wedge B)) \to C] \vee [((A_2 \wedge A) \wedge (B_2 \wedge B)) \to C] \vee [((A_3 \wedge A) \wedge (B_3 \wedge B)) \to D] \quad (27)$$

From (27) by applying (14), (15) and (20) coupled with the fact that $\circ_s \equiv \circ_\mu = \wedge$ we obtain

$$[(S(A_1 \wedge A_2, A) \wedge S(B_1 \wedge B_2, B)) \to C] \vee [(S(A_3, A) \wedge S(B_3, B)) \to D] \quad \langle \text{since } S = \wedge \rangle$$

$$= [(S(A_1^*, A) \wedge S(B_1^*, B)) \to C] \vee [(S(A_3, A) \wedge S(B_3, B)) \to D] \quad \text{QED}$$

4.1.2 Mamdani Model – Type II

In the above Mamdani Model, by replacing $g = \wedge, \mu = \vee$ and $S = \vee$ and retaining f to be an R – implication, it can similarly be shown using (23) that rule reduction of the proposed type is possible with $\circ_g = \vee$ and $\circ_s \equiv \circ_\mu = \vee$.

5. Other works along this line

In [7] the author employs the same idea, that of merging the rules with identical consequents. To this end, the author has defined new complimentation and also CNF Union (∪) and Intersection (∩) operations, that do not satisfy all the properties of S- and T-norms, respectively. The 'goodness' of the inference obtained by employing CNF ∪ and CNF ∩ are not discussed. Whether the merger envisaged is lossless in terms of inference is not addressed. Also rules with more than 1 antecedent have not been dealt with.

In [8], M.Setnes *et al.*, use a similarity measure to merge rules with fuzzy antecedents and/or consequents that are similar to each other above a specified threshold. Their main stated intention is the reduction in number of fuzzy sets used in the model. Again the issue of preservation of inference between the reduced rule base and the original fuzzy system is not addressed.

6. Advantages and Limitations

6.1 Advantages

+ The inference obtained from the original fuzzy rule base is preserved.

+ The method works for any type of membership functions.

+ Computationally efficient since there are only as many rules as the number of fuzzy sets that featured in the original rule base.

6.2 Limitations

- The approximation capability of the proposed restricted class of Fuzzy Systems is yet to be established.

- Merging of rules, in some cases, may turn out to be computationally intensive.

- In some instances, the above method may even increase the number of fuzzy sets defined on different input domains and thus may consume more memory.

7. References

[1] H.Ishibuchi, K. Nozaki, N.Yamamoto and H.Tanaka, " Selecting fuzzy if-then rules for classification problems using genetic algorithms", *IEEE Trans. on Fuzzy Syst.*, Vol. 3, 1995, pp. 260 – 270.

[2] C.L.Karr, " Applying genetics to fuzzy logic", *AI Expert*, Vol. 6, 1991, pp. 38 – 43.

[3] M.A.Lee and H.Takagi, " Integrating design stages of fuzzy systems using genetic algorithms", in *Proc. 2nd IEEE Conf. Fuzzy Systems 1993*, pp. 612 – 617.

[4] G.C.Mouzouris and J.M.Mendel, " Designing fuzzy logic systems for uncertain environments using a singular-value-QR decomposition method", in *Proc. 5th IEEE Conf. Fuzzy Systems 1996*, New Orleans, LA, pp. 295 – 301.

[5] L.X.Wang and J.M.Mendel, " Fuzzy basis functions, universal approximation and orthogonal least squares learning", *IEEE Trans. on Neural Networks*, Vol. 3, 1993, pp. 807 – 814.

[6] John Yen and Liang Wang, "Simplifying fuzzy rule-based models using orthogonal transformation methods", *IEEE Trans. on Syst. Man and Cyber. Part B*, Vol. 29, No.1, 1999, pp. 13 – 24.

[7] L. T. Koczy, " Fuzzy if…then rule models and their transformation into one another", *IEEE Trans. on Syst. Man and Cyber. Part A*, Vol. 26, No.5, 1996, pp. 621 – 637.

[8] Magne Setnes, R.Babuska, U.Kaymak and H.R. van Nauta Lemke, " Similarity measures in fuzzy rule base simplification", *IEEE Trans. on Syst. Man and Cyber. Part B*, Vol. 28, No.3, 1998, pp. 376 – 386.

[9] Esko Turuneen, " Algebraic Structures in Fuzzy Logic", *Fuzzy Sets and Systems* Vol. 52, 1992, pp. 183 – 188.

[10] M. A. Taylor, "The Generalized Equation of Bisymmetry: Solutions Based on Cancellative Abelian Monoids", *Aequationes Math.* Vol.57 1999, 288 - 302.

[11] M. A. Taylor, "On the Equation of Generalized Bisymmetry with Outer Functions Injective in Each Argument", *Aequationes Math.*,to appear.

[12] J. Aczel, "Bisymmetry and Consistent Aggregation: Historical Review and Recent Results", in *Choice, Decision, and Measurement (Morley, A. A. J. ed.)*, Lawrence Erlbaum Assoc. Publ., Mahwah, NJ, 1997, pp. 225 - 233.

[13] J. Aczel and Gy. Maksa, "Solution of the Rectangular m x n Generalized Bisymmetry Equation and of the Problem of Consistent Aggregation", *J. Math. Anal. Appl.* Vol. 203, 1996, pp. 104 - 126.

[14] J. Aczel, Gy. Maksa, and M. A. Taylor, "Equations of Generalized Bisymmetry and of Consistent Aggregation: Weakly Surjective Solutions Which May Be Discontinuous at Places", *J. Math. Anal. Appl.* Vol. 214, 1997, pp. 22 - 35.

[15] Magne Setnes and R.Babuska, "Rule Base Reduction: Some Comments on the Use of Orthogonal Transforms", *IEEE Trans. on Syst. Man and Cyber. Part C*, Vol. 31, No.2, 2001, pp. 199 – 206.