Chapter 1
Similarity Based Reasoning Fuzzy Systems and Universal Approximation

Sayantan Mandal and Balasubramaniam Jayaram

Abstract I
n this work, we show that fuzzy inference systems based on Similarity Based Reasoning (SBR) where the modification function is a fuzzy implication is a universal approximator under suitable conditions on the other components of the fuzzy system.

Key words: Similarity Based Reasoning, Fuzzy implications, Universal approximation.

1.1 Introduction

The term *approximate reasoning* (AR) refers to methods and methodologies that enable reasoning with imprecise inputs to obtain meaningful outputs [**?**]. AR schemes involving fuzzy sets are one of the best known applications of fuzzy logic in the wider sense. Fuzzy Inference Systems (FIS) have many degrees of freedom, viz., the underlying fuzzy partition of the input and output spaces, the fuzzy logic operations employed, the fuzzification and defuzzification mechanism used, etc. This freedom gives rise to a variety of FIS with differing capabilities. One of the important factors considered while employing an FIS is its approximation capability. Many studies have appeared on this topic and due to space constraints, we only refer the readers to the following exceptional review on this topic [**?**] and the references therein.

In this work, we consider a Similarity Based Reasoning (SBR) FIS where similarity between the inputs and the antecedents is used to subsequently

Department of Mathematics,
Indian Institute of Technology Hyderabad,
Yeddumailaram-502205, INDIA
e-mail: ma10p002@iith.ac.in, jbala@iith.ac.in

modify the consequents to obtain a final output. Such inference schemes are also known as plausible reasoning scheme [?]. After detailing the inference mechanism in an SBR, we show that when the modification functions are modeled based on fuzzy implications, under suitable conditions on the other components of an SBR, the FIS based on SBR does become a universal approximator, i.e., can approximate a continuous function over a compact set to arbitrary accuracy. Also we deal only with single variable functions, alternately where the rule base consists of Single Input Single Output (SISO) rules.

1.2 Preliminaries

We assume that the reader is familiar with the classical results concerning fuzzy set theory and basic fuzzy logic connectives, but to make this work more self-contained, we introduce some notations, concepts and results employed in the rest of the work.

1.2.1 Fuzzy Sets

If X is a non-empty set then we denote by $\mathcal{F}(X)$ the fuzzy power set of X, i.e., $\mathcal{F}(X) = \{A | A : X \to [0,1]\}$.

Definition 1. A fuzzy set A is said to be

- *normal* if there exists an $x \in X$ such that $A(x) = 1$,
- *convex* if X is a linear space and for any $\lambda \in [0,1]$, $x, y \in X$, $A(\lambda x + (1-\lambda)y) \geq \min\{A(x), A(y)\}$.

Definition 2. For an $A \in \mathcal{F}(X)$, the *Support*, *Height*, *Kernel* and *Ceiling* of A are denoted, respectively, as Supp A, Hgt A, Ker A and Ceil A and are defined as:

$$\begin{aligned} \text{Supp } A &= \{x \in X | A(x) > 0\} \ , \\ \text{Hgt } A &= \sup\{A(x) | x \in X\} \ , \\ \text{Ker } A &= \{x \in X | A(x) = 1\} \ , \\ \text{Ceil } A &= \{x \in X | A(x) = \text{Hgt } A\} \ . \end{aligned}$$

A is said to be *bounded* if Supp A is a bounded set. Note that for a normal fuzzy set Ker $A =$ Ceil A.

We denote the space of fuzzy sets which are bounded, normal, convex and continuous as $\mathcal{F}_{BNCC}(X)$. Clearly $\mathcal{F}_{BNCC}(X) \subseteq \mathcal{F}(X)$.

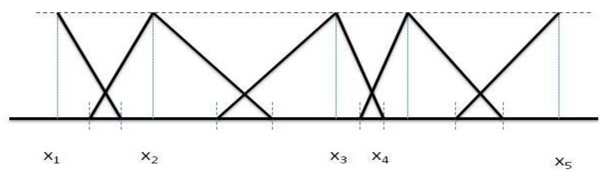

Fig. 1.1 An Illustrative Example for $\frac{1}{3}$-type partition in Definition **??**

.

Definition 3. Let $\mathcal{P}$ be an arbitrary collection of fuzzy sets of X, i.e, $\mathcal{P} = \{A_k\}_{k=1}^n \subseteq \mathcal{F}(X)$. $\mathcal{P}$ is said to form a *fuzzy partition* on X if

$$X \subseteq \bigcup_{k=1}^{n} \text{Supp } A_k \ .$$

In literature, a partition $\mathcal{P}$ of X as defined above is also called a **complete** partition.

Definition 4. A fuzzy partition $\mathcal{P} = \{A_k\}_{k=1}^n \subseteq \mathcal{F}(X)$ is said to be

- **consistent** if $A_k(x) = 1$ then $A_j(x) = 0$ for any $j \neq k$.
- **Ruspini partition** if $\sum_{k=1}^{n} A_k(x) = 1$ for every $x \in X$.

Definition 5. Let $\{x_k\}_{k=1}^n$ be a classical partition of X, i.e., $X = \bigcup_{k=1}^{n-2} [x_k, x_{k+1}) \cup [x_{n-1}, x_n]$. If $\mathcal{P} = \{A_k\}_{k=1}^n$ be a fuzzy partition of the space X in such a way that

- each A_k is normal at $x_k \in X$, i.e., $A_k(x_k) = 1$,
- Supp $A_k = \left(x_{k-1} + \frac{x_k - x_{k-1}}{3}, x_{k+1} - \frac{x_{k+1} - x_k}{3}\right)$ for $k = 2, \ldots, n-1$, while Supp $A_1 = \left[x_1, x_2 - \frac{x_2 - x_1}{3}\right)$ and Supp $A_n = \left(x_{n-1} + \frac{x_n - x_{n-1}}{3}, x_n\right]$,

we call this type of partition as $\frac{1}{3}$**-type partition**.

For instance, see Fig. **??** for $n = 5$.

1.2.2 Defuzzification

Often there is a need to convert a fuzzy set to a crisp value, a process which is called *Defuzzification*. This process of defuzzification can be seen as a mapping $g : \mathcal{F}(X) \longrightarrow X$. There are many types of defuzzification techniques available in the literature, see [**?**] for a good overview. In this work, we use the following defuzzifier extensively.

Example 1. For an $A \in \mathcal{F}(X)$, the *First of Maxima* (FOM) defuzzifier gives as output the smallest of all those values in X with the highest membership value, which can be mathematically expressed as

$$\mathrm{FOM}(A) = \min\{x | A(x) = \max_{w} A(w)\} . \tag{1.1}$$

Similarly the *Last of Maxima* (LOM) defuzzifier is defined as

$$\mathrm{LOM}(A) = \max\{x | A(x) = \max_{w} A(w)\} . \tag{1.2}$$

1.2.3 Fuzzy Logic Connectives

Definition 6 ([?]). A binary operation $T \colon [0,1]^2 \to [0,1]$ is called a *t-norm*, if it is increasing in both variables, commutative, associative and has 1 as the neutral element.

Definition 7 ([?]). A function $I \colon [0,1]^2 \to [0,1]$ is called a *fuzzy implication* if it is decreasing in the first variable, increasing in the second variable and $I(0,0) = 1$, $I(1,1) = 1$, $I(1,0) = 0$. The set of all fuzzy implications will be denoted by $\mathcal{I}$.

Definition 8 ([?]). A fuzzy implication $I \colon [0,1]^2 \to [0,1]$ is said to

- satisfy the *ordering property*, if

$$I(x,y) = 1 \Longleftrightarrow x \leq y \; , \qquad x, y \in [0,1] \; . \tag{OP}$$

- be a *positive fuzzy implication* if $I(x,y) > 0$, for all $x, y \in (0,1)$.

1.3 Fuzzy Inference Mechanism

Given two non-empty classical sets $X, Y \subsetneq \mathbb{R}$, a fuzzy Single Input Single Output (SISO) IF-THEN rule is of the form:

$$\textbf{IF } \tilde{x} \text{ is } A \textbf{ THEN } \tilde{y} \text{ is } B, \tag{1.3}$$

where $\tilde{x}$, $\tilde{y}$ are the linguistic variables and $A \in \mathcal{F}(X), B \in \mathcal{F}(Y)$ are the linguistic values taken by the linguistic variables. A knowledge base consists of a collection of such rules. Hence, we consider a rule base of n SISO rules which is of the form:

$$\textbf{IF } \tilde{x} \text{ is } A_i \textbf{ THEN } \tilde{y} \text{ is } B_i \ , \tag{1.4}$$

where $\tilde{x}$, $\tilde{y}$ and $A_i \in \mathcal{F}(X), B_i \in \mathcal{F}(Y), i = 1, 2, \ldots n$ are as mentioned above.

As an example, consider the rule

IF *Temperature* is *High* **THEN** *Fanspeed* is *Medium*.

Here *Temperature* and *Fanspeed* are the linguistic variables and *High*, *Medium* are the linguistic values taken by the linguistic variables in a suitable domain. Now given a single SISO rule (**??**) or a rule base (**??**) and given any input " $\tilde{x}$ is A'" , the main objective of an inference mechanism is to find B' such that " $\tilde{y}$ is B' ". Many types of inference mechanisms are available to us in [**?**], [**?**], [**?**], etc. Here we consider only the case of Similarity Based Reasoning.

1.4 Similarity Based Reasoning (SBR)

Consider the fuzzy if-then rule (**??**). Let the given input be $\tilde{x}$ is A'. Inference in Similarity Based Reasoning (SBR) schemes in AR is based on the calculation of a measure of compatibility or similarity $M(A, A')$ of the input A' to the antecedent A of the rule, and the use of a modification function J to modify the consequent B, according to the value of $M(A, A')$.

Some of the well known examples of SBR are Compatibility Modification Inference (CMI) [**?**], "Approximate Analogical Reasoning Scheme" (AARS) in [**?**] and "Consequent Dilation Rule" (CDR) in [**?**], Smets and Magrez [**?**], Chen [**?**], etc. In this section, we detail the typical inferencing mechanism in SBR, but only in the case of SISO fuzzy rule bases.

1.4.1 Matching function M

Given two fuzzy sets, say A, A', on the same domain, a matching function M compares them to get a degree of similarity, which is expressed as a real in the $[0, 1]$ interval. We refer to M as the Matching Function in the sequel. Formally, $M : \mathcal{F}(X) \times \mathcal{F}(X) \to [0, 1]$.

Example 2. Let X be a non-empty set and $A, A' \in \mathcal{F}(X)$. Below we list a few of the matching functions employed in the literature.

- Zadeh [**?**]: $M_{\mathbf{Z}}(A, A') = \max_{x \in X} \min(A(x), A'(x))$.

- Magrez - Smets [**?**]: Given a fuzzy negation N,

$$M_{\mathbf{M}}(A, A') = \max_{x \in X} \min(N(A(x)), A'(x)).$$

- Measure of Subsethood [**?**]: For an $I \in \mathcal{FI}$,

$$M_{\mathbf{S}}(A, A') = \min_{x \in X} I(A'(x), A(x)).$$

Definition 9. Let $\mathcal{F}^* \subseteq \mathcal{F}(X)$ be an arbitrary collection (not necessarily a fuzzy partition) of fuzzy sets on X. M is said to be **consistent with** $\mathcal{F}^*$ if for any $A \in \mathcal{F}^*$,

$$M(A, A) = 1. \tag{MCF}$$

Definition 10. Let $\mathcal{P} = \{A_k\}_{k=1}^n \subseteq \mathcal{F}^*$ be the given fuzzy partition of X. Let $A' \in \mathcal{F}^*$. M is said to be **consistent with** $\mathcal{P}$ (and $\mathcal{F}^*$) if

$$\sum_{k=1}^{n} M(A', A_k) \leq 1. \tag{MCP}$$

Definition 11. The matching function M is said to be **Strong** if

$$\text{Ker } A \subseteq \text{Ker } B \text{ or Ker } B \subseteq \text{Ker } A \Longrightarrow M(A, B) = 1 \tag{MS}$$

Example 3. Let $X \subsetneq \mathbb{R}$ be any bounded interval and $\mathcal{F}^* = \mathcal{F}_{BNCC}(X)$. For a given fuzzy partition $\mathcal{P} = \{A_k\}_{k=1}^n \subseteq \mathcal{F}_{BNCC}(X)$, we define a maching function as,

$$M_{\mathcal{P}}(A_k, A') = \frac{Area(A' \cap A_k)}{Area(A')}, \quad A' \in \mathcal{F}_{BNCC}(X). \tag{1.5}$$

Clearly M satisfies (**??**), (**??**) and (**??**).

Example 4. Let $X \subsetneq \mathbb{R}$ be any bounded interval. Let the antecedent fuzzy sets $\{A_k\}_{k=1}^n = \mathcal{P}_X \subseteq \mathcal{F}^*(X)$ partition the input space X such that it forms a partition of the type defined in Definition **??**.

Now, if $x' \in X$ is the input let $A' \in \mathcal{F}(X)$ be the fuzzified input such that A' attains normality at x', i.e., $A'(x') = 1$. Then the matching function defined as $M(A', A) = A(x')$ for any $A \in \mathcal{F}(X)$ has the property (**??**).

1.4.2 Modification Function J

Let A' be the fuzzy input and $s = M(A, A') \in [0, 1]$, a measure of the compatibility of A' to A.

The modification function J is again a function from $[0, 1]^2$ to $[0, 1]$ and, given the rule (**??**), modifies $B \in \mathcal{F}(Y)$ to $B' \in \mathcal{F}(Y)$ based on s, i.e., the consequent in SBR, using the modification function J, is given by

$$B'(y) = J(s, B(y)) = J(M(A, A'), B(y)), \quad y \in Y.$$

In AARS [?] the following modification operators have been used:

(i) $J_{\mathbf{ML}}(s, B) = B'(x) = \min\{1, B(x)/s\}$, $x \in X$;
(ii) $J_{\mathbf{MVR}}(s, B) = B'(x) = s \cdot B(x)$, $x \in X$.

In CMI [?] and CDR [?] J is taken to be a fuzzy implication operator. In fact, $J_{\mathbf{ML}}(s, B) = I_{\mathbf{GG}}(s, B)$, where $I_{\mathbf{GG}}$ is the Goguen implication [?].

1.4.3 Aggregation Function G

In the case of multiple rules

$$R_i\text{: IF } \tilde{x} \text{ is } A_i \text{ THEN } \tilde{y} \text{ is } B_i, \qquad i = 1, 2, \ldots, m,$$

we infer the final output by aggregating over the rules, using an associative operator $G\colon [0,1]^2 \to [0,1]$ as follows:

$$B'(y) = G_{i=1}^{m} \Big(J\big(M(A_i, A'), B_i(y)\big) \Big), \quad y \in Y. \tag{1.6}$$

Usually, G is a t-norm, t-conorm or a uninorm [?].

1.5 Fuzzy Systems $\mathcal{F}$ based on SBR

An SBR fuzzy inference system can be represented by the hexatuple $\mathbb{F} = \{\mathcal{R}(A_i, B_j), f, M, J, G, g\}$ where

- $\mathcal{R}$ is the fuzzy if-then rule base formed from the fuzzy partitions $\{A_i\}, \{B_j\}$ on X, Y, respectively,
- $f : X \longrightarrow \mathcal{F}(X)$ is called the fuzzification mapping that maps an element $x \in X$ to a fuzzy set of $\mathcal{F}(X)$,
- M is matching function,
- J is modification function,
- G is aggregation function, and
- $g : \mathcal{F}(Y) \to Y$ is any defuzzifier, that converts the output fuzzy set to a crisp value $y \in Y$.

We consider $\mathbb{F}$ with the following assumptions on the different components / elements.

1.5.1 The Fuzzy Partitions A_i, B_j

Let $X, Y \subsetneq \mathbb{R}$ be arbitrary but fixed and let $\mathcal{F}^*(Z) = \mathcal{F}_{BNCC}(Z)$, where $Z = X$ or Y.

Let the antecedent fuzzy sets $\{A_k\}_{k=1}^{n} = \mathcal{P}_X \subseteq \mathcal{F}^*(X)$ partition the input space X such that it forms a partition of the type defined in Definition **??**, which also implies it is complete.

Similarly, let the consequent fuzzy sets $\{B_j\}_{j=1}^{m} = \mathcal{P}_Y \subseteq \mathcal{F}^*(Y)$ form a complete and Ruspini partition of the output space Y.

1.5.2 The Fuzzified Input A'

Let us consider a fuzzification $f : X \longrightarrow \mathcal{F}^*(X)$ that maps $x' \in X$ to a fuzzy set of $A' \in \mathcal{F}^*(X) = \mathcal{F}_{BNCC}(X)$ such that

$$\text{Supp }(f(x') = A') \cap \text{Supp } A_k \neq \emptyset,$$

for some $A_k \in \mathcal{P}_X$. Moreover it is assumed that A' intersects only two of the adjacent fuzzy sets A_k i.e, Supp $A' \cap$ Supp $A_k \neq \emptyset$ if and only if $k = m, m+1$ for some $m \in \mathbb{N}_{n-1}$.

Note that it is with this fuzzified input A' the antecedents A_i of the different rules are matched against.

Example 5. Let $\{x_k\}_{k=1}^{n}$ be a crisp partition of X. Let$\{A_k\}_{k=1}^{n}$ partitioning the input space X be such that $A_k \in \mathcal{P}_X$ and forms a fuzzy partition of the type defined in Definition **??**. Then if we take

$$|\text{ Supp } A'| \leq \frac{1}{3} \cdot \min_{i=1}^{l} \Big\{ |x_{i+1} - x_i| \Big\},$$

then A' intersects atmost two of the adjacent fuzzy sets A_k.

1.5.3 The Operations M, J, G

We choose a matching function M such that M is Consistent w.r.to the partition $\mathcal{P}_X$ given in Section **??**, i.e M satisfies both (**??**) and (**??**).

We choose the modification function J to be a fuzzy implication, i.e., $J \in \mathcal{FI}$. For notational convenience we will denote it by " $\longrightarrow$ " in the sequel.

The aggregation function G is any t-norm T.

1.5.4 The Fuzzy Output B'

With the above assumptions, the output fuzzy set B' for a given crisp input x' (or fuzzy input A') takes the form as given in the following lemma:

Lemma 1. *With the operations of of the SBR FIS* (**??**) *as in Sections* **??** - **??** *the fuzzy output of the SBR FIS* (**??**)*, for a given input $x' \in X$ is given by*

$$B'(y) = T\left[s_m \longrightarrow B_m(y), s_{m+1} \longrightarrow B_{m+1}(y)\right] \ , \tag{1.7}$$

where $s_m = M\left(A', A_m\right)$ and $s_{m+1} = M\left(A', A_{m+1}\right)$.

Proof. With the above operations M, J, G the fuzzy output for a given input $x' \in X$ is given by (**??**) as follows:

$$\begin{aligned} B'(y) = &T\big[M(A', A_1) \longrightarrow B_1(y)), M(A', A_2) \longrightarrow B_2(y), \ldots, \\ &\ldots, M(A', A_n) \longrightarrow B_n(y)\big]. \end{aligned}$$

We can write the above as

$$B'(y) = T_{k=1}^{n}[M(A', A_k) \longrightarrow B_k(y)] \ . \tag{1.8}$$

By the choice of our fuzzification based on our above notations on A', A_k, viz., that A' intersects only two adjacent fuzzy sets among the $\{A_k\}$, say A_m, A_{m+1}, we have that $M(A', A_k) = 0$ for all $k \neq m, m+1$. Note also that $I(0, y) = 0 \longrightarrow y = 1$ for any $y \in [0, 1]$. Now, the fuzzy output $B'(y)$ for any $y \in Y$ which is given by (**??**) becomes

$$\begin{aligned} B'(y) &= T_{k=1}^{n}[M(A', A_k) \longrightarrow B_k(y)] \ , \\ &= T\Big[T_{k \neq m, m+1}\big(M(A', A_k) \longrightarrow B_k(y)\big), \\ &\qquad M\left(A', A_m\right) \longrightarrow B_m(y), M\left(A', A_{m+1}\right) \longrightarrow B_{m+1}(y)\Big] \\ &= T\Big[M\left(A', A_m\right) \longrightarrow B_m(y), M\left(A', A_{m+1}\right) \longrightarrow B_{m+1}(y)\Big] \\ &= T\left[s_m \longrightarrow B_m(y), s_{m+1} \longrightarrow B_{m+1}(y)\right] \ = (\mathbf{??}) \ . \end{aligned}$$

1.5.5 The Defuzzified Output $g(x')$

We have chosen the modification function J to be a fuzzy implication, i.e., $J = I \in \mathcal{FI}$. Assuming that the considered modification function J has (**??**), we define the defuzzification function g appropriately so that g is continuous. In the following, we discuss the explicit formulae for g. Note that g is also known as the system function of the fuzzy system $\mathbb{F}$ [**?**], [**?**].

1.6 SBR Fuzzy Systems and Universal Approximation

In this section, we show that $\mathbb{F} = \{\mathcal{R}(A_i, B_i), M, J, G, g\}$ such that the fuzzy partitions $\{A_k\}, \{B_k\}$ and the operations M, J, G, g as given in Sections **??** - **??** are universal approximators, i.e., they can approximate any continuous function over a compact set to arbitrary accuracy.

Theorem 1. *For any continuous function $h\colon [a, b] \to \mathbb{R}$ over a closed interval and an arbitrary given $\epsilon > 0$, there is an SBR fuzzy system $\mathbb{F} = \{\mathcal{R}(A_i, B_i), M, J, G, g\}$ with M having the property* (**??**) *w.r.to $\mathcal{P}_X = \{A_i\}$, J having* (**??**)*, G being a t-norm and g as given in* (**??**) *or* (**??**) *such that* $\max\limits_{x\in[a,b]} |h(x) - g(x)| < \epsilon$.

Proof. We prove this result in the following steps.

Step I : Choosing the points of normality

Since h is coninuous over a closed interval $[a, b]$, h is uniformly continuous on $[a, b]$. Thus for a given $\epsilon > 0$ there exists $\delta > 0$ such that

$$|w - w'| < \delta \Longrightarrow |h(w) - h(w')| < \frac{\epsilon}{2}\ .$$

Step I *(a)*: A Coarse Initial Partition

With the $\delta = \delta(\epsilon)$ defined above and taking $l = \left\lceil \frac{b-a}{\delta} \right\rceil$ we now choose $w_i \in X, i = 1, 2, \ldots l$, such that $|w_i - w_{i+1}| < \delta$.

Let $z_i = h(w_i)$, the value h takes at the above chosen w_i, for $i = 1, 2, \ldots l$. We call these points w_i and z_i the points of normality on the input space and the output space respectively.

In Fig. **??**, *the points $w_1, w_2, \ldots, w_{11}$ and the points $z_1, z_2, \ldots z_8$ are the points of normality in the input and the output spaces, respectively.*

Step I *(b)*: Redundancy Removal and Reordering

Let us choose the distinct z_i's from the above and sort them in ascending order. Let $\sigma : \mathbb{N}_l \longrightarrow \mathbb{N}_k$ denote the above permuation map such that $z_i = u_{\sigma(i)}$, for $i = 1, 2, \ldots l$ and $u_j,\ j = 1, 2, \ldots, k$ are in ascending order.

By rearranging the z_i's in ascending order and renaming them we have obtained: $u_1 = z_1, u_2 = z_8, u_3 = z_6, u_4 = z_5, u_5 = z_7, u_6 = z_2, u_7 = z_4, u_8 = z_3$.

Step I *(c)*: Refinement of the input space partition:

Thus for each $i = 1, 2, \ldots, l$ we have $h(w_i) = z_i = u_{\sigma(i)}$. However, note that consecutive points of normality w_i, w_{i+1} in the input space need not be mapped to consecutive points of normality $u_{\sigma(i)}, u_{\sigma(i)+1}$ or $u_{\sigma(i)}, u_{\sigma(i)-1}$.

In Fig. **??**, *$h(w_1) = u_1$ and $h(w_2) = u_6$. Thus for the consecutive points w_1 and w_2 the function values are u_1 and u_6, which are not consecutive.*

To ensure the above, we further refine the input space partition. To this end, we refine every sub-interval $[w_i, w_{i+1}]$, for $i = 1, 2, \ldots l - 1$ as follows. Note that $h(w_{i+1}) = u_{\sigma(i+1)}$.

Refinement Procedure:

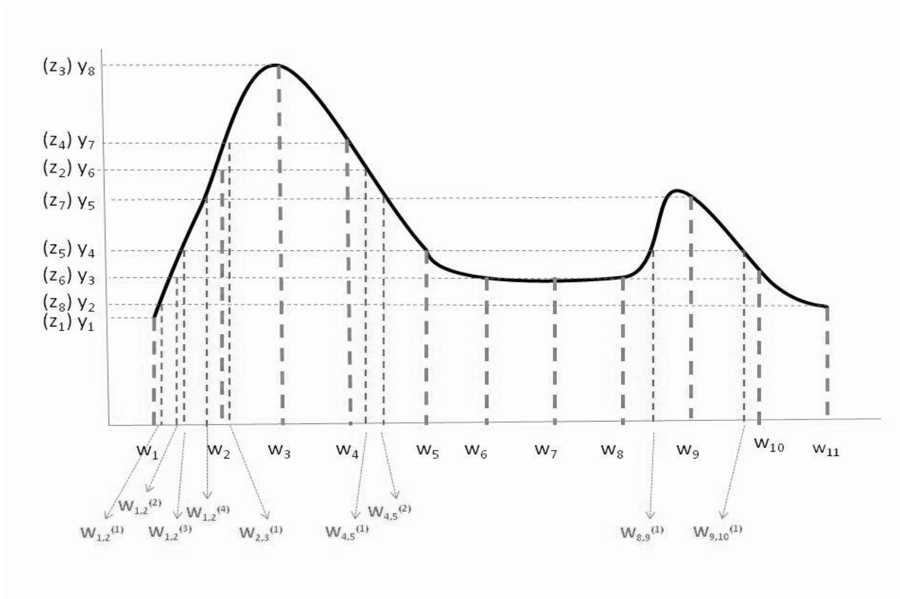

Fig. 1.2 An Illustrative Example for **Step I** in the proof of Theorem **??**.

For every $i = 1, 2, \ldots l-1$ do the following:

(i) If $u_{\sigma(i+1)} = u_{\sigma(i)+1}$ or $u_{\sigma(i)-1}$ then we do nothing.
(ii) Let $u_{\sigma(i+1)} = u_{\sigma(i)+p}$, where $p \geq 2$. For every $u \in \{u_{\sigma(i)+1}, u_{\sigma(i)+2}, \ldots, u_{\sigma(i)+p-1}\}$ we find a point $v \in [w_i, w_{i+1}]$ such that $h(v) = u$. Note that the existence of such a $v \in [w_i, w_{i+1}]$ is guaranteed by the continuity - essentially the ontoness - of the function h. If $u = u_{\sigma(i)+q}$, for some $1 \leq q \leq p-1$, then we denote the point v as $w^{(q)}_{i,i+1}$.
(iii) Similarly, let $u_{\sigma(i+1)} = u_{\sigma(i)-p}$, where $p \geq 2$. For every $u \in \{u_{\sigma(i)-1}, u_{\sigma(i)-2}, \ldots, u_{\sigma(i)-p+1}\}$ we find a $v \in [w_i, w_{i+1}]$ such that $h(v) = u$. Once again, if $u = u_{\sigma(i)-q}$, for some $1 \leq q \leq p-1$, then we denote v as $w^{(q)}_{i,i+1}$.

From Fig. **??**, *it can be seen that we have inserted points* $w^1_{1,2}, w^2_{1,2}, w^3_{1,2}, w^4_{1,2} \in [w_1, w_2]$. *Proceeding similarly, the following sub-intervals, shown in Fig.* **??**, *have been refined:* $[w_2, w_3], [w_4, w_5], [w_8, w_9]$ *and* $[w_9, w_{10}]$.

Step I *(d)*: Final Points of Normality:

Once the above process is done, we again rename the points of normality $w^{(q)}_{i,i+1}$ in the input space X in ascending order as $x_1, x_2, \ldots, x_n (n \geq l)$ and the $u_{\sigma(i)}$'s of the the output space as $y_1, y_2, \ldots y_k$.

Step II : Construction of the Fuzzy Partitions

In the next step, we construct fuzzy sets on both the input and output spaces with the above obtained x_i's and y_j's as the points of normality, as given below.

Step II *(a)*: Fuzzy Partition on the input space We construct l fuzzy sets such that

- each A_i is centered at x_i,
- Supp $A_i = \left(x_{i-1} + \frac{x_i - x_{i-1}}{3}, x_{i+1} - \frac{x_{i+1}-x_i}{3}\right)$ for $i = 2, \ldots, l-1$, while Supp $A_1 = \left[x_1, x_2 - \frac{x_2 - x_1}{3}\right)$ and Supp $A_l = \left(x_{l-1} + \frac{x_l - x_{l-1}}{3}, x_l\right]$,
- each A_i is normal at x_i, i.e., $A_i(x_i) = 1$,
- each A_i is a continuous convex fuzzy set,
- $\{A_i\}_{i=1}^{l}$ form a partition as defined in Definition **??**.

For instance, if each of the A_i's ($i = 2, \ldots, l-1$) is a triangular fuzzy set and A_1, A_l are half-triangular with all of them attaining normality at x_i then clearly we can construct $\{A_i\}_{i=1}^{l}$'s partitioning the input space X as in Definition **??** *and are continuous, convex, of finite support and $A_i(x_i) = 1$.*

Step II *(b)*: Fuzzy Partition on the output space

Now we have the output space partition points as $y_1, y_2, \ldots y_k$. We partition the output space such that $B_1, B_2, \ldots B_k$ form a Ruspini partition (as above) with $B_j(y_j) = 1, \quad j = 1, 2, \ldots k$. Here obviously,

$$|y_j - y_{j-1}| < \frac{\epsilon}{2}, \quad j = 1, 2, \ldots k.$$

Further, let the fuzzy sets $\{B_j\}_{j=1}^{k}$ be continuous, convex and of finite support along the same lines as the A_i's above, i.e., Supp $B_1 = [y_1, y_2)$, Supp $B_j = (y_{j-1}, y_{j+1}), j = 2, 3, \ldots k-1$, Supp $B_k = (y_{k-1}, y_k]$.

Step III: Construction of the *smooth* rule base

We construct the rule base with l rules of the following form:

$$\text{IF } x \text{ is } A_i \text{ THEN } y \text{ is } B_i\ , \quad i = 1, 2, \ldots l, \tag{1.9}$$

where the consequent B_i in the i-th rule is chosen such that $i = j$ is the index of that $y_j = h(x_i)$, where x_i is the point at which A_i attains normality.

Note that, since h is continuous, by the above assignment of the rules, we have that rules whose antecedents are adjacent also have adjacent consequents, i.e., for any $i = 1, 2, \ldots l-1$ we have Supp $B_i \cap$ Supp $B_{i+1} \neq \emptyset$. Thus the constructed rule base is smooth as defined in [**?**].

Step IV : Approximation capability of the output

Now we consider an SBR fuzzy system with Multiple SISO rules of the form (**??**). Let $x' \in X$ be the given input. Clearly, $x' \in [x_m, x_{m+1}]$ for some $m \in \mathbb{N}_l$. Now as in section **??**, we fuzzify x' in such a way that the fuzzified input A' (with $A'(x') = 1$) intersects atmost two of the A_i's, say, A_m, A_{m+1}.

For instance, one could take A' as in Example **??**.

So we have the following,

$$
\begin{aligned}
B'(y) = T\big[M\,(A', A_m) \longrightarrow B_m(y),\\
M\,(A', A_{m+1}) \longrightarrow B_{m+1}(y)\big]\\
= T\big[s_m \longrightarrow B_m(y), s_{m+1} \longrightarrow B_{m+1}(y)\big]\ ,
\end{aligned}
$$

where $s_m = M\,(A', A_m)$ and $s_{m+1} = M\,(A', A_{m+1})$. Note that by our assumption on M, we have that $s_m + s_{m+1} \leq 1$.

The output fuzzy set B' is given by (**??**). We consider the kernel of B', i.e., Ker $B' = \{y : B'(y) = 1\}$. We choose the defuzzified output y' such that it belongs to Ker B'.

Since T is a t-norm, we know that $T(p,q) = 1$ if and only if $p = 1$ and $q = 1$. Noting that J has (OP), i.e., $p \longrightarrow q = 1 \Leftrightarrow p \leq q$ and $s_m + s_{m+1} \leq 1$, we have

$$
\begin{aligned}
\text{Ker } B' &= \{y : B'(y) = 1\}\\
&= \{y : s_m \longrightarrow B_m(y) = 1\} \bigcap \{y : s_{m+1} \longrightarrow B_{m+1}(y) = 1\}\\
&= \{y : s_m \leq B_m(y)\} \bigcap \{y : s_{m+1} \leq B_{m+1}(y)\}\ .
\end{aligned}
$$

Let $\alpha_m = \min\{\alpha : s_m \longrightarrow \alpha = 1\}$ and $\beta_{m+1} = \min\{\beta : s_{m+1} \longrightarrow \beta = 1\}$. Since J has (OP), clearly $\alpha_m = s_m$ and $\beta_{m+1} = s_{m+1}$.

By the continuity and convexity of B_m, B_{m+1} there exist $a_m, b_m, a_{m+1}, b_{m+1}$ such that $B_m(a_m) = B_m(b_m) = s_m$ and $B_{m+1}(a_{m+1}) = B_{m+1}(b_{M=1}) = s_{m+1}$. By the monotonicity of the implication in the second variable, for every $y \in [a_m, b_m]$ we have that $s_m \to B_m(y) = 1$ and for every $y \in [a_{m+1}, b_{m+1}]$ we have that $s_{m+1} \to B_{m+1}(y) = 1$. Thus,

$$
\begin{aligned}
\{y : s_m \leq B_m(y)\} = [a_m, b_m]\ , \qquad \text{and}\\
\{y : s_{m+1} \leq B_{m+1}(y)\} = [a_{m+1}, b_{m+1}]\ .
\end{aligned}
$$

$$
\text{Hence, Ker } B' = \{y : B'(y) = 1\} = [a_m, b_m] \bigcap [a_{m+1}, b_{m+1}]. \tag{1.10}
$$

Claim: Ker $B' = [a_{m+1}, b_m] \neq \emptyset$.

Firstly, note that for any $s_m \in [0,1]$ by the normality of B_m we have that $B_m(y_m) = 1$ and hence $y_m \in \{y : s_m \leq B_m(y)\} = y_m \in [a_m, b_m] \neq \emptyset$. Similarly, $y_{m+1} \in [a_{m+1}, b_{m+1}] \neq \emptyset$. it suffices to show that $a_{m+1} \leq b_m$ from whence Ker $B' = [a_{m+1}, b_m]$.

Note that since $m < m+1$, $y_m < y_{m+1}$ and from $a_{m+1} \in$ Supp B_{m+1} we have that $y_m \leq a_{m+1} \leq y_{m+1}$. Similarly, $y_m \leq b_m \leq y_{m+1}$. Hence, $y_m \leq a_{m+1}, b_m \leq y_{m+1}$.

Since B_{m+1} is monotonic on $[y_m, y_{m+1}]$,

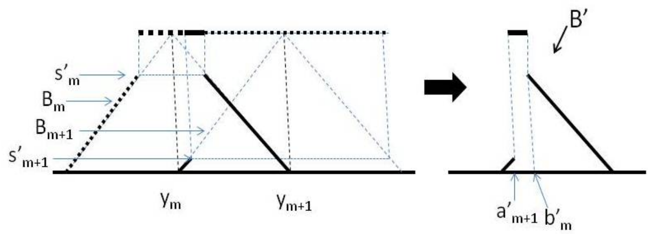

Fig. 1.3 The Output fuzzy set B'.

$$\begin{aligned} a_{m+1} > b_m &\text{ implies } B_{m+1}(a_{m+1}) \geq B_{m+1}(b_m) \\ &\text{ implies } s_{m+1} \geq 1 - B_m(b_m) \\ &\text{ implies } s_{m+1} \geq 1 - s_m \\ &\text{ implies } s_m + s_{m+1} \geq 1 \ . \end{aligned}$$

Since M satisfies (**??**), $s_m + s_{m+1} \leq 1$ and hence $s_m + s_{m+1} = 1$. Now,

$$\begin{aligned} s_m + s_{m+1} = 1 &\text{ implies } B_{m+1}(a_{m+1}) + B_m(b_m) = 1 \\ &\text{ implies } B_{m+1}(a_{m+1}) = 1 - B_m(b_m) \\ &\text{ implies } B_{m+1}(a_{m+1}) = B_{m+1}(b_m) \\ &\text{ implies } b_m \in [a_{m+1}, b_{m+1}], \text{ i.e., } a_{m+1} \leq b_m \ . \end{aligned}$$

Now, we define $g(x')$ as either of the following - (**??**) or (**??**):

$$y' = g(x') = FOM(B'(y)) = a_{m+1} \quad (1.11)$$
$$y' = g(x') = LOM(B'(y)) = b_m \quad (1.12)$$

Now from the above we have the system function as, $y' = g(x') = a_{m+1}$ or b_m. Now clearly, $a_{m+1}, b_m \in [y_m, y_{m+1}]$ and hence,

$$|y_m - g(x')| < \frac{\epsilon}{2} \quad \text{or} \quad |y_{m+1} - g(x')| < \frac{\epsilon}{2} \ .$$

WLOG, let $|y_m - g(x')| < \frac{\epsilon}{2}$ i.e, $|y_m - y'| < \frac{\epsilon}{2}$. Now since $x' \in [x_m, x_{m+1}]$, we have $|h(x') - y_m| < \frac{\epsilon}{2}$. Finally we have the following,

$$\begin{aligned} |g(x') - h(x')| &= |y' - h(x')| \\ &\leq |y' - y_m| + |y_m - h(x')| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} < \epsilon. \end{aligned}$$

Since x' is arbitrary we have, $\max_{x \in [a,b]} |h(x) - g(x)| < \epsilon$.

Remark 1. Note that with g as in (**??**) or (**??**) and since M satisfies (**??**), if $x' = x_k \in X$ we have $M(A', A_k) = 1$ and we obtain $B' = B_k$, i.e., $g(x') = y_k$ and the interpolativity of the inference is preserved.

References

1. D. Driankov, H. Hellendoorn, M. Reinfrank, An introduction to fuzzy control (2nd ed.), Springer-Verlag, London, UK, UK, 1996.
2. D. Tikk, L. T. Kóczy, T. D. Gedeon, A survey on universal approximation and its limits in soft computing techniques, Int. J. Approx. Reasoning 33 (2) (2003) 185–202.
3. D. Dubois, H. Prade, The generalized modus ponens under sup-min composition: a theoretical study, in: Approximate Reasoning in Expert Systems, Elsevier/North-Holland.
4. S. Roychowdhury, W. Pedrycz, A survey of defuzzification strategies, International Journal of Intelligent Systems 16 (6) (2001) 679–695.
5. E. P. Klement, R. Mesiar, E. Pap, Triangular Norms, Vol. 8 of Trends in Logic, Kluwer Academic Publishers, Dordrecht, 2000.
6. M. Baczyński, B. Jayaram, Fuzzy Implications, Vol. 231 of Studies in Fuzziness and Soft Computing, Springer, 2008.
7. W. Bandler, L. J. Kohout, Semantics of implication operators and fuzzy relational products, International Journal of Man-Machine Studies 12 (1) (1980) 89 – 116.
8. L. A. Zadeh, Outline of a new approach to the analysis of complex systems and decision processes, IEEE Transactions on Systems, Man and Cybernetics.
9. P. Magrez, P. Smets, Fuzzy modus ponens: A new model suitable for applications in knowledge-based systems, Int. J. Intell. Syst. 4 (2) (1989) 181–200. doi:10.1002/int.4550040205.
10. V. Cross, T. Sudkamp, Fuzzy implication and compatibility modification, in: Fuzzy Systems, 1993., Second IEEE International Conference on, 1993, pp. 219–224 vol.1.
11. I. Turksen, Z. Zhong, An approximate analogical reasoning approach based on similarity measures, IEEE Transactions on Systems, Man and Cybernetics.
12. N. N. Morsi, A. A. Fahmy, On generalized modus ponens with multiple rules and a residuated implication, Fuzzy Sets and Systems 129 (2) (2002) 267 – 274.
13. S.-M. Chen, A new approach to handling fuzzy decision-making problems, in: Multiple-Valued Logic, 1988., Proceedings of the Eighteenth International Symposium on, 1988.
14. L. A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning-III, Information Sciences 9 (1975) 43–80.
15. Y.-M. Li, Z.-K. Shi, Z.-H. Li, Approximation theory of fuzzy systems based upon genuine many-valued implications: SISO cases, Fuzzy Sets Syst. 130 (2002) 147–157.
16. Y.-M. Li, Z.-K. Shi, Z.-H. Li, Approximation theory of fuzzy systems based upon genuine many-valued implications: MIMO cases, Fuzzy Sets Syst. 130 (2002) 159–174.
17. M. Štěpnička, B. D. Baets, Monotonicity of implicative fuzzy models, in: FUZZ-IEEE 2010, IEEE International Conference on Fuzzy Systems, Barcelona, Spain, 18-23 July, 2010, Proceedings, 2010, pp. 1–7.