Conditions for Inference Invariant Rule Reduction in FRBS by combining rules with identical consequents

Jayaram Balasubramaniam

Department of Mathematics and Computer Sciences,
Sri Sathya Sai Institute of Higher Learning
Prasanthi Nilayam, A.P. -515134, India
jbala@ieee.org

Abstract: 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. Many rule reduction methods have been proposed in the literature and can be broadly classified into Lossless *or* Lossy *with respect to the inference, based on whether the outputs of the original and the reduced rule bases are identical or not. In a typical Multi-Input-Single-Output fuzzy system the number of rules far exceeds the number of fuzzy sets defined on the output domain. This suggests that the rule base can be partitioned into sets of rules, each set being mapped to a single consequent fuzzy set. In this paper, we investigate the conditions on the inference operators employed in a fuzzy system that enable "lossless" merging of rules with identical consequents.*

After briefly surveying the many techniques that have been proposed towards reducing the number of rules, we propose a general framework for Inference in Fuzzy Systems and then propose some sufficiency conditions on this general framework that give us a class of Fuzzy Systems that allow lossless rule reduction of the type mentioned above. We then explore these conditions in the setting of Fuzzy Logic. We find that R- and S-implications play a very critical role. We give examples from the above class of Fuzzy Systems. In this study we apply the above technique only on rules whose antecedents and consequents are fuzzy sets.

Keywords: Fuzzy Systems, Rule Reduction, Residuated Implications, Strong 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 a combinatorial explosion in the number of rules. On the other hand the number of output / consequent fuzzy sets remains a constant and is usually far less than the number of rules. This suggests that the rule base can be partitioned into sets of rules, each set being mapped to a single consequent fuzzy set. Thus the rules, though with different antecedents, but with identical consequents can be merged into a single rule. But such merger of rules, though reduces the number of rules may not be *lossless*, i.e., the inference obtained from the original rule base and the reduced rule base for a given input

may not be identical. In this paper, we investigate the conditions on the inference operators employed in a fuzzy system that enable "lossless" merging of rules with identical consequents. This provides us with a class of Fuzzy Systems in which the antecedents of the rules with identical consequent can be combined to reduce the number of rules in an inference invariant manner.

In section 2 we give a brief survey of the various rule reduction techniques proposed in the literature. In section 3 we propose a general framework for Inference in Fuzzy Systems and in section 4 we give sufficiency conditions on the inference framework that ensure lossless rule reduction of the type mentioned above. In section 5 we explore each of these conditions in the setting of Fuzzy Logic. In section 6 we give a few examples from the above class of SISO Fuzzy Systems that satisfy the above sufficiency conditions.

2 Rule Reduction as an Issue

2.1 Rule Reduction Techniques in the Literature

For an n-input Multi-Input Single-Output (MISO) fuzzy system, with n_i membership functions defined on each of the input domains $X_i (i = 1, 2, \ldots, n)$, we have $m = n_1 \times n_2 \times \ldots \times n_n = \prod_{i=1}^{n} n_i$ rules. Thus an increase in the number of input variables and/or the number of membership functions in the input domains quickly lead to a 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 significantly to the inference.

- Elimination of redundant rules based on some criteria.

- Merger of rules that share some common property.

2.1.1 Rule Reduction while Building a Fuzzy Rule Base

While trying to build a minimal fuzzy system, the authors in [52, 63] have employed Genetic Algorithm (GA) or GA-type optimisation to eliminate redundant rules and/or identify important or significant rules.

In [44] the authors have converted a linear fuzzy system in which the growth of the parameters with respect to inputs is exponential to an equivalent non linear fuzzy system in which their growth is linear. Works have also appeared that reduce the number of rules by reducing the number of input variables through Mathematical Fusion or through Symbolic Fusion, which involves the use of multi-dimensional fuzzy sets. In [70] a fuzzy binary box tree data structure has been proposed. In [43] the authors have designed a Fuzzy Logic Controller (FLC) based on Variable Structures techniques to be assured of Stability. They have reduced the number of rules from m^n to mn, where there are n input domains and m fuzzy sets on each domain.

2.1.2   Rule Reduction in an existing Fuzzy Rule Base

Towards reducing the number of rules in an existing fuzzy rule base, L.T. Koczy and Hirota [51] reduced a dense rule base to sparse rule base, containing the essential information in the original rule base, and all other rules were replaced by the Interpolation Algorithm that can recover them to a certain accuracy prescribed before reduction.

Following the Selection of significant rules or elimination of redundant rules, Rule Reduction has been addressed in [46,48,52] using GA and Evolutionary Algorithms, in [61,79,82] using Orthogonal Transformations, in [13] using Singular Value Decomposition, in [72] using Linear Matrix Inversion. [66] employs a Similarity Measure to prune the rules.

In [67] the authors 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.

In cases where coupling effects between different inputs are small, the design of an MISO fuzzy system has been reduced to that of designing a set of SISO fuzzy systems, in a decentralised fashion, each SISO fuzzy system being designed for a pair of input-output variables. Many approaches based on the approximation or decomposition of multi-dimensional fuzzy relations into two-dimensional ones have been studied [19,47]. In [41] the conditions for reducing multi-dimensional fuzzy relations into two-dimensional ones are studied for systems using max-min composition operator. However, such approximation may lead to unsatisfactory results if some peculiarities of the process are neglected.

In hierarchical fuzzy controllers introduced in [62] the number of rules increase linearly with the number of system inputs, but the decision of where the different variables are to be put in the hierarchy is often a difficult process.

2.2   Need for Lossless Rule Reduction Techniques

Many of the rule reduction methods in the literature give rise to an approximation error, i.e., the inference obtained from the original rule base and that obtained from the reduced rule base may not be the same.

In [14] Baranyi et al, discuss the trade off between Approximation Accuracy and Complexity. See also [50] for a discussion on the trade off between computation time and precision. Thus the approximation accuracy achieved should not be sacrificed in the process of complexity reduction. All these necessitate a study on rule reduction techniques that are lossless with respect to inference.

2.2.1   Lossless Rule Reduction Techniques in the literature

A few of the rule reduction techniques that are lossless are listed below. We define "lossless" in the sense that, the inference obtained from the original rule base and that obtained from the reduced rule base is identical.

In [64] an enhanced two-level Boolean Synthesis methodology is employed, where in, a given fuzzy rule with fuzzy connectives is mapped to a corresponding expression with boolean connectives, with each input fuzzy set being given a label. The method seeks to reduce the number of connectives employed in the antecedents of the rule. In [45] the authors in order to apply Karnaugh maps for rule reduction represent the linguistic values on a domain as 0 or 1. Though the reduced rule base can infer "sensibly" even if the original rule base were incomplete, if the output is identical in rules where one or more antecedents are different the method does not merge these rules and thus the rule reduction is incomplete.

In [22] the authors represent a Fuzzy System as a Fuzzy Inference Graph and try to minimise the number of nodes - rules - by a two step process. Again the rule reduction is incomplete since non-interacting antecedents are not combined even though their outputs are identical and also it is lossless only for the min implication operator.

In [23] the authors have proposed a novel, though much debated [24, 25, 30, 59], rule configuration called the Union Rule Configuration, wherein the growth in number of rules is only linear instead of exponential, but the proposed method is applicable only if there is monotonicity or ordering among inputs and membership functions and a one-one correspondence between input and output membership functions.

In [49] the author follows a similar approach as ours, that of merging rules with identical consequents by proposing new fuzzy operations where certain properties of regular fuzzy operations have been either relaxed or not imposed.

In [16] Baranyi et al., discuss both exact and non-exact reduction methods using Singular Value Decomposition methods, where by removing only the zero-Singular values one obtains lossless rule reduction and in the case when all Singular values below a threshold are discarded, the error bounds for some special types of fuzzy systems are also given in [11, 12, 80, 81]. Also [14, 15, 17, 18] discuss complexity reduction in Fuzzy Rule Bases using SVD. [13, 69, 82] give an excellent review of rule reduction techniques based on Orthogonal Transformations and discuss their goodness.

2.3   Our Approach towards Lossless Rule Reduction

The approach we take towards Lossless Rule Reduction is to merge rules with identical consequents even with different antecedents. We do not propose any new fuzzy operations to this end, but obtain some conditions that the different operators employed in a fuzzy inference system should satisfy. Also the final reduced rule base, obtained by employing our method, will contain only as many rules in the rule base, as there are output membership functions that featured in the original rule base. If there are n input domains and m input fuzzy sets in each domain the total number of rules that give a complete rule base is m^n. The best theoretical limit, so far, of a reduced rule base is mn [43]. With our method it reduces to k, where k is the number of output fuzzy sets that featured in the original rule base, and typically $k \ll m$.

3   A General Framework for Inferencing in Fuzzy Systems

First we give some preliminaries on Fuzzy Logic Operators that will be required in the rest of this work. As usual we will denote by I the unit interval $[0, 1]$.

3.1   Fuzzy Logic Operators

Definition 1 ([37], Definition 1.1, Pg 3)*. A Negation N is a function from I to I such that:*

- $N(0) = 1$; $N(1) = 0$;

- *N is non-increasing.*

A negation N is called strict if in addition N is strictly decreasing and continuous. A strong negation N is a strict negation N that is also involutive, i.e., $N(N(x)) = x$, $\forall x \in I$.

Definition 2 ([34] Definition 2.1 Pg 6)*. A t-norm T is a function from I^2 to I such that $\forall a$, b, $c \in I$,*

- $T(a, 1) = a,$

- $T(a, b) = T(b, a),$

- $T(a, T(b, c)) = T(T(a, b), c),$

- $T(a, b) \leq T(a, c)$ whenever $b \leq c.$

Definition 3 ([34] Definition 3.1 Pg 10)*. A t-conorm S is a function from I^2 to I such that $\forall a$, b, $c \in I$,*

- $S(a, 0) = a,$

- $S(a, b) = S(b, a),$

- $S(a, S(b, c)) = S(S(a, b), c),$

- $S(a, b) \leq S(a, c)$ whenever $b \leq c.$

Definition 4 ([34] Definitions 6.1 Pg 17 & 6.11 Pg 18)*. A t-norm T is said to be*

- ***Continuous*** *if it is continuous in both the arguments;*

- ***Archimedean*** *if for each $(x, y) \in (0, 1]^2$ there is an $n \in \mathbb{N}$ with $x_T^{(n)} < y$, where $x_T^{(n)} = T(\underbrace{x, \cdots, x}_{n \ times});$*

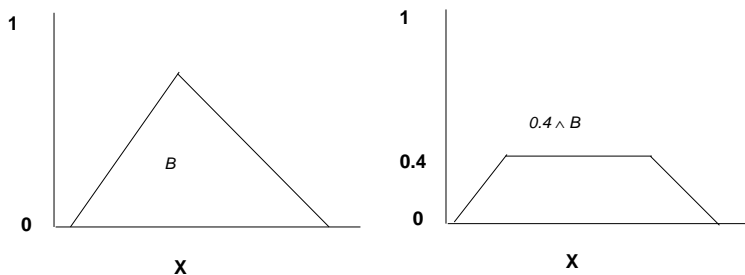- ***Strict*** *if T is continuous and stricly monotone, i.e., $T(x, y) < T(x, z)$ whenever $x > 0$ and $y < z;$*

Figure 1: Fuzzy Set B (left) and the fuzzy set $0.4 \wedge B$ (right)

- **Nilpotent** *if T is continuous and if each $x \in (0,1)$ is such that $x_T^{(n)} = 0$ for some $n \in \mathbb{N}$.*

Definition 5. *If $B : X \rightarrow I$, $a \in I$, and R is any binary operator on I, i.e., $R : I \times I \rightarrow I$, then $R(a, B)$ is a fuzzy set on X, i.e., $R(a, B) : X \rightarrow I$, defined as $R(a, B)(x) = R(a, B(x))$, $\forall x \in X$.*

Remark 1. *Thus R can also be seen as $R : I \times \widetilde{F}(X) \rightarrow \widetilde{F}(X)$- where $\widetilde{F}(X)$ denotes the set of all fuzzy sets on X. For example if $R(a, b) = \min(a, b)$ then in Figure 1 we have $B \in \widetilde{F}(X)$ and $R(0.4, B) = \min(0.4, B) = 0.4 \wedge B \in \widetilde{F}(X)$, i.e. $R(0.4, B)(x) = \min(0.4, B(x)) = 0.4 \wedge B(x)$, for all $x \in X$.*

Definition 6. *If $A, B : X \rightarrow I$, and R is any binary operator on I, i.e., $R : I \times I \rightarrow I$, then $R(A, B)$ is a fuzzy set on X, i.e., $R(A, B) : X \rightarrow I$, defined as $R(A, B)(x) = R(A(x), B(x))$, $\forall x \in X$.*

Remark 2. *Thus R can also be seen as $R : \widetilde{F}(X) \times \widetilde{F}(X) \rightarrow \widetilde{F}(X)$- where $\widetilde{F}(X)$ denotes the set of all fuzzy sets on X.*
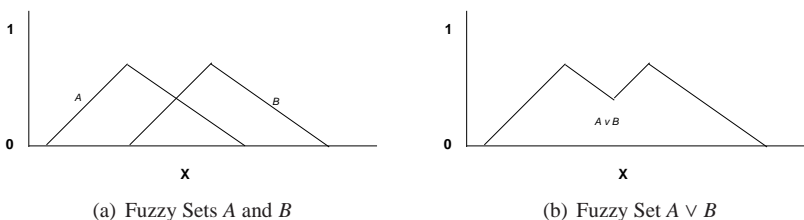

(a) Fuzzy Sets A and B (b) Fuzzy Set $A \vee B$

Figure 2: Fuzzy Sets A and B

For example if $R(a, b) = \max(a, b)$ then in Figure 2(a) we have $A, B \in \widetilde{F}(X)$ are fuzzy sets on X and Figure 2(b) gives $A^* = R(A, B) = \max(A, B) = A \vee B \in \widetilde{F}(X)$.

Definition 7 ([37] Definition 1.15, Pg 22). *A function $J : I^2 \rightarrow I$ is called a fuzzy implication if it has the following properties:*

$$J(p, r) \geq J(q, r) \; if \; q \geq p, \tag{J1}$$

$$J(p, r) \geq J(p, s) \; if \; r \geq s, \tag{J2}$$

$$J(0, r) = 1, \; \forall r \in I, \tag{J3}$$

$$J(p, 1) = 1, \; \forall p \in I, \tag{J4}$$

$$J(1, 0) = 0. \tag{J5}$$

The following are the two important classes of fuzzy implications well-established in the literature:

Definition 8 ([37] Definition 1.16, Pg 24). *An S-implication $J_{S,N}$ is obtained from a t-conorm S and a strong negation N as follows:*

$$J_{S,N}(a, b) = S(N(a), b), \forall a, b \in I. \tag{1}$$

Definition 9 ([37] Definition 1.16, Pg 24). *An R -implication J_T is obtained from a t-norm T as its residuation as follows:*

$$J_T(a, b) = Sup\{x \in I : T(a, x) \leq b\}, \forall a, b \in I. \tag{2}$$

R- and S-implications satisfy (J1) - (J5). Tables 1 and 2 list few of the well-known S-implications and R-implications, respectively.

Name	$S(a, b)$	$N(a, b)$	$J_{S,N}(a, b)$
Dienes	$max(a, b)$	$1 - a$	$max(1 - a, b)$
Reichenbach	$a + b - ab$	$1 - a$	$1 - a + ab$
Lukasiawicz	$min(1, a + b)$	$1 - a$	$min(1, 1 - a + b)$

Table 1: Some of the well known S-implications with their corresponding t-conorms

t-norm	$T(a, b)$	Implication	$J_T(a, b)$
Lukasiawicz	$max(0, a + b - 1)$	Lukasiawicz	$min(1, 1 - a + b)$
Mamdani	$min(a, b)$	Godel	$\begin{cases} 1, & \text{if } a \leq b \\ b, & \text{otherwise} \end{cases}$
Larsen	$min(1, a + b)$	Goguen	$\begin{cases} 1, & \text{if } a \leq b \\ b/a, & \text{otherwise} \end{cases}$

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

3.2   Fuzzy If-Then Rules

A linguistic statement "x *is* A" is interpreted as the variable x taking the linguistic value A. For example, if x denotes "Temperature" (on a suitable domain), then it can assume the following linguistic values A, viz., high, more or less high, medium, cool, very cold, etc. Each of the linguistic values (say cool) is represented by a fuzzy set on the domain X of the linguistic variable x, i.e., $A : X \rightarrow I$. The shape of the graph of the function represents the concept (say high temperature). The concept of high temperature is again context-dependent. For example, high temperature (fever) for a human being is different from the high temperature in a blast furnace, and accordingly the domain of the linguistic variable is selected.

A Fuzzy If-Then rule is of the form

$$\text{If } x \text{ is } A \text{ Then } y \text{ is } B, \tag{3}$$

where x, y are variables and A, B are linguistic expressions / values assumed by the linguistic variables. For example,

$$\text{"If } x \text{ (temperature) is } A \text{ (High)}$$
$$\text{Then } y \text{ (Pressure) is } B \text{ (Low)"}$$

The above is an example of a SISO rule. A Two-Input Single-Output rule is of the form

$$R_1 : \text{If } x \text{ is } A \text{ and } y \text{ is } B \text{ Then } z \text{ is } C,$$

where again A, B, C are linguistic values taken by the linguistic variables x, y, z over their respective domains.

3.3   Different Stages in the inferencing of a Fuzzy System

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

$$
\begin{aligned}
R_1 \quad &: \quad \textit{If } x_1 \textit{ is } A_1^1, \ \ldots, x_n \textit{ is } A_n^1 \textit{ Then } y \textit{ is } B_1 \\
&\vdots \\
R_j \quad &: \quad \textit{If } x_1 \textit{ is } A_1^j, \ \ldots, x_n \textit{ is } A_n^j \textit{ Then } y \textit{ is } B_j \\
&\vdots \\
R_m \quad &: \quad \textit{If } x_1 \textit{ is } A_1^m, \ \ldots, x_n \textit{ is } A_n^m \textit{ Then } y \textit{ is } B_m
\end{aligned}
\tag{4}
$$

where $A_i^j \in \widetilde{F}(X_i)$ for $i = 1, 2, \ldots, n$ are the antecedent fuzzy sets over the n non-empty domains $X_1, X_2, \ldots, X_n$. For $j = 1, 2, \ldots, m$, B_j can be a fuzzy set on the non-empty output domain Y, i.e., $B_j \in \widetilde{F}(Y)$, as in the case of a Mamdani Fuzzy System, or $B_j \in Y$ as in the case of a constant-output Takagi-Sugeno-Kang Fuzzy Systems.

In the following we propose a general framework for Inference in Fuzzy Rule Based Systems that captures the working of both the established models of Fuzzy Systems - TSK and Mamdani models of Inference. Towards this end, a Fuzzy System can be seen to consist of the following 5 stages:

3.3.1 Fuzzifier

If the given input is a crisp number $x \in X$, it is fuzzified to get a fuzzy set $\widetilde{X}$ on the corresponding input space X, i.e., $C : X \rightarrow \widetilde{F}(X)$, where $C(x) = \widetilde{X}$. Thus given a vector of crisp points $\overline{x} = [x_1, x_2, \ldots, x_n]$, where $x_i \in X_i$, for every input space X_i, we get a vector of Fuzzy sets $\overline{X} = [\widetilde{X}_1, \widetilde{X}_2, \ldots, \widetilde{X}_n]$. The often used [40,79] *Singleton Fuzzifier* of a crisp number x is given as

$$\widetilde{X}(y) = \begin{cases} 1 & \text{if } y = x \\ 0 & \text{otherwise} \end{cases} \qquad (SF^*)$$

Remark 3. *It can be readily seen that the above stage of "Fuzzifier" is the reverse of "Defuzzification" - wherein we obtain a crisp number from a fuzzy set (See 3.3.6 below). Though in many actual implementations of Fuzzy Systems a crisp value is directly given as input, the above stage has been added for generality. Also many times the input given to a fuzzy system is not precise owing to many types of observation errors. For example, a reading from a sensor that becomes an input for the controlling fuzzy system may be inherently imprecise due to instrument errors. In such cases a fuzzy set about the reading may be a more realistic input. In this paper, crisp inputs are identified with their fuzzified version as given by (SF^*).*

3.3.2 Matching

The input fuzzy sets $(\widetilde{X}_1, \widetilde{X}_2, \ldots, \widetilde{X}_n)$ are matched against their corresponding if-part fuzzy sets in each of the rule antecedents in the Fuzzy System, i.e.

$$M : \widetilde{F}(X_i) \times \widetilde{F}(X_i) \rightarrow I \qquad (5)$$

where $M(A_i^j, \widetilde{X}_i) = a_i^j$ for A_i^j and $\widetilde{X}_i \in \widetilde{F}(X_i)$, $j = 1, \ldots, m$.

A few matching functions used in the literature are given later in section 5.4.1.

3.3.3 Combining

In a multi-antecedent fuzzy system, the various matching degrees a_i^j of the n input fuzzy sets to the antecedent of the jth fuzzy if-then rule are combined to give the *"fit values"* μ_j ,

$$\mu : I^n \rightarrow I \qquad (6)$$

where $\mu(a_1^j, \ldots, a_n^j) = \mu_j$, $j = 1, 2, \ldots, m$. μ can be any t- or t-conorms (see Section 3.1).

3.3.4   Rule Firing

The combined value μ_j fires the rule consequent or the output fuzzy set B_j of the jth rule. This B_j can be a fuzzy set on Y,i.e.,$B_j \in \widetilde{F}(Y)$, or a value in Y, i.e.,$B_j \in Y$. Thus we have

$$f : I \times Z \to Z \qquad (7)$$

- When $Z = \widetilde{F}(Y)$ - the set of all fuzzy sets on the output domain Y, i.e.,$B_j \in \widetilde{F}(Y)$, $f(\mu_j, B_j) = f_j \in \widetilde{F}(Y)$ and is defined as $f(\mu_j, B_j(y)) = f_j(y), \forall y \in Y$

- When $Z = Y$, the output domain itself, i.e., $B_j \in Y$, then $f_j \in Y$ and is defined as $f(\mu_j, B_j) = f_j$.

Usually, $f = \pi$, the product, is commonly employed when $B_j \in Y \subseteq \mathbb{R}$, while a t-norm or any Fuzzy Implication Operator (Section 3.1) is the prefered choice if $B_j \in \widetilde{F}(Y)$.

3.3.5   Aggregation of Individual Inferences:

The fired output fuzzy sets (or crisp real numbers) f_j, $j = 1, 2, \ldots, m$ are then aggregated to obtain the final inferred fuzzy set (or crisp real number)

$$g : Z^m \to Z \qquad (8)$$

where again

- If $Z = \widetilde{F}(Y)$, the infered output set $g(f_1, \ldots, f_m) = B \in \widetilde{F}(Y)$. One can use any of the fuzzy logic operators, t- or t-conorms, to obtain $B \in \widetilde{F}(Y)$.

- If $Z = Y \subseteq \mathbb{R}$ then the Weighted Average or the Weighted Sum are the commonly used aggregation operators for g.

3.3.6   Defuzzification

When $Z = \widetilde{F}(Y)$, $g(f_1, \ldots, f_m) = B \in \widetilde{F}(Y)$ and we need to defuzzify B - a fuzzy set on Y - to a single value $b \in Y$, using an appropriate defuzzification method h as follows:

$$h : \widetilde{F}(Y) \to Y \qquad (9)$$

The Centre of Area or the Mean of Maxima methods [42, pp. 336 - 338] are the most widely used Defuzzification methods.

The different stages and the corresponding mappings capturing their actions are given in Table 3.

3.4   Different Models of Fuzzy System in the literature

Following are the two most established models of Fuzzy Systems:

Fuzzifier:	$\widetilde{X} = C(x)$	$C : X \to \widetilde{F}(X)$
Matching:	$a_i^j = M(A_i^j, \widetilde{X}_i)$	$M : \widetilde{F}(X_i) \times \widetilde{F}(X_i) \to I$
Combining:	$\mu_j = \mu(a_1^j, \ldots, a_n^j)$	$\mu : I^n \to I$
Firing:	$f_j = f(\mu_j, B_j)$	$f : I \times Z \to Z, Z = \widetilde{F}(Y) \ or \ Y$
Aggregation:	$B = g(f_1, \ldots, f_m)$	$g : Z^m \to Z$
Defuzzification:	$b = h(B)$	$h : \widetilde{F}(Y) \to Y$

Table 3: Different stages of a Fuzzy System

3.4.1 Mamdani Fuzzy System

E.H. Mamdani and S. Assilian [57] proposed the first type of Fuzzy Rule Based Systems. The rules in a Mamdani Fuzzy System are specified linguistically both for antecedents and consequents. Given a vector of crisp inputs $\overline{x'} = [x_1', x_2' \ldots, x_n']$, where $x_i' \in X_i$, the final output fuzzy set B on Y for the fuzzy rule base in (4) is obtained as follows:

$$B(y) \quad = \bigvee_{j=1}^m \{[\bigwedge_{i=1}^n a_i^j] \wedge B_j(y)\}, \forall \, y \in Y \tag{10}$$

where $a_i^j = A_i^j(x_i')$.

Though the Mamdani model is usually used with crisp inputs, it can handle both crisp and fuzzy inputs. In the case of a fuzzy inputs, say $x_1 \ is \ A_1, \ \ldots, x_n \ is \ A_n$, where A_i is a fuzzy set on the domain X_i, the final output fuzzy set B on Y for the fuzzy rule base in (4) is given by (10), but with a_i^j given by (11)

$$a_i^j = max_{x \in X_i}\{min(A_i^j(x), A_i(x))\} \tag{11}$$

Also in the case of a crisp input, the crisp input can be singleton fuzzified by *(SF*)* (Section 3.3.1) into a fuzzy set and can be given as an input to the fuzzy system. Thus given a vector of crisp inputs $\overline{x'} = [x_1', x_2', \ldots, x_n']$, where $x_i' \in X_i$, for every input space X_i, we get a vector of fuzzy inputs $\overline{X} = [\widetilde{X_1}, \widetilde{X_2}, \ldots, \widetilde{X_n}]$. It can be easily seen that if instead of the crisp inputs x_i', if their corresponding singleton fuzzified inputs are given, i.e., $A_i = \widetilde{X}_i$ are inputs, $a_i^j = A_i^j(x_i') = max_{x \in X_i}\{min(A_i^j(x), A_i(x))\} = max_{x \in X_i}\{min(A_i^j(x), \widetilde{X}_i(x))\}$. Thus we can always consider an input for the Mamdani model of fuzzy system to be fuzzy, with the understanding that any crisp input is singleton fuzzified according to *(SF*)* and (10) can be employed with (11).

Let the Matching Function $max_{x \in X}\{min(A(x), B(x))\}$ of two fuzzy sets $A, B : X \to I$ be denoted by $M_1(A, B)$. Now, comparing the inference in (10) to the different stages in Section 3.3, it can be seen that $M = M_1, \mu = \wedge, f = \wedge, g = \vee$ and $Z = \widetilde{F}(Y)$.

3.4.2   Takagi - Sugeno - Kang Fuzzy System

Instead of working with the linguistic rules of the kind employed in Mamdani Fuzzy
Systems, Takagi and Sugeno [71] proposed a new model based on rules whose an-
tecedent is composed of linguistic variables and the consequent is represented by a
real function of the input variables. TSK model differs from the Mamdani model
both in the form of their rules and the inference operators used. If in the case of
Mamdani model of a SISO fuzzy system a fuzzy rule has the form (3)

$$If \ x \ is \ A \ Then \ y \ is \ B$$

where A and B are fuzzy sets on X and Y, respectively, then in the case of the TSK
model the rules have the form (12)

$$If \ x \ is \ A \ Then \ y = b(x) \tag{12}$$

and the input is a crisp value for x. Their conclusion contains the real valued func-
tion $b(x)$ and not a fuzzy set. This function can be non-linear, although usually
linear functions are applied. Then the TSK rules have the form:

$$If \ x \ is \ A \ Then \ y = px + q \tag{13}$$

where the input is a crisp value for x and p, q are constants. In general the rules
of a SISO and MISO TSK fuzzy systems are of the form given by (14) and (15),
respectively.

$$R_j \ : \quad If \ x \ is \ A_j \ Then \ y \ = b_j(x) \tag{14}$$

$$R_j \ : \quad If \ x_1 \ is \ A_1^j, \ \ldots, \ x_n \ is \ A_n^j \ Then \ y \ = b_j(\overline{x}) \tag{15}$$

for $j = 1, \ldots, m$ and the input vector $\overline{x'} = [x_1', x_2', \ldots, x_n']$ and each x_i' is a crisp value
in X_i for $i = 1, \ldots, n$.

   Let us again consider a fuzzy rule base of m rules of the form (15) and a vector
of crisp inputs $\overline{x'} = [x_1', x_2', \ldots, x_n']$, where $x_i' \in X_i$, be given. In the TSK model of
fuzzy systems, the final crisp output is obtained as the Weighted Sum of *"fit values"*
and the rule consequents as given in (16).

$$F(\overline{x'}) = \sum_{j=1}^{m} \mu_j(\overline{x'}) \cdot b_j(\overline{x'}) \tag{16}$$

where $\mu_j(\overline{x'}) = \Pi_{i=1}^n a_i^j = \Pi_{i=1}^n A_i^j(x_i') = A_1^j(x_1') \cdot A_2^j(x_2') \cdot \ldots \cdot A_n^j(x_n')$.

   As in Section 3.4.1, by taking the singleton fuzzified crisp input vector $\overline{x'}$,
as given by *(SF*)*, it can be seen that, if $A_i = \widetilde{X}_i$ are inputs, $a_i^j = A_i^j(x_i') = $
$max_{x \in X_i}\{min(A_i^j(x), A_i(x))\} = max_{x \in X_i}\{min(A_i^j(x), \widetilde{X}_i(x))\}$. Thus again one can always
consider the singleton fuzzified fuzzy set $\widetilde{X}_i$ of a crisp input x_i' as being the input for
a TSK model of fuzzy system. Also product is the antecedent combiner, i.e., $\mu = \Pi$.
Though the product between the "fit value" of the given input to the antecedents of

rule j, $\mu_j(\overline{x'})$ and its consequent $b_j(\overline{x'})$ is an effect of the Weighted Sum aggregation employed and is not a rule connective, per se, one can perhaps consider it such so that $f = \pi$ for the TSK model in the above framework, i.e., $f : I \times Z \to Z$ is such that $f(\mu_j(\overline{x'}), b_j(\overline{x'})) = \mu_j(\overline{x'}) \cdot b_j(\overline{x'})$, where $Z = Y \subseteq \Re$, the actual domain of the output fuzzy sets.

Now, comparing the inference in (16) to the different stages in Section 3.3, it can be seen that $M = M_1, \mu = \pi, f = \pi, g = \Sigma$ and $Z = Y \subseteq \Re$.

From the above two sections, it is clear that the different stages in the inference of an output, given an input, in a fuzzy system can be mapped to different functions capturing the actions performed at every stage.

Definition 10. *A model of Inference in a fuzzy system is given by the quintuple $Q = \{M, \mu, f, g, Z\}$ where M, μ, f, g are the corresponding operators of the above framework and Z is the domain of consequents of the rule.*

Thus Mamdani Model of inference in a fuzzy system is defined as the quintuple $Q_M = \{M_1, \wedge, \wedge, \vee, \widetilde{F}(Y)\}$ while the TSK model of inference in a fuzzy system is given by $Q_{TSK} = \{M_1, \prod, \prod, \Sigma, Y \subseteq \Re\}$. We do not consider the fuzzifier stage since a crisp input to the fuzzy system can be thought of as a singleton fuzzified input fuzzy set using (SF^*). Table 4 summarises the above discussion, where $\prod =$ Product, $\Sigma =$ Sum, $\vee =$ max, $\wedge =$ min.

Name / Type	M	μ	f	g	Fuzzifier	Z
TSK	M_1	$\prod$	$\prod$	Σ	SF^*	$Y \subseteq \Re$
Mamdani	M_1	$\wedge$	$\wedge$	$\vee$	SF^*	$\widetilde{F}(Y)$

Table 4: M, μ, f, g and Z for the different models of fuzzy systems in Section 3.4

3.5 A Rule Reduction Technique for a Class of Fuzzy Systems

More often than not, the number of fuzzy sets, k, defined on the single output domain Y, is typically much less than the number of rules m, i.e., $k \ll m$. This suggests that the antecedents of more than one rule lead to the same consequent. To eliminate this redundancy, we propose a new type of off-line rule reduction where the rules with the same consequent 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 inference obtained from the original rule base and that obtained from the reduced rule base should be identical. This necessitates the functions M, μ, f and g to possess some properties. These are explored in the next section.

Remark 4. *Also in the rest of the paper, we will only consider fuzzy rules (SISO or MISO, as the case may be) of the type where both the antecedents and consequents*

*are fuzzy sets on their respective domains. The inputs may be crisp or fuzzy. In the
case of crisp inputs, we will consider their fuzzified version as obtained using (SF^*)
on the input.*

4 Conditions on the General Framework for Lossless Rule Reduction

The rule reduction procedure we propose is an offline procedure, i.e., from the given
original rule base we club the rules with same consequent, but different antecedents,
to produce new rules to replace old rules. In this, section we determine the structure
of the antecedents of the newly formed rules. The following theorem gives sufficient
conditions that the operators of the above proposed framework should satisfy to
obtain *lossless* or *inference invariant* rule reduction by combining antecedents of
rules that have identical consequents.

4.1 The Restrictions on g, f, μ and M :

Theorem 1. *Let a model of inference in a fuzzy system be defined by $Q = \{M, \mu, f, g, Z = \widetilde{F}(Y)\}$, where Y is the output domain. The following conditions
on the operators M, μ, f and g are sufficient to ensure inference invariant rule re-
duction, by combining antecedents of rules that have identical consequents, in any
MISO fuzzy system:*

*There exist operators g, o_g, o_μ which are commutative and associative binary
operators on I and for any $a, b, a_1, a_2, b_1, b_2 \in I$, A_1, A_2, A which are fuzzy sets
defined on an input domain X and $C \in \widetilde{F}(Y)$,*

$$g[f(a, C), f(b, C)] = f(a \ o_g \ b, C) \tag{17}$$

$$\mu(a_1, b_1) \ o_g \ \mu(a_2, b_2) = \mu(a_1 \ o_\mu \ a_2, b_1 \ o_\mu \ b_2) \tag{18}$$

$$M(A_1, A) \ o_\mu \ M(A_2, A) = M(A_1 \ o_\mu \ A_2, A) \tag{19}$$

Remark 5. *In the LHS of (19) o_μ is a binary operator on I while in the RHS of (19)
o_μ is the extension of o_μ to fuzzy sets on X (See Definition 6 and Remark 2 in Section
3.1).*

Proof. Without loss of generality, let us take a 2-input 1-output fuzzy system con-
sisting of three rules, where X_1 and X_2 are the input domains and Y the output
domain. Consider the fuzzy system given by the following rules, written in a sim-
plified form:

$$R_1 : A_1, B_1 \rightarrow C$$
$$R_2 : A_2, B_2 \rightarrow C \tag{20}$$
$$R_3 : A_3, B_3 \rightarrow D$$

where A_1, A_2, A_3 are fuzzy sets on X_1; B_1, B_2, B_3 are fuzzy sets on X_2 and C, D are fuzzy sets on Y.

Let us consider the inference in the above MISO - fuzzy system in the presence of an input, say "x_1 *is A and* x_2 *is B*", which is represented as $X = (A, B)$, where $A \in \widetilde{F}(X_1)$ and $B \in \widetilde{F}(X_2)$. The MISO inference from the original rule base (20) is given as:

$$g\{ \quad f[\mu(M(A_1, A), M(B_1, B)), C],$$
$$f[\mu(M(A_2, A), M(B_2, B)), C],$$
$$f[\mu(M(A_3, A), M(B_3, B)), D]\} \tag{21}$$

Letting $\mu(M(A_1, A), M(B_1, B))$ as a and $\mu(M(A_2, A), M(B_2, B))$ as b we have from (17), with g being associative,

$$(21) \quad = \quad g\{g\{f[\mu(M(A_1, A), M(B_1, B)), C],$$
$$f[\mu(M(A_2, A), M(B_2, B)), C]\},$$
$$f[\mu(M(A_3, A), M(B_3, B)), D]\}$$
$$= \quad g\{f[\mu(M(A_1, A), M(B_1, B))$$
$$o_g \, \mu(M(A_2, A), M(B_2, B)), C],$$
$$f[\mu(M(A_3, A), M(B_3, B)), D]\} \tag{22}$$

Again letting $M(A_i, A) = a_i \in I, M(B_i, B) = b_i \in I, \ i = 1, 2$, we have using (18) and (19)

$$(22) \quad = \quad g\{f[\mu(M(A_1, A) \, o_\mu \, M(A_2, A),$$
$$M(B_1, B) \, o_\mu \, M(B_2, B)), C],$$
$$f[\mu(M(A_3, A), M(B_3, B)), D]\} \tag{23}$$
$$= \quad g\{f[\mu(M(A_1 \, o_\mu \, A_2, A), M(B_1 \, o_\mu \, B_2, B)), C],$$
$$f[\mu(M(A_3, A), M(B_3, B)), D]\} \tag{24}$$

Thus the rule base in (20) can be reduced to the following rule base containing just two rules:

$$R_1^* \quad : A_1 \, o_\mu \, A_2, B_1 \, o_\mu \, B_2 \rightarrow C \tag{25}$$
$$R_3 \quad : A_3, B_3 \rightarrow D$$

It can be easily seen that for a given input $X = (A, B)$, the inference obtained from the reduced rule base (25) under the given model of inference Q is identical to (24). □

The above requirements on the general framework give us a class of Fuzzy Systems that allow lossless rule reduction by combining rules with same consequent.

5   Analysis of the Requirements for Lossless Rule Reduction

In this section, we explore each of the above conditions for Lossless Rule Reduction in the setting of Fuzzy Logic operators.

5.1   Conditions on g, o_g and o_μ

In this study we consider only continuous t-norms and t-conorms for g, o_g and o_μ, which are by definition both commutative and associative. This enables us to extend them to functions from I^n to I in the case of n-input domains.

5.2   On the Equation (17)

Typically in a Fuzzy System f (the rule firing operation) is interpreted either as a t-norm, for example Mamdani's *min*, or a Fuzzy Implication operator. In this work we investigate the solution of (17) both with f as a Fuzzy Implication Operator and as a t-norm. We explore equation (17) in the following way:

- Fix f to be in a specific class of Fuzzy Implications or any t-norm and vary g and hence o_g over conitnuous t-norms and t-conorms.

In this study, we consider the two most established and well studied families of Fuzzy Implication Operators, viz., R- and S-Implications. (See Definitions 8 and 9 in Section 3.1).

When f is fixed in (17), g and o_g are taken to be any S/T-norms, we have the following 4 possibilities:

$$f(T(p,q),r) = S(f(p,r),f(q,r)) \tag{26}$$
$$f(S(p,q),r) = T(f(p,r),f(q,r)) \tag{27}$$
$$f(T_1(p,q),r) = T_2(f(p,r),f(q,r)) \tag{28}$$
$$f(S_1(p,q),r) = S_2(f(p,r),f(q,r)) \tag{29}$$

5.2.1 $f = J$ a Fuzzy Implication

Fixing f to be any Fuzzy Implication J, we get the following four equations from the above:

$$J(T(p,q),r) = S(J(p,r),J(q,r)) \tag{30}$$
$$J(S(p,q),r) = T(J(p,r),J(q,r)) \tag{31}$$
$$J(T_1(p,q),r) = T_2(J(p,r),J(q,r)) \tag{32}$$
$$J(S_1(p,q),r) = S_2(J(p,r),J(q,r)) \tag{33}$$

Recently with $f = J$ interpreted as an R- or an S-implication and $g = S$, an t-conorm and $o_g = T$, a t-norm, Trillas and Alsina [76] have investigated (30) and proven the following:

Theorem 2. *An S- or an R-implication J satisfies (30) iff S = max and T = min.*

In [7] the authors have proven the following Theorem 3 concerning equation (31) obtained by letting $f = J$ to be an R- or an S-implication and $o_g = S$, an t-conorm and $g = T$, a t-norm,

Theorem 3. *An S- or an R-implication J satisfies (31) iff S = max and T = min.*

Also we have the following result:

Lemma 1. *For no Fuzzy Implication J, t-norm T (t-conorm S, respectively) do equations (32) ((33), respectively) hold.*

Proof. Let $p = 1, q = r = 0$. Then using the property of t-norms, $T(1, 0) = 0$ and (J3), we have that,

$$LHS\ of\ (32) \quad = \quad J(T_1(1, 0), 0) = J(0, 0) = 1$$
$$RHS\ of\ (32) \quad = \quad T_2(J(1, 0), J(0, 0)) = T_2(0, 1) = 0$$

LHS = RHS implies that $1 = 0$, which is absurd. Similarly, that (33) does not have a solution can be seen by again fixing $p = 1, q = r = 0$.   □

5.2.2 $f = T$ **a t-norm**

In [76] it is also shown that (30) does not hold for the Mamdani's Minimum $f = J = \wedge$ and the Larsen's Product $f = J = \prod$ operators. That (26) and (27) do not hold when f is any t-norm T can be easily seen by taking $p = r = 1$ and $q = 0$. Thus fixing $f = T$ to be a t-norm, we need to consider only the equations (28) and (29) which become:

$$T(T_1(p, q), r) = T_2(T(p, r), T(q, r)) \tag{34}$$
$$T(S_1(p, q), r) = S_2(T(p, r), T(q, r)) \tag{35}$$

We have the following theorems:

Theorem 4. *(34) is valid iff when $T_1 \equiv T_2 = min$.*

Proof. Claim: $T_1 \equiv T_2$ on $I \times I$.
    Let $r = 1$. Then $\forall p, q \in I$, we have
LHS of (34) = $T(T_1(p, q), 1) = T_1(p, q)$
RHS of (34) = $T_2(T(p, 1), T(q, 1)) = T_2(p, q)$ = LHS of (34) $\forall p, q \in I$ iff $T_1 \equiv T_2$.
    Now, let $p = q = 1, r \in I$. Then
LHS of (34) = $T(T_1(1, 1), r) = T(1, r) = r$.
RHS of (34) = $T_1(T(1, r), T(1, r)) = T_1(r, r) = r,\ \forall r \in I$ iff $T_1 = min$, the only idempotent t-norm.   □

Theorem 5. *(35) is valid iff when $S_1 \equiv S_2 = S$ and T distributes over S.*

Proof. Claim: $S_1 \equiv S_2$ on $I \times I$.

Let $r = 1$. Then $\forall p, q \in I$, we have

LHS of (35) $= T(S_1(p, q), 1) = S_1(p, q)$

RHS of (35) $= S_2(T(p, 1), T(q, 1)) = S_2(p, q) = $ LHS of (35) $\forall p, q \in I$ iff $S_1 \equiv S_2$.

Thus the equation (35) becomes

$$T(S(p, q), r) = S(T(p, r), T(q, r)) \tag{36}$$

which is true iff T distribuites over S.   □

Corollary 1. *(36) is true if S = max.*

5.3   On the Equation (18)

Continuing along the same vein, we have investigated the generalised bisymmetry
equation (18)

$$\mu(a_1, b_1) \ o_g \ \mu(a_2, b_2) = \mu(a_1 \ o_\mu \ a_2, b_1 \ o_\mu \ b_2)$$

involving o_g, μ and o_μ, with $a_1, a_2, b_1, b_2 \in I$.

Definition 11. *A function $B : [a, b]^2 \to [a, b]$ is said to be bisymmetric if*

$$B(B(x, y), B(u, v)) = B(B(x, u), B(y, v)), \ \ \forall x, y, u, v \in [a, b].$$

For a comprehensive coverage on Bisymmetry Equations refer [1 - 4, 74]. Also
[38,39] list many results on bisymmetry equations on the unit interval. Allowing
o_g, μ and o_μ to be t- and t-conorms, we get the following 8 possible cases in all,
which for convenience we have grouped into two sets:

Group 1

$$T_1(T_2(a_1, b_1), T_2(a_2, b_2)) = T_2(T_3(a_1, a_2), T_3(b_1, b_2)) \tag{37}$$

$$S_1(S_2(a_1, b_1), S_2(a_2, b_2)) = S_2(S_3(a_1, a_2), S_3(b_1, b_2)) \tag{38}$$

Group 2

$$T_1(S(a_1, b_1), S(a_2, b_2)) = S(T_3(a_1, a_2), T_3(b_1, b_2)) \tag{39}$$

$$T_1(T_2(a_1, b_1), T_2(a_2, b_2)) = T_2(S(a_1, a_2), S(b_1, b_2)) \tag{40}$$

$$T_1(S_1(a_1, b_1), S_1(a_2, b_2)) = S_1(S_2(a_1, a_2), S_2(b_1, b_2)) \tag{41}$$

$$S_1(T(a_1, b_1), T(a_2, b_2)) = T(S_2(a_1, a_2), S_2(b_1, b_2)) \tag{42}$$

$$S_1(S_2(a_1, b_1), S_2(a_2, b_2)) = S_2(T(a_1, a_2), T(b_1, b_2)) \tag{43}$$

$$S_1(T_1(a_1, b_1), T_1(a_2, b_2)) = T_1(T_2(a_1, a_2), T_2(b_1, b_2)) \tag{44}$$

We show that only 2 of the above 8 equations, the ones belonging to *Group 1* have solutions, while the rest of the equations belonging to *Group 2* do not have solutions as given by the following theorems, the proofs of which can be found in the Appendix.

Theorem 6. *If T_1, T_2 and T_3 are any t-norms then the equation (37) obtained by letting $o_g = T_1, \mu = T_2$ and $o_\mu = T_3$ in (18) is valid iff $T_1 \equiv T_2 \equiv T_3$ on I^2 .*

Theorem 7. *If S_1, S_2 and S_3 are any t-conorms then the equation (38) obtained by letting $o_g = S_1, \mu = S_2$ and $o_\mu = S_3$ in (18) is valid iff $S_1 \equiv S_2 \equiv S_3$ on I^2 .*

Theorem 8. *The equations belonging to* Group 2 *do not have solutions.*

5.4　　On the Equation (19)

In this section, we investigate equation (19), namely,

$$M(A_1, A) \ o_\mu \ M(A_2, A) = M(A_1 \ o_\mu \ A_2, A)$$

where M is a matching function that compares two fuzzy sets on the same domain, i.e., $M : \widetilde{F}(X) \times \widetilde{F}(X) \rightarrow I$, with o_μ a t- or t-conorm, in which case we get the following equations (45) and (46):

$$T[M(A_1, A), M(A_2, A)] = M(T(A_1, A_2), A) \tag{45}$$

$$S[M(A_1, A), M(A_2, A)] = M(S(A_1, A_2), A) \tag{46}$$

5.4.1　　A few Matching functions existing in the literature

Below we list a few of the matching functions commonly employed in the literature.

- Zadeh's Sup-min : $M_1(A, A') = max_x \ min(A(x), A'(x))$

- Magrez - Smets' Measure [56]: $M_2(A, A') = max_x \ min(\overline{A(x)}, A'(x))$, where $\overline{A(x)}$ is the negation of $A(x)$.

- Sup-T : $M_3(A, A') = max_x \ T(A(x), A'(x))$, where T is any t-norm.

- Sup-T-N: $M_4(A, A') = max_x \ T(\overline{A(x)}, A'(x))$.

- Inf- max : $M_5(A, A') = min_x \ max(A(x), A'(x))$.

- Inf - max- N: $M_6(A, A') = min_x \ max(\overline{A(x)}, A'(x))$.

- Inf-S : $M_7(A, A') = min_x \ S(A(x), A'(x))$, where S is any t-conorm.

- Inf - S - N: $M_8(A, A') = min_x \ S(\overline{A(x)}, A'(x))$.

Note: M_3 and M_4 (M_7 and M_8) are generalisations of M_1 and M_2 (M_5 and M_6), respectively, while M_5, M_6, M_7 and M_8 are duals of M_1, M_2, M_3 and M_4.

The proofs of the following results can be found in the Appendix.

Theorem 9. M_1, M_2, M_3 and M_4 satisfy equation (46) iff $S = max$.

Theorem 10. M_5, M_6, M_7 and M_8 satisfy equation (45) iff $T = min$.

Remark 6. M_1, M_2, M_3 and M_4 (M_5, M_6, M_7 and M_8) do not satisfy (45) (resp.
(46)) since there does not exist any t-conorm S (resp. t-norm T) such that
$S(min_x a_x, min_x b_x)$ $=$ $min_x S(a_x, b_x)$ (such that $T(max_x a_x, max_x b_x)$ $=$
$max_x T(a_x, b_x)$). We refer the readers to [20, 21] for the corresponding proofs.

Combining the results of section 5.1 - 5.4, we get the following table - Table 5 -
of operators available for equations (17),(18) and (19), where J_S, J_R denote S- and
R-implication, while T and S denote a t-norm and t-conorm, respectively.

f	g	o_g	μ	o_μ	Conditions	Examples of M
J_S or J_R	$\vee$	$\wedge$	$\wedge$	$\wedge$	-	M_5, M_6, M_7, M_8
J_S or J_R	$\wedge$	$\vee$	$\vee$	$\vee$	-	M_1, M_2, M_3, M_4
T	$\wedge$	$\wedge$	$\wedge$	$\wedge$	-	M_5, M_6, M_7, M_8
T	S	S	S	S	T dist over S	$M_1, M_2, M_3, M_4; S = \vee$

Table 5: Table of operators for (17),(18) and (19) to be satisfied

6   Examples of a few Fuzzy Systems from the above class

In this section we show how the results from Section 5 can be applied to particular
models of inferencing in Fuzzy Systems. For throughout this section, we consider
the following SISO fuzzy system with 3 rules as given in (47).

$$R_1 : A_1 \to B$$
$$R_2 : A_2 \to B \qquad\qquad (47)$$
$$R_3 : A_3 \to C$$

where A_1, A_2, A_3 are fuzzy sets on X; B, C are fuzzy sets on Y and $\to$ is any rule
firing operation relating the antecedent to the consequent.

6.1   Mamdani Model of Inference in Fuzzy Systems

Consider the following set of m Single-Input Single-Output (SISO) fuzzy if-then
rules of Mamdani type:

$$If\ x\ is\ A_j\ Then\ y\ is\ B_j,\ \ j = 1, 2, \cdots, m$$

where A_i, B_i are fuzzy sets on the input and output domains X, Y, respectively. From Section 3.4.1, we know a Single-Input Single-Output (SISO) Mamdani type Fuzzy System has the final output fuzzy set B given by

$$B(y) \quad = \bigvee_{j=1}^{m}\{[A_j(x) \wedge B_j(y)]\}, \forall \ y \in Y \tag{48}$$

which corresponds to $Q_M = \{M_1, na, \wedge, \vee, \widetilde{F}(Y)\}$.

Remark 7. *Since in the case of SISO rule base, the antecedent combiner μ does not play a role we have indicated it as Not Applicable - na - in Q_M.*

6.1.1 Lossless Rule Reduction in Mamdani Model of Inference in SISO Fuzzy Systems

Theorem 11. *Inference Invariant Rule Reduction is possible in Mamdani Model of Inference, in the case of SISO fuzzy rules, by combining the antecedents of rules that have identical consequent.*

Proof. We know that f - the rule firing operator - is the t-norm *min* in (48). In the presence of an input, say *x is A*, denoted as $X = A$, we have from (48), the final output fuzzy set B' is given by

$$\begin{aligned} B'(y) \quad &= \quad [M_1(A_1, A) \wedge B] \\ &\quad \vee [M_1(A_2, A) \wedge B] \\ &\quad \vee [M_1(A_3, A) \wedge C] \end{aligned} \tag{49}$$

¿From (49) by the distributivity of $\wedge$ over $\vee$ we have (50),

$$\begin{aligned} B'(y) \quad &= \quad \{[M_1(A_1, A) \ \vee (M_1(A_2, A)] \wedge B\} \\ &\qquad \vee \ [M_1(A_3, A) \wedge C] \tag{50} \\ &= \quad [M_1(A_1 \vee A_2, A) \wedge B] \\ &\qquad \vee \ [(M_1(A_3, A) \wedge C] \tag{51} \\ &= \quad [M_1(A_1^*, A) \wedge B] \vee [(M_1(A_3, A) \wedge C] \tag{52} \end{aligned}$$

We know from Theorem 9 that $M_1(A_1, A) \vee M_1(A_2, A) = M_1(A_1 \vee A_2, A)$, using which we obtain (51) from (50). In (52) $A_1^* = A_1 \vee A_2$, which is again a fuzzy set on X, by Definition (6) and Remark 2.

Thus instead of the SISO fuzzy rule base of 3 rules (47), the following reduced rule base with two rules can be used, without any loss of inference for a given input, while employing the Mamdani Model of Inference.

$$\begin{aligned} R_1 \quad &: \quad A_1^* \rightarrow B \\ R_3 \quad &: \quad A_3 \rightarrow C \end{aligned} \tag{53}$$

$\square$

6.2 General Mamdani Model of Inference in Fuzzy Systems

A slight generalisation of the Mamdani model of inference can be seen as follows:
Let $Q_M^T = \{M, na, T^*, S^*, \widetilde{F}(Y)\}$ denote a General Mamdani Model of Inference
where T^* is any t-norm that distributes over the t-conorm S^*. Then the following
can be easily shown as above:

Theorem 12. *Inference Invariant Rule Reduction is possible in Mamdani Model of
Inference, in the case of SISO fuzzy rules, by combining the antecedents of rules that
have identical consequent, if the Matching function M is such that*

$$S^*[M(A_1, A), M(A_2, A)] = M(S^*[A_1, A_2], A). \tag{54}$$

In the case $S^* = max$, the matching function M, among others, can be one of
M_1, M_2, M_3, M_4.

6.3 Modified Mamdani Model of Inference in Fuzzy Systems

By a Modified Mamdani Model of Inference we refer to the following quintuple
$Q_M^J = \{M, na, J, \wedge, \widetilde{F}(Y)\}$, where M is any Matching function and J is either an
R- or an S-implication. In this model of inference an Implication Operator J is
employed to relate the antecedent and the consequent of the fuzzy rules. The final
output fuzzy set B in Q_M^J for a SISO rule base is given by

$$B(y) \quad = \bigwedge_{j=1}^{m} \{[A_j(x) \rightarrow B_j(y)]\}, \forall \, y \in Y \tag{55}$$

where $\rightarrow$ is either an R- or an S-implication.

Recently Li et al [54,55] have shown that the above Modified Mamdani Model
of Inference in Fuzzy Systems with R- or S-implications for the rule firing operation
and with Trapezoidal or Triangular membership functions are Universal Approxi-
mators both in the case of SISO and MISO fuzzy systems. (In the case of MISO
systems the antecedent combiner $\mu = \wedge$). These considerations make Modified
Mamdani Model of Inference very attractive. In this section we show that lossless
rule reduction is possible in Modified Mamdani Model of Inference in the SISO
case.

6.3.1 Lossless Rule Reduction in Modified Mamdani Model of Inference in SISO Fuzzy Systems

Theorem 13. *Inference Invariant Rule Reduction is possible in Modified Mamdani
Model of Inference, in the case of SISO fuzzy rules, by combining the antecedents of
rules that have identical consequent, if the Matching function M obeys (56).*

$$M(A_1, A) \vee M(A_2, A) = M(A_1 \vee A_2, A). \tag{56}$$

Proof. Let us now interpret f - the rule firing operator - as an R- or an S-implication
in (55). Let us again consider the above SISO fuzzy system of 3 rules as given in

(47), but where the $\rightarrow$ is any R- or S-implication. Also let the Matching function M obey (56).

In the presence of an input, say x is A, denoted as $X = A$, we have from (55), the final output fuzzy set B' is given by

$$B'(y) = [M(A_1, A) \rightarrow B] \wedge [M(A_2, A) \rightarrow B]$$
$$\wedge [M(A_3, A) \rightarrow C] \tag{57}$$

¿From (57) by using Theorem 3 we obtain (58),

$$\begin{aligned} B'(y) &= \{[M(A_1, A) \vee (M(A_2, A)] \rightarrow B\} \\ &\quad \wedge [M(A_3, A) \rightarrow C] \end{aligned} \tag{58}$$
$$\begin{aligned} &= [M(A_1 \vee A_2, A) \rightarrow B] \\ &\quad \wedge [M(A_3, A) \rightarrow C] \end{aligned} \tag{59}$$
$$= [M(A_1^*, A) \rightarrow B] \wedge [M(A_3, A) \rightarrow C] \tag{60}$$

We obtain (59) by using the fact that M obeys (56). In (60) $A_1^* = A_1 \vee A_2$, which is again a fuzzy set on X, by Definition (6) and Remark 2. Thus again instead of the SISO fuzzy rule base of 3 rules (47), we have the reduced rule base (53).   □

Examples of Matching functions M that satisfy (56) are M_1, M_2, M_3, M_4.

In general, the number of rules can be reduced to k , where k is the number of output fuzzy sets that featured in the original rule base. Most importantly, this type of rule reduction is lossless w.r.to inference and Table 6 summarises the above discussion for the SISO case with the following:

Condition (i) T^* distributes over S^*

Condition (ii) M satisfies (54).

Condition (iii) M satisfies (56).

Name / Type	Q	M	f	g	$o_g = o_\mu$	Conditions
Original Mamdani	Q_M	M_1	$\wedge$	$\vee$	$\vee$	-
General Mamdani	Q_M^T	M	T^*	S^*	S^*	(i) and (ii)
Modified Mamdani	Q_M^J	M	J	$\wedge$	$\vee$	(iii)

Table 6: M, f, g, o_g and o_μ for the different models of inference discussed in Sections 6.1 - 6.3

7   Conclusion

In this work we have proposed a simple rule reduction technique that combines rules with identical consequents, which is lossless with respect to inference. Towards this

end, we proposed a general framework for Inference in Fuzzy Systems and imposed certain requirements on the different inference operators employed in a Fuzzy System. Also we have explored these requirements in the setting of Fuzzy Logic Operators. We have also given a few examples of Models of Inference in Fuzzy Systems that have the required properties for inference invariant rule reduction. We note a few observations below:

- Merging of rules, in some cases, may turn out to be computationally intensive, but this is a one time off-line job which even though might reduce interpretability will make computation of inferences much faster. Perhaps the original rule base can still be preserved for interpretability considerations.

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

In this work, we have considered only S- and R-implications for the fuzzy implication J and have shown the important role played by their distributivity over t-norms and t-conorms in the inference scheme in Section 6.3. Recently, there are a few more families that have been proposed, viz., U-implications and the residual implications of uninorms J_{U*} in [29] and the recently proposed families of f-generated implications J_f and g-generated implications J_g by Yager in [85] and h-generated implications J_h in [8], [9]. The distributivity of J_{U*} and U-implications over uninorms - which are generalisations of t-norms and t-conorms (see [84]) - is studied in [27] and [28] while that of J_f over t-norms and t-conorms is done in [9]. Hence these families of fuzzy implications can also be employed for the inference scheme in Section 6.3.

In this work we have considered in detail the proposed rule reduction technique only in the SISO case explicitly. Recently we have done some work on rule reduction in the MISO case also, as has been demonstrated within the scope of Similarity Based Reasoning in [10].

Acknowledgments

References

[1] J. Aczel, Lectures on Functional Equations and their Applications, Academic Press (1966).

[2] 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. 203 (1996) 104 - 126.

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

[4] 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. 214 (1997) 22 - 35.

[5] Michal Baczynski, On a class of Distributive Fuzzy Implications , Intl. Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 9 (2) (2001) 229 - 238.

[6] Michal Baczynski, Contrapositive Symmetry of Distributive Fuzzy Implications , Intl. Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 10 (Suppl) (2002) 135 -147.

[7] Balasubramaniam.J., and C.Jagan Mohan Rao, On the Distributivity of Implication Operators over T- and t-conorms , IEEE Transaction on Fuzzy Systems, 12 (2) (2004) 194 - 198.

[8] Balasubramaniam.J., "Contrapositive symmetrization of fuzzy implications - Revisited", Fuzzy Sets and Systems 157 (2006) 2291 – 2310.

[9] Balasubramaniam.J., "Yagers new class of implications J_f and some classical tautologies"', Info. Sci. *In Press*.

[10] Balasubramaniam.J., Rule Reduction for Efficient Inferencing in Similarity Based Reasoning, Int. Jl. Approx. Reas., *Submitted*.

[11] P. Baranyi and Y. Yam, Singular value-based approximation with Takagi-Sugeno type fuzzy rule base, Proc. of the 6th IEEE Int. Conf. on Fuzzy Systems (FUZZ-IEEE'97), volume I, Barcelona, Spain, 1997, pp. 265 - 270.

[12] P. Baranyi and Y. Yam, Singular value-based approximation with non-singleton fuzzy rule base, Proc. of the 7th Int. Fuzzy Systems Association World Congress (IFSA'97), volume II, Prague, Czech Republic, 1997, pp. 127 - 132.

[13] P.Baranyi and Y.Yam, Fuzzy Rule Base Reduction in Fuzzy IF-THEN Rules in Computational Intelligence : Theory and Applications, Eds. D.Ruan and E.E.Kerre, Kluwer, (2000) 135 - 160.

[14] P.Baranyi et al, Trade-off between Approximation Accuracy and Complexity : HOSVD based Complexity Reduction, Periodica Polytechnica Ser. Transp. Eng. 29 (1-2) (2001) 3 - 26.

[15] P.Baranyi, P.Korondi and H.Hashimoto, Trade-off Between Approximation accuracy and complexity for TS fuzzy models , Asian Journal of Control 6 (21-1) (2004) 21 - 33.

[16] P.Baranyi, Y.Yam, D.Tikk and R.J.Patton, Trade-off between approximation accuracy and complexity: TS controller design via HOSVD based complexity minimization, Studies in Fuzziness and Soft Computing, Vol. 128. Interpretability Issues in Fuzzy Modeling, J. Casillas, O. Cordn, F.Herrera, L.Magdalena (Eds.), Springer-Verlag, 2003. pp. 249-277.

[17] P.Baranyi, Y.Yam, A.R.Várkonyi-Kóczy and R.J.Patton, SVD Based Reduction to MISO TS Fuzzy Models , IEEE Transaction on Industrial Electronics, Vol. 50. No. 1. February 2003, pp. 232-242.

[18] P.Baranyi, Y.Yam, A. R.Várkonyi-Kćzy, R.J.Patton, P.Michelberger and M.Sugiyama, SVD Based Reduction to TS Fuzzy Models , IEEE Transaction on Industrial Electronics, 49 (2) (2002) 433-443.

[19] Boverie,N., Narishkin,D., Lequelle,J., Titli,A., Fuzzy Control of Higher Order Systems using a Parallel Structure of second order blocks , in Preprints from IFAC World Congress, (1993) 573 - 576.

[20] Carlo Bertoluzza, On the Distributivity between t-Norms and t-Conorms, FUZZ-IEEE '93, 140-147.

[21] Carlo Bertoluzza, V. Doldi, "On the Distributivity between t-norms and t-conorms", Fuzzy Sets and Systems 142 (1) (2004) 85-104.

[22] C.Chantana, S.Tongsima and E.H.M.Sha, Minimization of Fuzzy Systems based on Fuzzy Inference Graphs , Research Report TR-96-6, University of Notre Dame.

[23] W.E.Combs and J.E.Andrews, Combinatorial rule explosion eliminated by a fuzzy rule configuration , IEEE Trans. Fuzzy Systems 6 (1998) 1-11.

[24] W.E.Combs, Author's reply , IEEE Trans. Fuzzy Systems 7 (1999) 371.

[25] W.E.Combs, Author's reply , IEEE Trans. Fuzzy Systems 7 (1999) 478 - 479.

[26] V.Cross, T.Sudkamp, "Fuzzy Implication and Compatibility Modification", FUZZ-IEEE '93, San Francisco, California March 28 - April 1, 1993 pp. 219 - 224.

[27] Daniel Ruiz, Joan Torrens, Distributive Strong Implications from Uninorms. Proc. AGOP-2005 (2005), 103-108.

[28] Daniel Ruiz, Joan Torrens, Distributive Residual Implications from Uninorms. Proc. EUSFLAT-2005 (2005).

[29] De Baets B., Fodor J., Residual operators of uninorms, Soft Computing 3 (1999), 89–100.

[30] S.Dick and A.Kandel, Comments on "Combinatorial rule explosion eliminated by a fuzzy rule configuration" , IEEE Trans. Fuzzy Systems 7 (1999) 475 - 477.

[31] Dubois.D. and H.Prade, " Fuzzy sets in approximate reasoningPart I: Inference with possibility distributions", Fuzzy Sets and Systems 40 (1991) 143202.

[32] Dubois.D. and H.Prade, "What are fuzzy rules and how to use them", Fuzzy Sets and Systems 84 (1996) 169185 (special issue in memory of Prof A.Kaufmann).

[33] Dubois D., Prade H. and Ughetto L. (1997). Checking the coherence and the reduncancy of fuzzy knowledge bases, IEEE Trans. on Fuzzy Systems, 5, 398-417.

[34] Erich Peter Klement, Radko Mesiar and E.Endre Pap, " Triangular norms. Position paper I: Basic Analytical and Algebric Properties", Fuzzy Sets and Systems 143 (2004) 5 - 26.

[35] Erich Peter Klement, Radko Mesiar and E.Endre Pap, "Triangular norms - Position paper II: General constructions and parameterized families", Fuzzy Sets and Systems 145 (2004) 411 - 438.

[36] Erich Peter Klement, Radko Mesiar and E.Endre Pap, " Triangular norms - Position paper III: continuous t-norms", Fuzzy Sets and Systems 145 (2004) 439 - 454.

[37] Fodor.J.C., and M. Roubens (1994), Fuzzy Preference Modelling and Multi-criteria Decision Support, Kluwer, Dordrecht.

[38] Janos C Fodor, On Bisymmetry operations on the Unit Interval , Proc. IFSA '97, 7th Int. Fuzzy Systems Association World Congress (1997) 277 - 280.

[39] J. Fodor and J.-L. Marichal, On nonstrict means, Aequationes Mathematicae 54 (3) (1997) 308-327.

[40] S. Galichet L. Foulloy, Fuzzy Controllers: synthesis and equivalences, IEEE Trans. Fuzzy Systems 3 (1995) 140-148.

[41] Gegov,A. and Frank,P., Reduction of Multidimensional Relations in Fuzzy Control Systems , Systems and Control Letters, 25 307 - 313.

[42] George.J.Klir and Bo Yuan, Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice Hall Inc., Englewoods, 1995.

[43] J.S.Glower and J.Munighan, Designing Fuzzy Controllers from a Variable Structures Standpoint , IEEE Trans. Fuzzy Systems 5 (1) (1997) 138 - 144.

[44] Guven.K.Mustafa and Kevin M. Passino, Avoiding Exponential Parameter Growth in Fuzzy Systems , IEEE Trans. Fuzzy Systems 9 (1) (2001) 194 - 199.

[45] C.C.Hung and B.R.Fernandez, Minimizing Rules of Fuzzy Logic System by using a Systematic Approach , FUZZ-IEEE '93, 38 - 48.

[46] H.Ishibuchi, K. Nozaki, N.Yamamoto and H.Tanaka, Selecting fuzzy if-then rules for classification problems using genetic algorithms , IEEE Trans. on Fuzzy Systems 3 (1995) 260 - 270.

[47] Jia,L., and Zhang.X., Identification of Multivariable Fuzzy Systems through Fuzzy Cell Mapping , Preprints from IFAC World Congress (1993) 389 - 393.

[48] C.L.Karr, Applying genetics to fuzzy logic , AI Expert 6 (1991) 38 - 43.

[49] L. T. Koczy, Fuzzy if-then rule models and their transformation into one another , IEEE Trans. on Syst. Man and Cyber. Part A 26 (5) (1996) 621 - 637.

[50] L.T.Koczy and A.Zorat, Optimal Fuzzy Rule Bases - the Cat and Mouse Problem, FUZZ - IEEE '96, 1865 - 1870.

[51] L.T.Koczy and K.Hirota, Size Reduction by Interpolation in Fuzzy Rule Bases, IEEE Trans. on Syst. Man and Cyber. Part B 27 (1) (1997) 14 - 25.

[52] A.Krone, P.Krause and T.Slawinski, A New Rule Reduction Method for finding Interpretable and Small Rule Bases in High Dimensional Search Spaces, FUZZ-IEEE 2000, 694 - 699.

[53] M.A.Lee and H.Takagi, Integrating design stages of fuzzy systems using genetic algorithms, FUZZ-IEEE '93, 612 - 617.

[54] Y.M.Li, Z.K.Shi, Z.H.Li, Approximation Theory of Fuzzy Systems based upon genuine many-valued implication - SISO Case, Fuzzy Sets and Systems, 130 (2) (2002) 147 - 158.

[55] Y.M.Li, Z.K.Shi, Z.H.Li, Approximation Theory of Fuzzy Systems based upon genuine many-valued implication - MIMO Case, Fuzzy Sets and Systems, 130 (2) (2002) 159 - 174.

[56] P.Magrez, Ph.Smets, Fuzzy Modus Ponens: a new model suitable for applications in Knowledge-bases systems, Int. Jl. of Intelligent Systems 4 (1989) 181 - 200.

[57] E.H. Mamdani and S.Assilian, "An experiment in Linguistic Synthesis with a Fuzzy Logic Controller", Int. Jl. of Man - Machine Studies 7 (1975) 1 - 13.

[58] Mamdani. E. H., Application of Fuzzy Logic in approximate Reasoning using Linguistic Systems, IEEE Transactions on Computers 26 (1997) 1182 - 1191.

[59] J.M.Mendel and Q.Liang, Comments on "Combinatorial rule explosion eliminated by a fuzzy rule configuration", IEEE Trans. Fuzzy Systems 7 (1999) 369 - 371.

[60] N.N.Morsi, A.A.Fahmy, On generalised modus ponens with multiple rules and a residuated implication , Fuzzy Sets and Systems 129 (2) (2002) 267 - 274.

[61] G.C.Mouzouris and J.M.Mendel, Designing fuzzy logic systems for uncertain environments using a singular-value-QR decomposition method , in FUZZ-IEEE '96, 295 - 301.

[62] Raju,G. and Zhou,J, Fuzzy Logic processes controller, IEEE Intl. Conf. On Systems Engg., (1990) 145 - 147.

[63] H.Roubos and M.Setnes, Compact and Transparent Fuzzy Models and Classifiers through Iterative Complexity Reduction, IEEE Trans. Fuzzy Systems 9 (4) (2001) 516 - 524.

[64] R.Rovatti, R.Guerrieri and G.Baccarani, An Enhanced Two-Level Boolean Synthesis Methodology for Fuzzy Rules Minimization, IEEE Trans. Fuzzy Systems 3 (1995) 288 - 299.

[65] E.H.Ruspini, A new approach to clustering, Information and Control 15 (1) (1969) 22 - 32.

[66] M.Setnes, A.Koene, R.Babuska and P.Bruijin, Data-driven Initialisation and structure learning in fuzzy neural networks, FUZZ-IEEE '98, 1147 - 1152.

[67] 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 28 (3) (1998) 376 - 386.

[68] M.Setnes, V.Lacrose and A.Titli, Complexity Reduction Methods for Fuzzy Systems in Fuzzy Algorithms for Control, Eds. H.Verbruggen, H.J.Zimmermann and R.Babuska, (1999) 185 - 218.

[69] 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, 31 (2) (2001) 199 - 206.

[70] C.T.Sun, Rule-Base structure identification in an Adaptive-Network-Based Fuzzy Inference System, IEEE Trans. Fuzzy Systems 2 (1) (1994) 64 - 73.

[71] T. Takagi and M. Sugeno, Fuzzy identification of systems and its applications to modeling and control, IEEE Trans. Systems, Man and Cybernetics 15 (1) (1985) 116 - 132.

[72] K.Tanaka, T.Taniguchi and H.O.Wang, Generalised Takagi-Sugeno Fuzzy Systems : Rule Reduction and Robust Control, FUZZ-IEEE 2000, 688 - 693.

[73] M. A. Taylor, The Generalized Equation of Bisymmetry: Solutions Based on Cancellative Abelian Monoids, Aequationes Math. 57 (1999) 288 - 302.

[74] M. A. Taylor, The Aggregation Equation : Solutions with Non-intersecting Partial Functions in Functional Equations - Results and Advances, Eds. Z. Daroczy and Z. Pales, Kluwer Academic Publishers (2000) 1 - 10.

[75] E.Trillas and L.Valverde, On Implication and Indistinguishability in the setting of Fuzzy Logic, in Management Decision Support Systems Using Fuzzy Sets and Possibility Theory, Verlag TUV-Rhineland, 1985, pp. 198 - 212.

[76] E.Trillas, C.Alsina, On the Law $(p \wedge q) \to r \equiv (p \to r) \vee (q \to r)$ in Fuzzy Logic, IEEE Trans. Fuzzy Systems 10 (1) (2002) 84 - 88.

[77] I.B.Turksen and Z.Zhong, "An Approximate Reasoning schema based on Similarity measures and Interval valued fuzzy sets", IEEE Trans. on Systems, Man and Cybernetics, 18 (6) (1988) 1049 - 1056.

[78] Jose Villar, Conditions for Equivalence between the Compositional Rule of Inference and the Compatibility Modification Inference, FUZZ-IEEE '96 (1996) 444 - 449.

[79] L.X.Wang and J.M.Mendel, Fuzzy basis functions, universal approximation and orthogonal least squares learning, IEEE Trans. on Neural Networks, 3 (1993) 807 - 814.

[80] Yeung Yam, Fuzzy Approximation Via Grid Point Sampling and Singular Value Decomposition, IEEE Trans. on Syst. Man and Cyber. Part B 27 (6) (1997) 933 - 951.

[81] Yeung Yam, Baranyi, P. and Yang, C.T., Reduction of Fuzzy Rule Base via Singular Value Decomposition ,IEEE Transaction on Fuzzy Systems, 7 (2) (1999) 120-132.

[82] John Yen and Liang Wang, Simplifying fuzzy rule-based models using orthogonal transformation methods, IEEE Trans. on Syst. Man and Cyber. Part B 29 (1) (1999) 13 - 24.

[83] R.R.Yager, Alternative Structures for Knowledge Representation in Fuzzy Logic Controllers, in "Fuzzy Control Systems", Eds. A.Kandel and G.Langholz, CRC Press (1994) 99 - 137.

[84] Yager R.R., Rybalov A., Uninorm aggregation operators, Fuzzy Sets and Systems, 80 (1996) 111–120.

[85] Yager. R., "On some new classes of implication operators and their role in Approximate Reasoning", Info Sci. 167 (2004) 193 – 216.

[86] L.A.Zadeh, Outline of a new approach to the analysis of complex systems and decision processes , IEEE Trans. Systems, Man and Cybernetics 3 (1973) 28 - 44.

Proof. Proof of Theorem 6

Claim $T_1 \equiv T_2$*:* Let $a_1 = b_2 = 1$. Then $\forall a_2, b_1 \in I$, we have

$$
\begin{aligned}
LHS &= T_1(T_2(1, b_1), T_2(a_2, 1)) = T_1(b_1, a_2) \\
RHS &= T_2(T_3(1, a_2), T_3(b_1, 1)) = T_2(a_2, b_1) = T_2(b_1, a_2)
\end{aligned}
$$

which implies $T_1(b_1, a_2) = T_2(b_1, a_2) \ \forall a_2, b_1 \in I$ and thus $T_1 \equiv T_2 = T$. Now (37) becomes

$$
T(T(a_1, b_1), T(a_2, b_2)) = T(T_3(a_1, a_2), T_3(b_1, b_2)) \tag{61}
$$

Now, since T is a t-norm,

$$
T(T(a_1, b_1), T(a_2, b_2)) = T(T(a_1, a_2), T(b_1, b_2)) \tag{62}
$$

and we have from (61) and (62) that $T_1 \equiv T_2 \equiv T_3$ on I^2.   □

Proof. Proof of Theorem 8

Let us consider (40) from *Group 2*. Let $a_1 = b_2 = 1$ and $a_2, b_1 \in (0, 1)$. Then we have that

$$
\begin{aligned}
LHS &= T_1(T_2(1, b_1), T_2(a_2, 1)) = T_1(b_1, a_2) \\
RHS &= T_2(S(1, a_2), S(b_1, 1)) = T_2(1, 1) = 1
\end{aligned}
$$

which implies that $T_1(b_1, a_2) = 1$, with $a_2, b_1 \in (0, 1)$, which is absurd. Similarly, all the other equations, (39), (41) - (44) in *Group 2* can be shown to have no solutions.
                      □

Proof. Proof of Theorem 10

We give the proof for $M = M_7$. The proofs for $M = M_5, M_6$ and M_8 are similar.

$$
\begin{aligned}
LHS \ of \ (45) &= T[\inf_x S(A_1(x), A'(x)), \\
& \qquad \inf_x S(A_2(x), A'(x))] \tag{63} \\
&= \inf_x T[S(A_1(x), A'(x)), \\
& \qquad S(A_2(x), A'(x))] \tag{64} \\
&= \inf_x S(T[A_1(x), A_2(x)], A'(x)) \tag{65} \\
&= M_7[T(A_1, A_2), A'] \\
&= RHS \ of \ (45)
\end{aligned}
$$

Since $T(\inf_x a_x, \inf_x b_x) \equiv \inf_x T(a_x, b_x)$ iff $T = \min$, we have that (64) is equivalent to (63) iff $T = \min$. Also since any t-conorm S is distributive over $T = \min$ [20], we obtain (65) from (64).   □