

Paper:

# Fuzzy Rule Interpolation and Extrapolation Techniques: Criteria and Evaluation Guidelines

Domonkos Tikk<sup>\*1</sup>, Zsolt Csaba Johanyák<sup>\*2</sup>, Szilveszter Kovács<sup>\*3</sup>, and Kok Wai Wong<sup>\*4</sup>

<sup>\*1</sup>Dep. of Telecommunications and Media Informatics, Budapest University of Technology and Economics

Magyar tudósok krt. 2, H-1117 Budapest, Hungary

E-mail: tikk@tmit.bme.hu

<sup>\*2</sup>Institute of Information Technology, Kecskemét College

Izsáki út 10, H-6000 Kecskemét, Hungary

E-mail: johanyak.csaba@gamf.kefo.hu

<sup>\*3</sup>Department of Information Technology, University of Miskolc

H-3515 Miskolc-Egyetemváros, Hungary

E-mail: szkovacs@iit.uni-miskolc.hu

<sup>\*4</sup>School of Information Technology, Murdoch University

South St., Murdoch, Western Australia 6150, Australia

E-mail: k.wong@murdoch.edu.au

[Received December 22, 2010; accepted March 18, 2011]

**This paper comprehensively analyzes Fuzzy Rule Interpolation and extrapolation Techniques (FRITs). Because extrapolation techniques are usually extensions of fuzzy rule interpolation, we treat them both as approximation techniques designed to be applied where sparse or incomplete fuzzy rule bases are used, i.e., when classical inference fails. FRITs have been investigated in the literature from aspects such as applicability to control problems, usefulness regarding complexity reduction and logic. Our objectives are to create an overall FRIT standard with a general set of criteria and to set a framework for guiding their classification and comparison. This paper is our initial investigation of FRITs. We plan to analyze details in later papers on how individual techniques satisfy the groups of criteria we propose. For analysis, MATLAB FRI Toolbox provides an easy-to-use testbed, as shown in experiments.**

**Keywords:** fuzzy rule interpolation (FRI), fuzzy rule extrapolation (FRE), criteria of FRITs, evaluation guidelines of FRI methods

## 1. Introduction

Fuzzy-Rule-Based Systems (FRBSs) have been applied in applications such as control engineering, expert systems, pattern recognition, operation research, and decision support systems [1–7]. FRBS output is generated by an inference mechanism based on a knowledge base of IF–THEN rules. Originally based on Zadeh’s initial linguistic variable concept, rule bases were assembled using expert knowledge. This was replaced from the 90s – particularly in engineering applications – by automatic rule base construction extracting rules from numerical sample

data. Classical inference – Zadeh’s CRI [8], Mamdani-Larsen [9, 10], Takagi-Sugeno [11, 12], etc. – determine output by rule matching, i.e., matching observed input to rule premises and calculating conclusions as weighted combinations of rule consequents with nonzero matching in which weights depend on the degree of matching.

Due to incomplete knowledge, rule-base construction may produce sparse or incomplete rule bases that may be due to missing or insufficient expertise or available sample data for covering all possible input configurations. Such rule bases thus do not completely (or sufficiently) cover the range of possible input, as defined in Section 2. In sparse or incomplete rule bases, classical fuzzy inference approaches do not always generate meaningful output because actual input is not guaranteed to fire any rule. Such situations can, however, be treated by approximative approaches. The approximative approaches most commonly applied are Fuzzy Rule Interpolation (FRI) and Fuzzy Rule Extrapolation (FRE). FRE is usually an extension of FRI. In the literature, FRI attracted more interest and FRE was mostly neglected [13–17], even though the latter significantly broadens FRI applicability. When we refer here to FRITs, we mean an approximative technique dealing with sparse rule bases including both FRI and FRE; when approximation is concerned, we explicitly specify either interpolation or extrapolation.

The application of FRITs can also be justified on other grounds. The main motivation of the first proposed FRIT [18] originated in fuzzy system complexity [19] – the “curse of dimensionality” problem, i.e., the rule base size and inference algorithm complexity grow exponentially with input space dimensions. This issue is settled, in part, by omitting redundant or insignificant rules. The original rule base’s completeness cannot, however, be guaranteed after rule base reduction, i.e., the reduced rule base is often sparse.

Despite the numerous FRITs proposed since 1991,

a general set of criteria for FRIT consisting of both mathematical- and application-originated requirements does not exist. Our goal here is to correct this deficiency.

The several initiatives in the literature to set up FRIT criteria are divided into two groups – approach-oriented and summarization-oriented. The approach-oriented papers focus on a new FRI motivated by critiques of other methods or new analytical aspects, e.g., [15, 20, 21], using conditions in FRIT analysis that are, in our opinion, motivated to justify the proposed approach or aspect, i.e., are not comprehensive enough. Summarization-oriented papers include brief summaries mostly focusing on selected FRI aspects, e.g., [22–25], criticizable on the same ground. [21], for example, analyzes and compares proposed MACI’s general applicability, complexity, approximative power, and fuzziness of conclusion. Jenei [15] axiomatically treated FRITs by setting eight conditions on rule interpolation/extrapolation, focusing on applicability, consistency, functionality, and logic. Baranyi et al. [20] investigated the general methodology for applicability, consistency, and shape-preservation – all of which appear in the second, which are also summarizations.

As this paper’s main contribution, we define the set of criteria and properties an ideal FRIT must meet. These properties may serve as a standard and be used to guide FRIT classification and comparison. To facilitate comparison and evaluation, we recently created the MATLAB FRI Toolbox [26], which implements several important FRITs and can be extended by other contributors.

This paper is organized as follows. Section 2 defines the notion of a sparse rule base and characterizes FRIT applicability. Section 3 gives FRIT criteria and properties investigated in Section 4 by two examples using our MATLAB FRI Toolbox. Section 5 presents conclusions.

## 2. Sparse Rule Bases

We assume readers to be familiar with basic fuzzy set theory, e.g., [27].

### 2.1. Fuzzy Rule Base

The knowledge base for approaching fuzzy paradigm reasoning consists of fuzzy IF–THEN rules. The fuzzy rule ensemble partially maps areas between regions of input and output space, formalized in the model of available but possibly uncertain knowledge. Uncertainty is encoded in fuzzy set membership functions describing input and output space.

Let  $X_j$  ( $j = 1, \dots, n$ ) be input dimensions and  $Y$  output space, denoting the Cartesian product of input dimensions by  $X = \times_{j=1}^n X_j$ . A fuzzy IF–THEN rule is given as

$$R_i : \text{if } A_{i1} \wedge A_{i2} \wedge \dots \wedge A_{in} \text{ then } B_i \quad \dots \quad (1)$$

where antecedents  $A_{ij} \in \mathcal{F}(X_j)$ , consequents  $B_i \in \mathcal{F}(Y)$ , and  $\mathcal{F}(Z)$  denote the entirety of all fuzzy subsets of  $Z$ . We denote the ( $n$ -dimensional) Cartesian product of antecedents  $A_{ij}$ , ( $j = 1, \dots, n$ ) of rule  $R_i$  by  $A_{(i)}$ . Based on

the concept dominant in literature, we limit our investigation to rules of form Eq. (1), i.e., when fuzzy sets in the antecedent are connected by a conjunction. If other logical connectives or unary operators, such as disjunction and negation, are allowed, conflicts may occur in the rule base [28], which must then be resolved by inference mechanisms.

Multiple output rule bases – SIMO or MIMO – can be decomposed into single output rule bases thus, without loss of generality, only MISO rule bases are investigated.

We require that  $X_j$ ,  $j = 1, \dots, n$  and  $Y$  be bounded and gradual [19, 29]. This guarantees a total ordering for each by which a partial ordering can be introduced among  $\mathcal{F}(X_j)$  and  $\mathcal{F}(Y)$  elements. In practice,  $X$  and  $Y$  are typically compact subsets of  $\mathbb{R}^n$  and  $\mathbb{R}$ .

### 2.2. Rule Base Coverage

Let us characterize the applicability of rule-matching-based fuzzy inference mechanisms regarding rule base properties. We first define the activation degree of a particular rule and rule base, using the degree of matching.

**Definition 1.** *The activation degree (degree of matching) of rule  $R_i$  for  $n$ -dimensional observation  $A^*$  is*

$$\omega_{n,t}(R_i) = h(t(A_{i1}, A_1^*), \dots, t(A_{in}, A_n^*)), \quad \dots \quad (2)$$

where  $t$  denotes an arbitrary  $t$ -norm, and  $h$  can be an arbitrary  $t$ -conorm or aggregation operator.

In Eq. (2) the most typical  $h$  is the min function. Rules with nonzero activation degree are called activating, firing, or matching rules.

**Definition 2.** *The activation degree (degree of matching) of rule base  $\mathbf{R} = \{R_i | i = 1, \dots, r\}$  for observation  $A^*$  is equal to the highest rule activation degree in the rule base*

$$\omega(\mathbf{R}) = \max_{i=1}^r \omega(R_i). \quad \dots \quad (3)$$

Rule-matching-based fuzzy inference is applicable if it generates a conclusion. This condition holds if the rule base activation degree is nonzero for arbitrary observation. Practically speaking, it is reasonable to specify an  $\varepsilon > 0$  threshold for the minimal rule base activation degree.  $\varepsilon$  can also be interpreted as a prescribed minimal confidence value of the conclusion, so the applicability of such methods depend on the rule base, or more precisely, how the entirety of rule antecedents covers  $X$ , characterized by the next definition.

**Definition 3.** *Let  $\mathbf{R} = \{R_i | i = 1, \dots, r\}$  be a rule base with rules of form Eq. (1). If*

$$\forall A^* \in \mathcal{F}(x) : \omega(\mathbf{R}) \geq \varepsilon > 0 \quad \dots \quad (4)$$

*then  $\mathbf{R}$  forms an  $\varepsilon$ -cover of  $X$ .*

The  $\varepsilon$ -coverage of a rule base guarantees that the rule-base activation degree is at least  $\varepsilon$  for arbitrary observation, hence rule-matching-based models apply. We call such rule bases  $\varepsilon$ -dense.