



Approximate reasoning with fuzzy rule interpolation: background and recent advances

Fangyi Li^{1,2} · Changjing Shang² · Ying Li³ · Jing Yang^{2,3} · Qiang Shen² 

Published online: 5 June 2021
© The Author(s) 2021

Abstract

Approximate reasoning systems facilitate fuzzy inference through activating fuzzy if-then rules in which attribute values are imprecisely described. Fuzzy rule interpolation (FRI) supports such reasoning with sparse rule bases where certain observations may not match any existing fuzzy rules, through manipulation of rules that bear similarity with an unmatched observation. This differs from classical rule-based inference that requires direct pattern matching between observations and the given rules. FRI techniques have been continuously investigated for decades, resulting in various types of approach. Traditionally, it is typically assumed that all antecedent attributes in the rules are of equal significance in deriving the consequents. Recent studies have shown significant interest in developing enhanced FRI mechanisms where the rule antecedent attributes are associated with relative weights, signifying their different importance levels in influencing the generation of the conclusion, thereby improving the interpolation performance. This survey presents a systematic review of both traditional and recently developed FRI methodologies, categorised accordingly into two major groups: FRI with non-weighted rules and FRI with weighted rules. It introduces, and analyses, a range of commonly used representatives chosen from each of the two categories, offering a comprehensive tutorial for this important soft computing approach to rule-based inference. A comparative analysis of different FRI techniques is provided both within each category and between the two, highlighting the main strengths and limitations while applying such FRI mechanisms to different problems. Furthermore, commonly adopted criteria for FRI algorithm evaluation are outlined, and recent developments on weighted FRI methods are presented in a unified pseudo-code form, easing their understanding and facilitating their comparisons.

Keywords Approximate reasoning · Fuzzy rule interpolation · Weighted fuzzy interpolative reasoning · Sparse rule bases

✉ Qiang Shen
qq@aber.ac.uk

Extended author information available on the last page of the article

1 Introduction

Fuzzy set theory (Zadeh 1965) and its extension, fuzzy logic, have been successfully applied for many real-world problems (e.g., Ross 2005; Terano et al. 2014; Zimmermann 2011). In particular, fuzzy expert systems are popular due to their ability to exploit the tolerance for imprecision, partial truth and approximations, in an effect to achieve close resemblance with human activity and reasoning intuition. Many of which have been developed using the idea of approximate reasoning (also known as linguistic reasoning), reflecting the manner of human cogitation and leading to new, more human interpretable, intelligent systems.

In general, an approximate reasoning system can be formalized as a fuzzy *if-then* rule-based inference mechanism that derives a conclusion given an input observation. It consists of linguistic variables, fuzzy rules and a fuzzy inference method. Linguistic variables facilitate the interpretation of linguistic expressions in terms of fuzzy quantities of certain underlying mathematical semantics. Fuzzy inference rules are a set of rules that associate input and output variables of a given physical system or other phenomenon in determining their relationships, either learned from historical data or directly acquired from domain experts, or a mixture of both. Based on such rules, a fuzzy inference mechanism is encoded to implement the process of approximate reasoning, through manipulation among the fuzzy inference rules in response to any new input data. As such, fuzzy rule bases are the essential component of any approximate reasoning model, storing knowledge required to inference and determining what computational techniques to use.

Another main component of an approximate reasoning system is the mechanism that computes the output given an input and the rule base. A variety of potentially useful methods exist in the literature. Many have been seen to implement generalised *modus ponens*, mostly by following the basic idea of Compositional Rule of Inference (CRI, Zadeh 1973). The law of CRI has been widely and successfully applied. For example, the famous Mamdani's fuzzy logic controller (Mamdani and Assilian 1999) was implemented this way. Nevertheless, CRI can only derive reasonable and accurate outcomes in its full potential while working with dense, or complete, rule bases where rules available cover the entire problem space. This implies that any unknown observation must match at least one fuzzy rule in the rule base for the mechanism to work.

In many circumstances, a dense rule base cannot be realistically obtained, but only an incomplete rule base instead. A number of reasons may lead to such incomplete rule bases, with the most common ones being (Baranyi et al. 1999; Tikk and Baranyi 2000): (1) To utilise incomplete knowledge about the modelled problem, regardless of the means for the construction of the rule base, be it from human expertise or machine learning techniques; and (2) To reduce the number of rules in a rule base and hence, the complexity of the resultant fuzzy system.

Unfortunately, CRI is unable to draw a conclusion when a rule base is not dense but sparse. Sparse, or incomplete, rule bases considered here are not referring to the quantity of the rules in a given rule base, but to the coverage of the problem domain by the antecedents of the rules regarding the universe of discourse. That is, an input observation may have no overlap with any of the rules available and hence, no rule may be executed to derive the required consequent by directly applying CRI.

1.1 Research context

Resolving real-world problems frequently involves the use of such sparse rule bases, since a dense rule base may be impracticable in a multidimensional environment where the number of rules increases exponentially as the input variables and the fuzzy linguistic labels associated with each variable increase. It is therefore, desirable to develop more advanced inference mechanism to work with incomplete rule bases. Fuzzy rule interpolation (FRI) facilitates approximate reasoning in fuzzy rule-based systems when only sparse knowledge is available (Kóczy and Hirota 1993a, b). It addresses the key limitation of CRI that requires a dense fuzzy rule base which fully covers the entire problem domain. With a different underlying approach from CRI and many of its derivatives, FRI reasons through manipulation of rules that bear similarity with an unmatched observation without having to resort to direct pattern matching. This makes a significant breakthrough in fuzzy rule-based inference for situations where only an incomplete rule base is available. It works with sparse knowledge, attempting to reduce, if not to completely remove, the restriction of CRI for cases where no conclusion may be derived due to no rules matching a new observation. This offers an alternative way to infer an approximately interpolated outcome, accomplishing the so-called fuzzy interpolative reasoning.

FRI essentially makes two contributions to the development of fuzzy rule-based systems. It not only facilitates the assistance of reasoning on sparse rule bases (Burkhardt and Bonissone 1992), but also offers the potential for a reverse application where the rule base may be so dense that model simplification is required. That is, FRI can be utilised to simplify the complexity of fuzzy rule bases through say, a procedure of iteratively replacing two existing rules with an interpolated one (Koczy and Hirota 1997), thereby eliminating those fuzzy rules which may be approximated from their neighbouring ones. Nonetheless, this is not the motif of the great majority of FRI techniques developed in the literature and hence, is beyond the scope of this review. This paper focuses on the former issue that is performing inference with a sparse rule base. From this viewpoint, the goal of FRI is not to produce an interpolated rule through interpolative reasoning, but to compute an interpolated consequent that corresponds to the input observation. In so doing, FRI achieves the inference task with respect to the observations that originally have no conclusions to be drawn due to the sparseness of the fuzzy rule base.

As an inference mechanism, FRI starts to reach its goal from the selection of the nearest neighbouring (aka. the closest) rules in the sparse fuzzy rule base with regards the given unmatched observation. Such chosen rules form the basis for conducting fuzzy interpolation. Two major types have been seen in the literature to implement fuzzy interpolative reasoning: (1) the α -cut based interpolations and (2) the intermediate rule based interpolations. This grouping depends upon whether the computation of the interpolated result is accomplished through a process of construction and transformation of an intermediate rule first. As such, FRI methods may also be organised in two groups, respectively termed as non-transformation based and transformation based FRI (Chen and Adam 2018). The seminal work for fuzzy interpolative reasoning, as of the techniques reported in Kóczy and Hirota (1993a, b) and their extensions, form the most typical non-transformation based FRI. For those relying on transforming intermediate rules, a family of scale and move transformation-based FRI (termed as T-FRI), such as those given in Huang and Shen (2006, 2008), Jin et al. (2014), Yang et al. (2017) and Naik et al. (2017b), have been popularly studied and widely applied despite their relative recency. This survey will review the general FRI methodologies as its first main topic.

In resolving practical real-world problems, multidimensional input variables are a common issue. Fortunately, many FRI methods exist in the literature that are capable of dealing with interpolation, by the use of fuzzy rules that involve multiple antecedent variables. Nonetheless, there is a common problem existing in these FRI approaches, where the antecedent attributes within the rules are presumed to be of equal significance for interpolation. Thus, inaccurate and even incorrect interpolated outcomes may result since different domain attributes may generally make different contributions to the decision making process.

Recently however, a number of methods have been proposed for FRI working on multiple antecedents associated with different weights (e.g., Chen and Chang 2011b; Chen and Chen 2016; Chen et al. 2009; Cheng et al. 2015; Diao et al. 2014). They are making significant contributions to enhancing the inference accuracy by moving away from the assumption of equal antecedent weight when dealing with sparse rule bases. In order to achieve the goal of weighted FRI, two closely related questions are the main issues shared within the weighted FRI methods, which are (1) how the weights are generated; and (2) whether and how the weights are integrated within the underlying, non-weighted FRI. A variety of weighted FRI approaches have been developed to implement approaches that attempt to address these two concerns. To generate the weights of different rule antecedent variables, it can be fulfilled either by assigning the predefined weights by domain experts or by running automatic weight learning algorithms. The latter may be obviously preferred as predefining weights will require human intervention and hence, will adversely reduce the flexibility and automation level of the resulting fuzzy systems. Regarding the second question, the weights are integrated within the original FRI in rather different way when it is implemented in different weighted FRI approaches. These two issues can be dealt with independently, or be implemented in a form of the “wrapper” approach that mixes up FRI-based inference and learning weights from data. Discussions about the techniques for weighted FRI form the second focus of this survey.

The efficacy of the inference mechanism introduced by an FRI method may be reflected or revealed through their utilisation in resolving real-world application problems. As with many practical applications of classic fuzzy reasoning tools, FRI has reinforced the power of systems control, including successful examples: for the simulation of automated guided vehicles (Kovács and Kóczy 1999), for surveillance navigation control of mobile robots (Vincze and Kovács 2008), and for general behaviour-based control (Kovács and Kóczy 2004). The work on dynamic FRI (Naik et al. 2017b) offers significant opportunities for facilitating selection, combination and generalisation of informative, frequently used interpolated rules for enriching existing rule base while performing interpolation. It provides promising solutions to cyber-security problems, including: network security analysis, intelligent intrusion detection (Naik et al. 2017b) and firewall reinforcement (especially for Microsoft Windows Firewall) (Naik et al. 2017a). FRI also finds impressive results in performing practical pattern recognition tasks, examples include: classic classification and prediction problems (Li et al. 2018b, 2020a; Chen and Chen 2016) using weighted FRI techniques; computer vision and image super resolution (Yang et al. 2019); and disease diagnosis in general and mammographic mass risk analysis (Li et al. 2019) and colorectal polyp detection (Nagy et al. 2018) in particular. Further applications of FRI are found in function approximation (Wong and Gedeon 2000; Berecz 2009) and student academic performance evaluation (Johanyák 2010).

1.2 Motivational objectives

This paper aims to provide a comprehensive review of FRI techniques that enable approximate reasoning in the context of sparse rule bases, covering both the conventional FRI methods and the recent advances (particularly in weighted FRI mechanisms). To be more specific, the objectives of this survey are three-fold:

- To offer an up-to-date tutorial of the key developments regarding fuzzy rule-based inference which are tailored to situations where only incomplete domain knowledge is available;
- To provide a systematic comparison between different approaches, so that readers can have an informed choice of what may be the potentially suitable FRI technique(s) to apply given their specific domain problems; and
- To promote the advanced weighted FRI mechanisms to inform the readers about the benefits of using such most recent developed algorithms, which ensure not only the effectiveness of approximate reasoning but also the efficiency (as each time, only two nearest neighbouring rules are required to perform FRI).

1.3 Main contributions

By achieving the above objectives, this survey offers the following major contributions to the relevant literature: (C1) The basic idea of FRI techniques are summarised and the main FRI approaches categorised, with the properties of the representatives of each category discussed, offering a comprehensive tutorial of the FRI literature. (C2) The methodologies of weighted FRI are introduced systematically, in terms of both the mechanisms for weight generation and the techniques for weight integration within the underlying otherwise unweighted FRI methods, with all reviewed algorithms concisely described using an unified pseudocode format to ease understanding. (C3) The comparison between different FRI techniques are provided with respect to a series of commonly used criteria, and this is shown not only within each of the two main categories of approaches: unweighted or weighted, but also between the two categories themselves, highlighting the advantages of running a weighted FRI method over its unweighted counterpart.

1.4 Paper structure

The remainder of this paper is organised as follows. Section 2 first illustrates the basic notations utilised throughout the survey. Sections 3 and 4 review the techniques of unweighted FRI and weighted FRI, including a comparative analysis within each category. Section 4 also includes a qualitative comparison between the two major FRI categories themselves. Section 5 finally concludes this survey and points out interesting further work in this exciting field. To aid in better digesting and henceforth, in facilitating drawing informed conclusions from different parts of this paper, the structural organisation of this survey is illustrated Fig. 1, including an outline of the key content and concepts contained within each section, in relation to the major contributions as identified above.

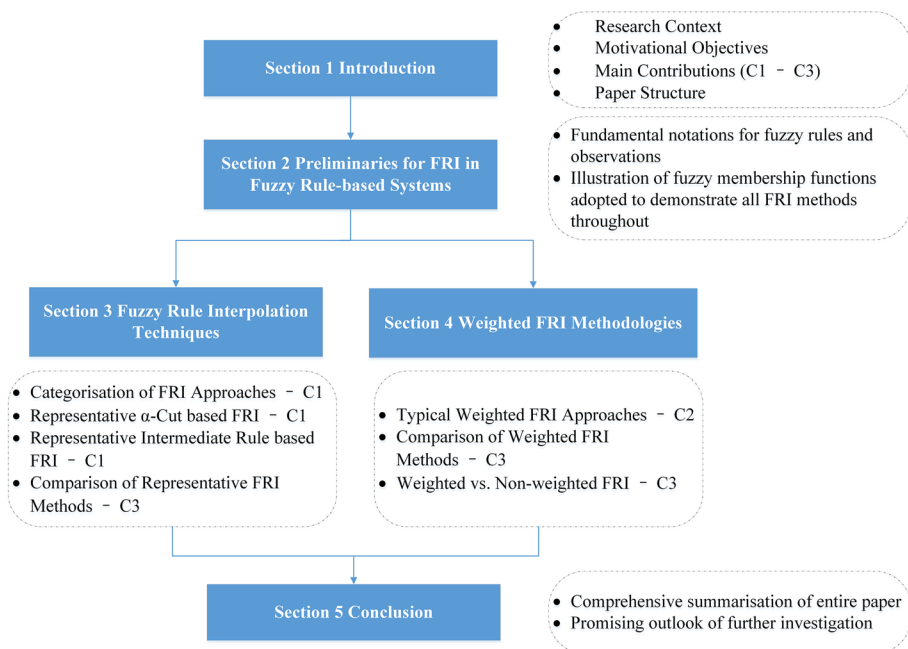


Fig. 1 Organisation of survey structure

2 Preliminaries for FRI in fuzzy rule-based systems

This is a significant undertaking to survey a large body of literature about fuzzy rule interpolation, concerning both recent and current developments over a diversity of different approaches. Thus, the typical structure used for presenting literature reviews in the field of computational intelligence is adopted in this work. In particular, preliminary theoretical foundations that are shared by the different approaches to be reviewed are provided herein, in an effort to help reduce potential repetitions in the subsequent main body of the survey.

This section first explains the fundamental notations adopted to express fuzzy rules and observations, both being the key components in any FRI system that performs fuzzy reasoning with incomplete knowledge. This is followed by an illustration of fuzzy membership functions (MFs), which are used to describe the antecedent and consequent parts of a fuzzy rule or a given observation. Then, the main categorisation of MFs is outlined, with triangular MFs being highlighted and described in slightly more detail.

In general, a fuzzy rule-based system, where FRI works within, has as its key component a set of *if-then* rules, each of which takes fuzzy or crisp terms that represent specifications of the input variables and associates these with the output of a certain problem description. In general, a rule may involve multiple output attributes as well as multiple input variables, but a multiple output rule can always be equivalently expressed by several single output rules. Without losing generality, only rules which have a single output class are considered in this work.

Formally, a typical fuzzy rule model essentially contains two key elements $\langle R, Y \rangle$ in describing a given problem: A non-empty finite set of domain attributes $Y = A \cup \{z\}$,

where $A = \{a_j | j = 1, 2, \dots, m\}$ represents the set of input antecedent attributes and z stands for the consequent, and a non-empty finite set of fuzzy rules $R = \{r^1, r^2, \dots, r^N\}$. In many conventional fuzzy rule-based systems, including systems implemented with FRI techniques, a given rule $r^i \in R$ and an observation o^* are often expressed generally as follows:

$$\begin{aligned} r^i &: \text{if } a_1 \text{ is } A_1^i \text{ and } a_2 \text{ is } A_2^i \text{ and } \dots \text{ and } a_m \text{ is } A_m^i, \text{ then } z \text{ is } B^i \\ o^* &: a_1 \text{ is } A_1^* \text{ and } a_2 \text{ is } A_2^* \text{ and } \dots \text{ and } a_m \text{ is } A_m^* \end{aligned} \quad (1)$$

where A_j^i and A_j^* denote the fuzzy set values taken by the antecedent attribute a_j in r^i and o^* , respectively; and B^i represents the fuzzy set value of the consequent attribute z in r^i .

In the above rule representation, the logical conjunctive term *and* between each conjunct pair of the propositions that an antecedent attribute takes a fuzzy value indicates that the fuzzy values taken by the two attributes form a compound fuzzy value. The compound value is computed as the result of applying a t-norm operator (e.g., the popular operation *min*) to the two conjunct values (Bede 2013). Thus, the “if-part” of a fuzzy rule can be interpreted in the following, where the operator \wedge may be simply implemented with *min*:

$$A_1^i(a_1) \wedge A_2^i(a_2) \wedge \dots \wedge A_m^i(a_m) \quad (2)$$

This interpretation equally applies to the interpretation of the term *and* in an observation. However, a given observation is deliberately denoted differently from the rule in the above, where the value of any attribute involved is labelled with an asterisk. This is to explicitly differentiate it from the antecedent part of a rule, as unlike any rule the observation does not have the “then-part”. It is this “conclusion” part that is unknown and that is to be inferred, using an FRI technique if there is no rule to match the observation. Throughout this paper, unless otherwise stated, any notion that is attached with an asterisk sign implies that it is part of an observation.

Fuzzy values of both the rule antecedents and the consequent are in general represented by fuzzy sets. The concept of fuzzy sets was introduced by Zadeh (1965). Informally, the definition of a fuzzy set given by Zadeh can be stated as follows: A fuzzy set is a class with a continuum of membership grades. Thus, a fuzzy set A in a universe of discourse X is characterised by a membership function (MF) A which associates each element $x \in X$ with a real number $A(x) \in [0, 1]$. This is interpreted such that $A(x)$ is the membership grade of x belonging to the fuzzy set A (Bede 2013).

As can be seen from the above, the MF $A : X \rightarrow [0, 1]$ distinguishes the fuzzy sets from the classical boolean sets. Unlike a classical set with clear boundaries, i.e., $x \in A$ or $x \notin A$, which excludes any other possibility, the property of the membership function enables fuzzy sets to model partial degrees to which a variable or attribute is deemed to take a certain underlying real or categorical value. Such fuzzy sets are often assigned with linguistic terms to help capture and reflect human interpretation of imprecise measurements or descriptions.

Particularly, when the universe of discourse X consists of the real line \mathbb{R} , any type of continuous functions can be used as an MF, provided that a set of parameters is given to specify the appropriate meanings of the MF. In this case, it is impractical to list all the pairs defining an MF, even if imposing the constraint that all MFs are convex in topology to ease the expression of common sense interpretation of belongingness. Fortunately, only a small number of types of MF that are typically used in practice. Basically, there are two main categories in terms of their properties: smoothness and linearity, which are: (1) polygonal (piecewise linear) fuzzy sets, including triangular shaped, trapezoidal shaped, hexagonal

shaped MF, etc., and (2) nonlinear fuzzy sets, typically including Gaussian, Generalised bell-shaped, and Sigmoid MFs.

Polygonal fuzzy sets are generally represented by their characteristic points (CPs) in ascending order [which are defined mathematically as the odd points of the membership function (Huang and Shen 2008)], and nonlinear ones by the defining parameters that are used to specify each nonlinear function. The choice of different MFs relies on the specific requirements of a given application. Amongst the family of all possible functions, triangular MFs and trapezoidal MFs have been used most extensively, especially for real-time implementations, thanks to their simple representation and computational efficiency. In the literature, generally speaking, different MFs have been exploited to implement the proposed FRI methods. Nonetheless, procedures employing triangular MFs and/or trapezoidal MFs may be seen as specific cases of those which utilise more complex polygonal fuzzy sets. It is difficult to have a generic closed form representation that unifies all FRI processes as they are dependent upon the MFs used (Yang and Shen 2013). In this work, for illustrative and demonstrative consistency and simplicity, as well as for their popularity in the literature, triangular MFs are employed to describe and contrast all FRI methods.

As shown in Fig. 2, a normal and convex triangular fuzzy set A is illustrated with its three ascending-ordered CPs, i.e., (a_1, a_2, a_3) , where the first and third CP stand for the two extreme points of the support with a membership value of 0 and the middle one stands for the normal point of the fuzzy set with a membership of 1. For a fuzzy rule base consisting of rules in the form as per Eq. (1), the triangular fuzzy values A_j^i , A_j^* , B^i , and the consequent B^* to be computed by an FRI process ($i = 1, 2, \dots, N, j = 1, 2, \dots, m$) are therefore, represented by their corresponding CPs: $(a_{j1}^i, a_{j2}^i, a_{j3}^i)$, $(a_{j1}^*, a_{j2}^*, a_{j3}^*)$, (b_1^i, b_2^i, b_3^i) , and (b_1^*, b_2^*, b_3^*) , respectively.

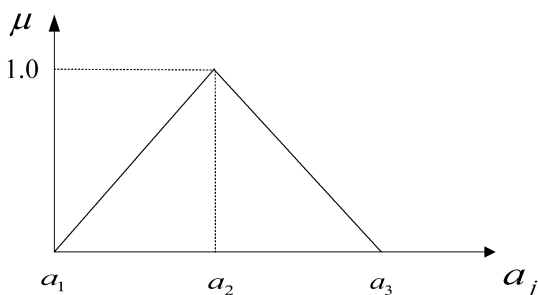
3 Fuzzy rule interpolation techniques

This section categorises and details the representatives of classical FRI methods, with typical pros and cons of different approaches discussed.

3.1 Categorisation of FRI approaches

In the literature, various FRI approaches have been proposed following the seminal work of Kóczy and Hirota (1993a, b), to perform fuzzy interpolative reasoning. In general, the existing methodologies can be grouped into two categories:

Fig. 2 Normal and convex triangular membership function



1. α -cut/Non-transformation based FRI, see Table 1 for a summary with Table 2 listing further developments belonging to this category.
2. Intermediate rule/Transformation based FRI, see Table 3 for a summary with Table 4 listing a particular family of scale and move transformation based FRI (denoted as T-FRI hereafter) which are the most popular in the recent literature.

Table 1 α -cut (non-transformation) based FRI methods

Methods	Characteristics
Huang et al. (2004), Kóczy and Hirota (1993a, b) and Ughetto et al. (2000)	FRI with only two fuzzy rules
Chang et al. (2008), Chen and Chen (2016), Chen et al. (2013a, 2015), Chen and Lee (2011), Cheng et al. (2015, 2016); Kovács (2006), Yang and Shen (2013)	FRI with multiple fuzzy rules
Chen and Chen (2016), Chen et al. (2013a) and Cheng et al. (2015)	FRI with fuzzy rules weighted
Chen and Lee (2011)	FRI with interval type-2 fuzzy sets (Mendel et al. 2006)
Chen et al. (2015)	FRI with rough-fuzzy sets
Chen and Adam (2017) and Cheng et al. (2016)	FRI with adaptivity
Hsiao et al. (1998)	Exploiting slopes of fuzzy sets to obtain valid conclusions

Table 2 Family of KH FRI

Methods	Characteristics
Kóczy and Hirota (1997, 1991, 1993a, b, c), Kóczy et al. (2000a, b)	Foundational linear KH FRI based on computation of α -cut levels
Vass et al. (1992)	Extended KH FRI with reduction of invalid conclu- sions
Baranyi et al. (1999), Tikk (1999), Tikk and Bara- nyi (2000) and Yam et al. (1999)	Modified α -cut based method based on coordinate modification
Tikk et al. (1997, 1999, 2002)	Stabilised (general) KH interpolation
Wong et al. (2000, 2005)	Modified α -cut based multidimensional scheme

Table 3 Intermediate rule (transformation) based FRI methods

Methods	Characteristics
Wu et al. (1996)	Using similarity transfers to guarantee valid inter- polation
Baranyi et al. (1995, 1996a, b, 1998, 2004), Baranyi and Kóczy (1996a, b)	Adopting generalised concept for interpolation and extrapolation
Kawaguchi and Miyakoshi (2000a, b, 2001) and Kawaguchi et al. (1997)	Performing B-spline based interpolation
Chen et al. (2016), Chen and Shen (2017), Huang and Shen (2006, 2008), Jin et al. (2014, 2020), Li et al. (2020a, 2018b), Naik et al. (2017b), Shen and Yang (2011), Yang et al. (2017), Yang and Shen (2011)	Running FRI with scale and move transformation (T-FRI)
Das et al. (2019)	Using fuzzy geometry for linear FRI

Table 4 Family of scale and move transformation based FRI (T-FRI)

Methods	Characteristics
Huang and Shen (2006)	Foundational T-FRI working with two given neighbouring rules involving multiple antecedent variables in various fuzzy membership functions (e.g., complex polygon, Gaussian or bell-shaped)
Huang and Shen (2008)	Extended T-FRI facilitating both interpolation and extrapolation involving multiple fuzzy rules, with each rule consisting of multiple antecedents
Jin et al. (2014, 2019)	Backward T-FRI allowing missing antecedent values directly related to the consequent to be interpolated from known antecedents and consequent, supporting backward interpolation and extrapolation involving multiple multi-antecedent fuzzy rules
Yang et al. (2017), Yang and Shen (2011)	Adaptive T-FRI being capable of restoring system consistency once contradictory results reached during interpolation
Chen et al. (2016)	Rough-fuzzy T-FRI allowing representation, handling and utilisation of different levels of uncertainty in knowledge
Chen and Shen (2017)	Extended T-FRI with interval type-2 fuzzy sets
Naik et al. (2017b)	Dynamic T-FRI facilitating selection, combination, and generalisation of informative, frequently used interpolated rules for enriching existing rule base while performing interpolation
Li et al. (2017, 2018a, b, 2020a)	Weighted T-FRI allowing attribute weights to be generated given a sparse rule base, being integrated within non-weighted FRI for classification and prediction
Chen et al. (2020)	Extended T-FRI being implemented with Takagi Sugeno Kang (TSK) fuzzy models

This categorisation is made depending upon whether processes to construct and then, to utilise a so-called intermediate fuzzy rule are involved in order to derive an interpolated result.

The α -cut based FRI approaches, also known as non-transformation based methods, directly interpolate the results based on the computation of each α -cut level given at least two fuzzy rules adjacent to an unmatched observation. Considerable work has been reported on this type of approach at the early stage of the investigation of fuzzy interpolative reasoning. In particular, the very first proposed, termed the KH method after the name of its inventors (Kóczy and Hirota 1993a, b), is the most typical α -cut based algorithm for FRI. As indicated earlier, Table 1 also summarises several other alternative α -cut based methods from different perspectives.

For the group of transformation-based approaches, they work by first computing an intermediate rule. The required consequent to an unknown observation is obtained through a two-step procedure by manipulating selected neighbouring rules to the observation. An intermediate rule is artificially constructed such that its antecedent is as “close” (given a certain distance metric, often the Euclidean one) to the observation as possible. An intermediate consequent is computed from the constructed rule antecedent. Observing that there may still exist a difference between the antecedent of the intermediate rule and the observation, the second step works based on the principle of analogical reasoning mechanism (Bouchon-Meunier and Valverde 1999; Turksen and Zhong 1988). It derives the conclusion by transforming the intermediate consequent in terms of the similarity measured

between the antecedent of the intermediate rule and the observation, in an analogical manner as transforming the intermediate rule antecedent to the given observation. As one of the outstanding intermediate rule based FRI methods, the foundational T-FRI methodology was first introduced by Huang and Shen (2006, 2008). Many follow-on developments and modifications to this seminal approach have been proposed over the last two decades.

Apart from the above two major groups of FRI methods to conduct fuzzy interpolative reasoning, there are alternative FRI techniques, as summarised in Table 5. This shows the diversity of this interesting research area. For the purpose of demonstrating the basic ideas of typical FRI methods, several commonly used FRI approaches from each of the two main categories are reviewed below. As indicated previously, the triangular fuzzy membership functions, as defined in Sect. 2 are employed throughout, unless otherwise stated, both for consistency in demonstration of the ideas and for efficiency in computation.

3.2 Representative α -cut based FRI

As the seminal approach to FRI, the α -cut based interpolation is essentially a fuzzy extension of the classical linear interpolation of given points that are linked with fuzzy rules. The interpolated result is generated through the computation and then, the aggregation of linear interpolation at each α -cut level. Theoretically, in the case of arbitrary shaped convex normalised fuzzy sets an infinite number of α -levels should be taken into consideration for an approximate conclusion. In practice however, to achieve an acceptable computational requirement, most α -cut based methods only take a finite number of α -levels (usually two, three or four) into account, with the resulting points being connected piecewise linearly to yield an approximation of the consequent.

3.2.1 KH: foundational linear FRI

This section first formulates the basic idea of the most famous α -cut based FRI, named KH linear FRI (after its inventors Kóczy and Hirota 1993a, b), in a general formation, followed by its practical implementation by the use of triangular membership functions in a multidimensional situation.

Table 5 Alternative FRI techniques

Methods	Characteristics
Kovács and Kóczy (1997a, b, c)	Interpolation based on approximation of vague environment of fuzzy rules with application to automatic guided vehicle systems
Bouchon-Meunier et al. (1999, 2001, 2000) Jenei (2001), Jenei et al. (2002)	Interpolative method based on graduality Axiomatic approach for interpolation and extrapolation of fuzzy quantities
Yam et al. (2000b), Yam and Kóczy (2000, 1998, 2001) and Yam et al. (2000a)	Cartesian based interpolation with each fuzzy set mapped onto a point in high dimensional Cartesian space
Yang et al. (2019, 2021)	Group rule interpolation for constructing Adaptive Neuro-Fuzzy Inference System (ANFIS) and an evolutionary computation-supported approach

3.2.1.1 Core principle The KH rule interpolation offers an initial proposal for fuzzy interpolative reasoning through manipulating α -cut distances. When a given observation fails to match any rule in the sparse rule base for firing, an interpolated consequent is constructed by performing a linear aggregation of the rule consequents of a number (usually two) of selected neighbouring rules closest to the observation. The aggregation operation complies with the general principle of similarity-based analogical reasoning, such that

The closer a rule's antecedent A^i (which is a logical aggregation of individual attribute values A_j^i) to the observation o^* , the closer the rule's consequent B^i to the outcome B^* that corresponds to o^* .

The similarity measure employed is specified by the use of fuzzy distances defined between a rule antecedent and the observation. That is, the smaller distance between A^i and o^* is, the more similar they are, with the corresponding B^i deemed to potentially make more contribution than otherwise to the consequent being sought.

Suppose that there are two rules r^i and r^j in the rule base R , which are formulated as shown in Eq. (1). Given an observation o^* [again, as per Eq. (1)], the notion of linear rule interpolation can be written as:

$$\frac{\tilde{d}(A^*, A^i)}{\tilde{d}(A^*, A^j)} = \frac{\tilde{d}(B^*, B^i)}{\tilde{d}(B^*, B^j)} \quad (3)$$

where

$$\tilde{d}(A^*, A^i) = \sqrt{\tilde{d}_{i1}^2 + \tilde{d}_{i2}^2 + \dots + \tilde{d}_{im}^2} \quad \tilde{d}_{it} = \tilde{d}(A_t^*, A_t^i), t = 1, 2, \dots, m \quad (4)$$

and \tilde{d} denotes the fuzzy distance between the two membership functions.

Fuzzy distance between two fuzzy sets is interpreted as a pair of *lower and upper fuzzy distances* between their α -cut sets, with respect to the *Resolution Principle* (Kóczy and Hirota 1993a). For a particular $\alpha \in [0, 1]$, the lower fuzzy distance $\tilde{d}_L(A, B)$ and upper fuzzy distance $\tilde{d}_U(A, B)$ are denoted as:

$$\tilde{d}_L(A, B) = D(\inf(A_\alpha), \inf(B_\alpha)) \quad \tilde{d}_U(A, B) = D(\sup(A_\alpha), \sup(B_\alpha)) \quad (5)$$

where D denotes the Minkowski distance, and $\inf(\cdot)$ and $\sup(\cdot)$ are the infimum and supremum of the α -cut concerned, respectively. Hence, the formula of linear rule interpolation [i.e., Eq. (3)] can be rewritten as:

$$\begin{aligned} \frac{\tilde{d}_L(A_\alpha^*, A_\alpha^i)}{\tilde{d}_L(A_\alpha^*, A_\alpha^j)} &= \frac{\tilde{d}_L(B_\alpha^*, B_\alpha^i)}{\tilde{d}_L(B_\alpha^*, B_\alpha^j)} \\ \frac{\tilde{d}_U(A_\alpha^*, A_\alpha^i)}{\tilde{d}_U(A_\alpha^*, A_\alpha^j)} &= \frac{\tilde{d}_U(B_\alpha^*, B_\alpha^i)}{\tilde{d}_U(B_\alpha^*, B_\alpha^j)} \end{aligned} \quad (6)$$

This leads to the solution for $\min\{B_\alpha^*\}$ and $\max\{B_\alpha^*\}$ being:

$$\begin{aligned} \min\{B_\alpha^*\} &= \frac{w_{\alpha L}^i \min\{B_\alpha^i\} + w_{\alpha L}^j \min\{B_\alpha^j\}}{w_{\alpha L}^i + w_{\alpha L}^j} \\ \max\{B_\alpha^*\} &= \frac{w_{\alpha U}^i \max\{B_\alpha^i\} + w_{\alpha U}^j \max\{B_\alpha^j\}}{w_{\alpha U}^i + w_{\alpha U}^j} \end{aligned} \quad (7)$$

where

$$\begin{aligned} w_{\alpha L}^i &= \frac{1}{\tilde{d}_L(A_\alpha^*, A_\alpha^i)} & w_{\alpha L}^j &= \frac{1}{\tilde{d}_L(A_\alpha^*, A_\alpha^j)} \\ w_{\alpha U}^i &= \frac{1}{\tilde{d}_U(A_\alpha^*, A_\alpha^i)} & w_{\alpha U}^j &= \frac{1}{\tilde{d}_U(A_\alpha^*, A_\alpha^j)} \end{aligned} \quad (8)$$

The α -cut of conclusion is then given by

$$B_\alpha^* = [\min\{B_\alpha^*\}, \max\{B_\alpha^*\}] \quad (9)$$

and the interpolated conclusion can therefore, be obtained by the use of *Resolution Principle*, such that

$$B^* = \bigcup_{\alpha \in [0,1]} B_\alpha^* \quad (10)$$

3.2.1.2 Multidimensional implementation The foundational KH FRI works effectively and efficiently for simple linear problems. It has been subsequently developed to address sparse rule interpolation in more complex situations, for instance involving multiple rules with multiple antecedent variables (see Tikk et al. 2002; Wong et al. 2005). Thanks to the piecewise linear property presumed by KH interpolation, given triangular membership functions, the interpolated outcome $B^* = (b_1^*, b_2^*, b_3^*)$ can be determined with its two α -cut sets (when α is 0 or 1), resulting in the three characteristic points taking the values of

$$b_t^* = \frac{\sum_{i=1}^n \frac{1}{\sqrt{\sum_{j=1}^m (a_{jt}^i - a_{jt}^*)^2}} b_t^i}{\sum_{i=1}^n \frac{1}{\sqrt{\sum_{j=1}^m (a_{jt}^i - a_{jt}^*)^2}}} \quad (11)$$

where n is the number of the neighbouring rules used for interpolation, m is the number of attributes in the rule, and $t = 1, 2, 3$. Such computation for the interpolated fuzzy set B^* reflects exactly the general situation as expressed by Eq. (7), where

$$\begin{aligned} b_1^* &= \min\{B_0^*\} \\ b_3^* &= \max\{B_0^*\}, & \alpha &= 0 \\ b_2^* &= \min\{B_1^*\} = \max\{B_1^*\}, & \alpha &= 1 \end{aligned} \quad (12)$$

3.2.2 CCL rule interpolation

As one of the most popularly used α -cut based FRI methods, the CCL rule interpolation (named after its inventors Chang et al. 2008) offers an alternative means for fuzzy interpolative reasoning that exploits the areas of the fuzzy sets involved in the rules and the (unmatched) observation. The idea is to preserve the logically consistent properties with respect to the *ratio of fuzziness* (RF), which is determined by the areas of the fuzzy sets concerned. That is, it pursues consistency of RF between the (to be) interpolated consequent over the observation and the consequent value over the antecedent value of each rule used for interpolation. More specifically, the RF between two fuzzy values A and B is defined by

$$RF(A, B) = \frac{S(A)}{S(B)} \quad (13)$$

where $S(A)$, $S(B)$ denote the area of the fuzzy set of A and that of B , respectively.

The CCL FRI method presents a flexible interpolative reasoning framework, allowing the use of different types of membership function (MF), including various polygonal typed and Gaussian shaped MFs. It can also handle general cases that involve multiple antecedent variables involved in multiple fuzzy rules. For simplification and consistency throughout, the core computations are summarised below in relation to the use of triangular fuzzy membership functions.

First, the normal point b_2^* of the (to be) interpolated consequent B^* is defined by linear interpolation, such that

$$S_K(B^*) = \begin{cases} \left(\sum_{j=1}^m S_K(A_j^*) \right) \times \left(\sum_{i=1}^n W_i \times \frac{S_K(B^i)}{\sum_{j=1}^m S_K(A_j^i)} \right), & \text{if } \exists ij S_K(A_j^i) > 0 \\ \frac{\sum_{j=1}^m S_K(A_j^*)}{m}, & \text{if } \forall ij S_K(A_j^i) = 0 \end{cases} \quad (14)$$

$$b_2^* = \sum_{i=1}^n W_i b_2^i \quad (15)$$

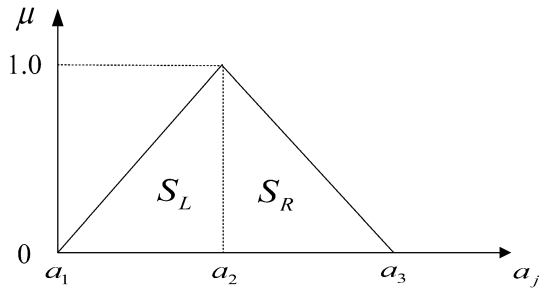
in which n is the number of selected rules for interpolation, and W_i is the aggregated rule weight, which is calculated by

$$W_i = \frac{\sum_{j=1}^m w_{ij}}{\sum_{i=1}^n \sum_{j=1}^m w_{ij}}, \quad w_{ij} = 1 - \left| \frac{a_{j2}^i - a_{j2}^*}{\max_{a_{j2}} - \min_{a_{j2}}} \right| \quad (16)$$

where $\max_{a_{j2}}$ and $\min_{a_{j2}}$ are used for normalisation, denoting the maximal and minimal value within $\{a_{j2}^i | i = 1, 2, \dots, n\}$.

Given the three characteristic points created from the two α -cut sets (when $\alpha = 0, 1$), a triangular fuzzy set is divided into two smaller sub-triangles, as shown in Fig. 3 (more triangular or even trapezoidal shaped sub-polygons may be generated for more complex polygonal fuzzy sets with many characteristic points, but the same idea is followed as

Fig. 3 Left area S_L and right area S_R of triangular fuzzy set



herein). From this, the left triangular area $S_L(B^*)$ (i.e., the part of the geometrical area of a triangular fuzzy set on the left hand side of the normal point) and the right triangular area $S_R(B^*)$, of the fuzzy set B^* are calculated by Eqs. (14 and 15), where for the subscript S_K , $K \in \{L, R\}$. This equation exactly reveals the basic idea of the CCL rule interpolation, where the RF from the observation viewpoint is constructed by the weighted aggregation of the RF of the involved rules, thereby leading to the derivation of the area of the interpolated fuzzy set.

Finally, the left and right extreme points of the support for the interpolated result B^* are derived from the resulting triangular areas as follows:

$$b_1^* = b_2^* - 2S_L(B^*), \quad b_3^* = b_2^* + 2S_R(B^*) \quad (17)$$

3.3 Representative intermediate rule based FRI

This section reviews the underlying interpolation mechanism of the intermediate rule (or transformation) based FRI. In particular, more detailed description is given to the scale and move transformation-based FRI approach since it has been continuously investigated for decades and widely applied.

3.3.1 Representative value of fuzzy set

Prior to going through the details of intermediate rule based FRI techniques, a very important concept needs to be introduced, which is adopted within this type of interpolation algorithm. This is the *Representative Value (Rep)* of a fuzzy set. There are actually many variations in the literature (e.g., Baranyi et al. 2004; Chen and Chang 2011b; Chen et al. 2009; Huang and Shen 2008), assigned with different names, such as representative value in Huang and Shen (2008), reference point in Baranyi et al. (2004), and characteristic value in Chen and Ko (2008) and Chen et al. (2009). Nonetheless, they imply similar interpretations as described below with the term of representative value.

The representative value of a fuzzy set is a single value assigned to help capture important information contained by the set in a simplified way, such as the “most typical” overall location of the fuzzy set in its domain and also, its geometric shape. In certain situations, the *Rep* value may be defined by the defuzzified value of the fuzzy set if that is preferred since there is no unified definition. What is important is within a particular FRI method, all *Rep* values are computed in the same way.

More formally, as with the most popular approach in the literature, given an arbitrary polygonal fuzzy set $A = (a_1, a_2, \dots, a_n)$ where $a_i, i = 1, 2, \dots, n$ denote the characteristic points of the polygonal, its representative value $Rep(A)$ is defined by:

$$Rep(A) = \sum_{i=1}^n w_i a_i \quad (18)$$

where w_i is the weight assigned to the characteristic point a_i per i . In particular, the simplest case, which is named the average Rep , is one so computed where all points take the same weight value, i.e., $w_i = 1/n$. For a triangular fuzzy set $A = (a_1, a_2, a_3)$ as shown in Fig. 2 (of Sect. 2), $Rep(A)$ is commonly and simply defined as follows in the literature (though its centre of gravity may also be used as an alternative if preferred):

$$Rep(A) = \frac{a_1 + a_2 + a_3}{3} \quad (19)$$

The definition of representative values for more complex membership functions can be found in Huang and Shen (2008).

Apart from its geometrical meaning, the Rep value also simplifies the definition of the distance between fuzzy sets, to measure the degree of “closeness”. A simple case of the distance between two fuzzy sets A and B can be defined by

$$d(A, B) = |Rep(A) - Rep(B)| \quad (20)$$

which is a crisp distance in contrast with α -cut distance based methods (Baranyi et al. 2004). The distance definition employed in a given FRI approach will be specified in each method later.

3.3.2 Scale and move transformation-based FRI (T-FRI)

The scale and move transformation based FRI (T-FRI) is one of the most general and advanced intermediate rule based FRI mechanisms. One of the key aims of this development has been to eliminate an important practical issue that earlier work of FRI had in that the interpolated outcomes were not guaranteed to be convex and in certain cases, not even a fuzzy set. The presentation of the fundamental idea of T-FRI is reported in Huang and Shen (2006, 2008), which can handle both interpolation and extrapolation of multiple multi-antecedent rules with complex polygon shaped, Gaussian and bell-shaped fuzzy membership functions. The following outlines the key computational steps of T-FRI working with multiple fuzzy rules where in general, multiple rule antecedents are involved in each rule.

Given a sparse rule base R and an observation o^* , in the form of Eq. (1), T-FRI works by running a computational process as highlighted in Fig. 4, involving four core procedures as summarised below.

- Step 1: Closest rules selection

This procedure is required as the basis upon which to perform FRI, when o^* does not match any of the rules in the rule base. Intuitively, it searches for a certain number of rules that are closest to the observation. The distance between an observation o^* and a rule r^q , or the distance

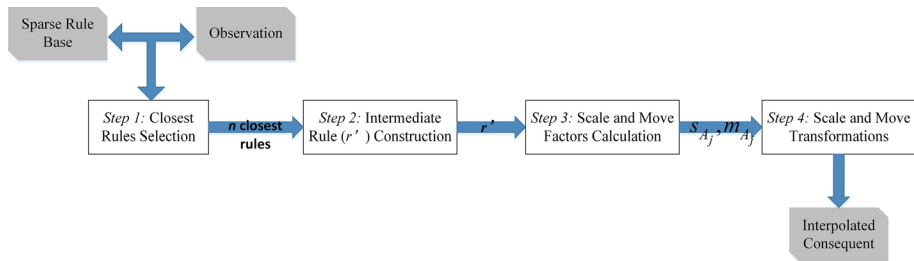


Fig. 4 Framework of scale and move transformation-based FRI

between any two rules $r^p, r^q \in R$, is determined by computing the aggregated distances over all the corresponding fuzzy values of the shared attributes between them:

$$d(v, r^q) = \frac{1}{\sqrt{m}} \sqrt{\sum_{j=1}^m d(A_j^v, A_j^q)^2} \quad (21)$$

where v is o^* or r^p (so A_j^v is A_j^* or A_j^p), depending on whether the distance is between an observation and a rule or between two rules. So, the n closest rules to o^* are those rules leading to the n smallest values of this distance measurement. In computing the aggregation regarding the above definition, the distance between a pair of antecedent fuzzy sets is calculated as below:

$$d(A_j^v, A_j^q) = \frac{|Rep(A_j^v) - Rep(A_j^q)|}{max_{A_j} - min_{A_j}} \quad (22)$$

This is implemented by the use of the Rep values of the corresponding fuzzy sets (defined in Sect. 3.3.1), representing the normalised result of the otherwise absolute distance, where max_{A_j} and min_{A_j} denote the maximal and minimal value of the attribute a_j , respectively. This normalisation is to ensure that all distance measures are compatible with each other over different attribute domains.

- Step 2: Intermediate fuzzy rule construction

From the above, n closest rules to a given observation can be chosen which have the minimal distances amongst all the rules with respect to the observation. Next, an intermediate fuzzy rule r' is constructed, forming the start point of the transformation process in T-FRI. In most applications of T-FRI, n is taken to be 2 purely for computational efficiency, but often at the expense of interpolative accuracy [if all rule antecedents are regarded as of having equal significance (Li et al. 2020b)].

The construction procedure computes the antecedent fuzzy sets $A'_j, j = 1, \dots, m$ and the corresponding consequent fuzzy set B' of the intermediate rule:

$$r': \text{ if } a_1 \text{ is } A'_1 \text{ and } a_2 \text{ is } A'_2 \text{ and } \dots \text{ and } a_m \text{ is } A'_m, \text{ then } z \text{ is } B'$$

which is a weighted aggregation of the n closest rules. Let $w_j^i, i \in \{1, \dots, n\}$, denote the weight to which the j th antecedent of the i th fuzzy rule contributes to the construction of the j th antecedent A_j' of the intermediate fuzzy rule:

$$w_j^i = \frac{1}{1 + d(A_j^i, A_j^*)} \quad (23)$$

where $d(A_j^i, A_j^*)$ is calculated as per Eq. (22). Then,

$$A_j' = A_j'' + \delta_{A_j} (\max_{A_j} - \min_{A_j}) \quad (24)$$

with

$$A_j'' = \sum_{i=1, \dots, n} \hat{w}_j^i A_j^i \quad (25)$$

where \hat{w}_j^i denotes the normalised weight and δ_{A_j} is a constant (termed the shift factor of A_j), defined respectively by

$$\hat{w}_j^i = \frac{w_j^i}{\sum_{i=1, \dots, n} w_j^i}, \quad \delta_{A_j} = \frac{|Rep(A_j^*) - Rep(A_j'')|}{\max_{A_j} - \min_{A_j}} \quad (26)$$

The consequent value of the intermediate rule is constructed in the same manner as above, that is

$$B' = B'' + \delta_z (\max_z - \min_z) \quad (27)$$

where \max_z and \min_z are the maximal and minimal values of the consequent attribute, B'' is the weighted aggregation of the consequent values of the n closest rules $B^i, i = 1, \dots, n$:

$$B'' = \sum_{i=1, \dots, n} \hat{w}_z^i B^i \quad (28)$$

with \hat{w}_z^i being the mean of the normalised weights associated with the antecedents \hat{w}_j^i in each rule:

$$\hat{w}_z^i = \frac{1}{m} \sum_{j=1}^m \hat{w}_j^i \quad (29)$$

and the shift factor δ_z of the consequent is the mean of $\delta_{A_j}, j = 1, \dots, m$

$$\delta_z = \frac{1}{m} \sum_{j=1}^m \delta_{A_j} \quad (30)$$

• Step 3: Scale and move factors calculation

As a transformation based FRI method, the goal of the transformation process T in T-FRI is to scale and move an intermediate fuzzy set A_j' , such that the transformed shape and representative value coincide with those of the observed value A_j^* . This process is implemented in the following two stages:

1. Scale operation from A'_j to \hat{A}'_j (denoting the scaled intermediate fuzzy set).

In order to implement this, the required scale rate s_{A_j} is first determined. As the specification of the scale (and the subsequent move) factors are dependent upon the fuzzy membership functions used. Given a triangular fuzzy set $A'_j = (a'_{j1}, a'_{j2}, a'_{j3})$, the scale rate s_{A_j} is:

$$s_{A_j} = \frac{a_{j3}^* - a_{j1}^*}{a'_{j3} - a'_{j1}} \quad (31)$$

which essentially expands or contracts the support length of $A'_j : a'_{j3} - a'_{j1}$ so that it becomes the same as that of A_j^* . The scaled intermediate fuzzy set \hat{A}'_j , which has the same representative value as A'_j , is then obtained such that

$$\begin{aligned} \hat{a}_{j1}' &= \frac{(1 + 2s_{A_j})a'_{j1} + (1 - s_{A_j})a'_{j2} + (1 - s_{A_j})a'_{j3}}{3} \\ \hat{a}_{j2}' &= \frac{(1 - s_{A_j})a'_{j1} + (1 + 2s_{A_j})a'_{j2} + (1 - s_{A_j})a'_{j3}}{3} \\ \hat{a}_{j3}' &= \frac{(1 - s_{A_j})a'_{j1} + (1 - s_{A_j})a'_{j2} + (1 + 2s_{A_j})a'_{j3}}{3} \end{aligned} \quad (32)$$

2. Move operation from \hat{A}'_j to A_j^* .

Given the scaled intermediate fuzzy set \hat{A}'_j , the move ratio m_{A_j} can then be determined. As indicated above, the move operation shifts the position of \hat{A}'_j to becoming the same as that of A_j^* , while maintaining its representative value $Rep(\hat{A}'_j)$. This is achieved using the move ratio m_{A_j} :

$$m_{A_j} = \begin{cases} \frac{3(a_{j1}^* - \hat{a}_{j1}')}{\hat{a}_{j2}' - \hat{a}_{j1}'} & \text{if } a_{j1}^* \geq \hat{a}_{j1}' \\ \frac{3(a_{j1}^* - \hat{a}_{j1}')}{\hat{a}_{j3}' - \hat{a}_{j2}'} & \text{otherwise} \end{cases} \quad (33)$$

This step computes and records all such scale rates and move ratios for use in the subsequent, and final, procedure to obtain the required consequent value, in response to the observation otherwise unmatched by any rule in the sparse rule base.

• Step 4: Scale and move transformations

After calculating the necessary scale and move factors (i.e., s_{A_j} and $m_{A_j}, j = 1, \dots, m$), this procedure completes the T-FRI process, deriving the required consequent value of B^* . This follows the intuition of similar observations leading to similar consequents, a heuristic fundamental to analogical approximate reasoning. For this, the transformation factors on the antecedent attributes are aggregated. In all conventional T-FRI methods, this is implemented by averaging them:

$$s_z = \frac{1}{m} \sum_{j=1}^m s_{A_j} \quad m_z = \frac{1}{m} \sum_{j=1}^m m_{A_j} \quad (34)$$

This entails the computation of scaled $\hat{B}' = (\hat{b}'_1, \hat{b}'_2, \hat{b}'_3)$:

$$\begin{aligned} \hat{b}'_1 &= \frac{(1 + 2s_z)b'_1 + (1 - s_z)b'_2 + (1 - s_z)b'_3}{3} \\ \hat{b}'_2 &= \frac{(1 - s_z)b'_1 + (1 + 2s_z)b'_2 + (1 - s_z)b'_3}{3} \\ \hat{b}'_3 &= \frac{(1 - s_z)b'_1 + (1 - s_z)b'_2 + (1 + 2s_z)b'_3}{3} \end{aligned} \quad (35)$$

where $B' = (b'_1, b'_2, b'_3)$ is the fuzzy value of the intermediate consequent previously computed. From this, again, by analogy to the transformation required for the antecedent to match the observation, move transformation is applied, resulting in the final, required interpolated consequent $B^* = (b^*_1, b^*_2, b^*_3)$:

$$\begin{aligned} b^*_1 &= \hat{b}'_1 + m_z \gamma \\ b^*_2 &= \hat{b}'_2 - 2m_z \gamma \\ b^*_3 &= \hat{b}'_3 + m_z \gamma \end{aligned} \quad \gamma = \begin{cases} \frac{\hat{b}'_2 - \hat{b}'_1}{3}, & \text{if } m_z \geq 0 \\ \frac{\hat{b}'_3 - \hat{b}'_2}{3}, & \text{otherwise} \end{cases} \quad (36)$$

For illustration, Fig. 5 outlines the scale and move transformation process (i.e., Steps 3 and 4 of a typical T-FRI method), where the scale and move factors of each rule antecedent are shown to be calculated in the upper box and the interpolated result is obtained by the corresponding transformations shown underneath. For conciseness, such a process can be collectively represented by: $B^* = T(B', s_z, m_z)$, emphasising on the significance of both scale and move transformations.

3.3.3 Representative modifications to scale and move transformation-based FRI

Following the generic and seminal ideas of the above-reviewed T-FRI approach, there have been a large family of works that have been proposed to further improve it, of which an overview is previously shown in Table 4. This section provides an outline of representative methods within this family.

- Adaptive T-FRI (Yang and Shen 2011; Yang et al. 2017)

This work is motivated by an observation that there may exist inconsistency in interpolated results after a sequence of T-FRI operations. The potential reasons have been analysed to include defective interpolated rules and inaccurate interpolative transformations. The adaptive fuzzy interpolation enhances the original T-FRI with the ability for identification and correction of defective rules in interpolative transformations, facilitating the removal of certain inconsistencies. This is accomplished through two sub-systems: (1) a diagnostic sub-system that is constructed by the use of the general diagnostic engine, where the inconsistent interpolated results are recorded in an ATMS (assumption-based truth maintenance

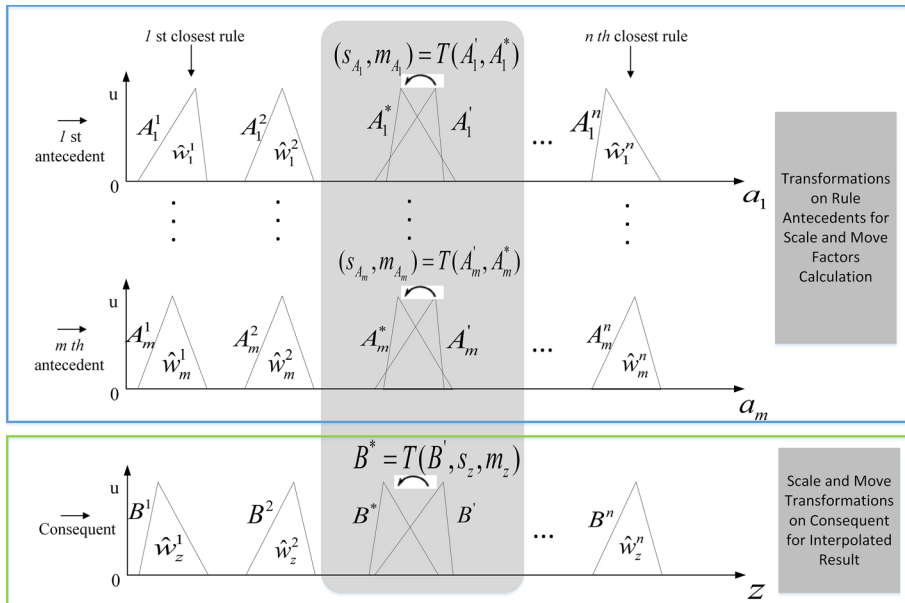


Fig. 5 Fuzzy rule interpolation via scale and move transformations

system) (deKleer 1986); and (2) a corrective sub-system that is derived from a fuzzy extension to the traditional numerical interpolation theory and its application in approximation computation. However, this work is focussed on the implementation of adaptive T-FRI that involves just two multiple-antecedent rules. Besides, further investigation is required to reveal whether it can handle situations where extrapolation is necessary (since the original T-FRI is able to deal with extrapolation in the same manner as with interpolation).

- Backward T-FRI (Jin et al. 2014, 2019)

Conventional FRI generally executes in a “forward” manner, where the consequent is required to be interpolated given the rule base and all antecedent attributes of an observation available. Nevertheless, situations may arise when certain crucial antecedents are absent from the given observation, which may also be involved in the subsequent interpolation process, thereby leading to the failure of the derivation of the final interpolated conclusion. This important issue is addressed by a modification of T-FRI, termed backward T-FRI, which provides a series of solutions for handling both single missing antecedent value and multiple missing antecedents problems. The single missing antecedent issue is resolved by implementing a four-step computation procedures (mirroring what is presented in Sect. 3.3.2) of the original T-FRI, resulting in the reverse calculation of the relative parameters corresponding to the unknown antecedent value. The general backward T-FRI with multiple missing antecedent values is addressed by two procedures: (1) a direct extension of the method for the single missing value case, by estimating parameter combinations that would lead to the closest resemblance of the original (missing) values; and (2) an approach for the removal of possible missing antecedent values through a process of verifying interpolative results against (past) observations. The modification for backward T-FRI is proven to preserve many crucial properties that the original T-FRI possesses, e.g., the capability in handling multiple multi-antecedent rules and the maintenance

of convexity and normality of interpolated results. Whilst backward T-FRI helps address the problem of missing antecedent attribute values it does not totally remove this problem, especially when the scale of missing values becomes substantial.

- Dynamic T-FRI (Naik et al. 2017b)

A great majority of the existing transformation based FRI mechanisms work on a static sparse rule base. However, the use of a static rule base may affect the effectiveness of FRI due to the absence of the most concurrent (dynamic) rules as the requirements of fuzzy systems may change over time. Yet, a volume of intermediate fuzzy rules are typically generated from this type of FRI methods while executing rule interpolation. Collectively, they may gradually cover regions that were uncovered by the original sparse rule base, thereby offering possibly valuable information for updating the static sparse rule base. From this observation, the work of Naik et al. (2017b) makes use of such intermediate rules which are otherwise discarded once the required outcomes have been obtained within the most of transformation based FRI methods, to develop a dynamic T-FRI mechanism. It enriches the rule base by refining and promoting the intermediate rules, gaining efficiency by allowing for more direct rule-firing while minimising future running of the interpolation procedure. It is implemented by first partitioning the interpolated rules into hypercubes, where the nonempty ones are fed as the input into a Genetic Algorithm-based clustering algorithm to find the “best” cluster arrangement. An iterative process is then run to select the densest clusters that have accumulated a sufficient number of candidate rules for achieving the rule aggregation and promotion. The practical significance of this approach is obvious. Further reinforcement is however, still possible, say, by employing more effective and efficient clustering and optimisation methods to replace the relevant components within the current implementation.

- Higher order T-FRI (Chen et al. 2016; Chen and Shen 2017)

A common implementation shared by most of the established FRI methods is that fuzzy membership functions in the rules and observations are generally defined with conventional type-1 fuzzy sets, for the interpretation and treatment of uncertainty. Very little of the existing FRI work can conjunctively handle more than one form of uncertainty in the rules or observations, despite there may be cases in which more complicated fuzzy set representations become necessary (Fu and Shen 2010). In response to this observation, a higher order FRI has been developed, allowing for the representation and manipulation of different types of uncertainty in knowledge within the common T-FRI framework. It works by first encoding uncertain knowledge with higher order representation and then, by deriving the final conclusions through performing higher order interpolation over models of such representation. In particular, two common types of technique for uncertainty representation are exploited, resulting in: (1) a rough-fuzzy set based rule interpolation approach (Chen et al. 2016), which facilitates the representation of uncertain fuzzy membership functions with rough-fuzzy approximations; and (2) an interval type-2 fuzzy set-based approach (Chen and Shen 2017), which works in the same manner as with the rough-fuzzy-based T-FRI. Within either method, the concept of representative value of a fuzzy set also plays an indispensable role as within the original type-1 T-FRI. Another type-2 fuzzy set-based FRI method can be found in Chen and Lee (2011). Such methods require relative modifications corresponding to each particular uncertainty representation, which inevitably increases the computational complexity as the cost for exchange of a much more general T-FRI mechanism that will collapse back to the type-1 method if all uncertainty involved can be sufficiently captured by type-1 fuzzy sets.

- Other T-FRI-like approaches

Apart from the above-outlined modifications to T-FRI that are directly investigated and improved upon the original T-FRI method, there are other proposals for reinforcing fuzzy interpolative reasoning which are analogous to the basic ideas of T-FRI (Li et al. 2005). For instance, in Chen and Adam (2018), ranking values of arbitrary polygonal fuzzy sets are used to express the characteristic points of the underlying membership functions, which are in turn used to play a similar role in the modified transformation-based FRI process as the *Rep* values do in T-FRI. In addition, the scale and move transformations involved in T-FRI are replaced with the distance ratio and move rate, respectively, to transform the constructed intermediate rule in an effort to obtain the final interpolated outcome.

Another variation of T-FRI is reported in Chen and Ko (2008), named CK FRI hereafter to acknowledge its inventors. The classical *Rep* values are substituted by characteristic values in this work, facilitating not only the simplified representation of a fuzzy set but also the definition of the distance between fuzzy terms. For situations where polygonal fuzzy sets are involved, the interpolated fuzzy set being sought is derived by calculating each of the characteristic points that are obtained from a series of α -cuts. Particularly, the normal points (of which the membership is 1) are first determined, aiding in any subsequent calculation of the remaining points. From this, two transformations, namely increment and ratio transformations, are executed to convert the average consequent into the final interpolated outcome with the similarity degree measured between these two analogous to that of the average antecedent and the observation. Improved on this work further, two enhanced transformations have been introduced (Chen et al. 2009) to support weighted approaches to FRI (that will be addressed separately later in this review).

3.3.4 Generalised function-based FRI

Bearing significant similarity with the intermediate rule based FRI methods as outlined above is another approach, which is herein referred to as generalised function-based for convenience. Example methods falling within this family include those reported by Baranyi et al. (1995, 1996a, b, 1998, 2004), Baranyi and Kóczy (1996a, b). Unlike the α -cut based interpolation algorithms, given an unmatched observation, this approach infers the conclusion based on the interpolation of fuzzy relations instead of using α -cut distances. It works through two major steps which are briefly outlined below for academic completeness. Further details can be referred to the relevant references provided.

Given two fuzzy rules (say, r^1, r^2) and an observation (o^*) in the form of Eq. (1), the core of a generalised function-based FRI method can be described through the following two stages.

3.3.4.1 Generation of interpolated firing rule The aim of this first stage is to create an intermediate rule r' , in such a way that the antecedent of r' is as “close” to that of the observation (A^*) as possible. Note that the term “close” here stands for the case where at least partial overlapping is ensured between the observation and the intermediate rule. This implies that the firing of the resulting intermediate rule can be subsequently conducted (see the next stage). Denote the procedure of this stage by

$$r' = f^{Interpolation}(r^1, r^2) \quad (37)$$

where $f^{Interpolation}$ represents a mapping from a pair of rules onto a set of all possible rules of the form as per Eq. (1). There are two types of algorithm that may be utilised to implement this stage of the approach:

- Fuzzy relation interpolation, which includes any of the solid cutting methods (Baranyi et al. 1995, 1996a, b; Baranyi and Kóczy 1996a, b), and those based on the fixed point law or fixed value law (Ding et al. 1989, 1992; Mukaidono et al. 1990; Shen et al. 1993, 1988).
- Semantic relation interpolation, which includes any of the semantic revision methods (Ding et al. 1989, 1992; Mukaidono et al. 1990; Shen et al. 1993, 1988), using the semantic revision principle to describe the relation between the antecedent and consequent fuzzy sets within an interpolated intermediate rule.

3.3.4.2 Inference with single rule firing This second stage is to fire the intermediate rule returned by the first. This is enabled by temporarily regarding the newly generated intermediate rule as one of the existing rules within the rule base, and also by computing the overlapping between the observation and the antecedent of the intermediate rule. The procedure implementing this stage can be generally denoted by

$$B^* = f^{Inference}(r', A^*) \quad (38)$$

Exactly what mechanism to implement rule-firing may vary with respect to different FRI methods in this family. Any method reported in Ding et al. (1989, 1992), Mukaidono et al. (1990) and Shen et al. (1993, 1988) may be utilised to directly fire the transformed intermediate rule to compute the final consequent value required.

For simplicity, the above description has been focussed on interpolation for the cases where a single rule antecedent is considered. As with α -cut based and T-FRI approaches, the generalised function-based mechanism has also been extended to performing FRI with multiple rule antecedents and fuzzy extrapolation. More details can be found in Baranyi et al. (2004) and other derivatives. Overall, the workflow of such a method can be illustrated in Fig. 6. Conceptually, this is of course very similar to the underlying approach taken by T-FRI.

3.3.5 Geometry-based linear FRI (GLFRI)

As can be seen from the above, numerous contributions have been made to achieve fuzzy interpolative reasoning that perform with respect to a given sparse rule base. The review so far has been organised in terms of two fundamentally different approaches: α -cut based and

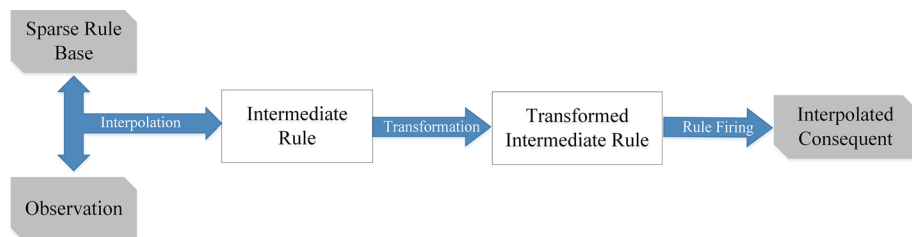


Fig. 6 Workflow of generalised function-based fuzzy rule interpolation

intermediate rule based. Whilst individual methods within each of these two groups share a number of similar characteristics, it is difficult to demonstrate their underlying interpolating patterns in a closed mathematical form (Yang and Shen 2013). A recent study presented in Das et al. (2019), referred to as geometry-based linear FRI (GLFRI), has attempted to solve this problem using the theory of fuzzy geometry, through which the basic idea of FRI can be geometrically visualised.

GLFRI belongs to the intermediate rule based FRI scheme, requiring a two-step procedure to implement fuzzy interpolative reasoning. This method first derives an intermediate rule as a convex combination of adjacent rules, then the final conclusion is obtained from the intermediate rule through executing geometrical transformation. In particular, fuzzy rules are interpreted as fuzzy sets that are represented as certain fuzzy points in a multi-dimensional space. As such, all the fuzzy rules expressed in the form of multidimensional fuzzy points may be joined by fuzzy line segments. The first step of the interpolative process is then to construct an intermediate rule corresponding to the observation, which is implemented via a convex combination of the neighbouring rules of the observation. Since the fuzzy rules have been equivalently represented as fuzzy points, the point signifying the required intermediate rule can be obtained by the use of classic linear interpolation of the crisp values by manipulating the fuzzy line segments formed by the point expressing the observation and those neighbouring points (i.e., rules) to it. In the second step, predefined expansion or contraction parameters are calculated to transform the resulting intermediate point, by comparing the difference between the antecedent part of this point and the observation, which are subsequently utilised to transform the consequent of intermediate rule to derive the expected conclusion.

Whilst GLFRI is promising in presenting the computational process of FRI mathematically in a closed form, with geometrical visualisation, this initial attempt takes into consideration of only convex normalised fuzzy sets with singleton core, further extension of this study involving non-normal and non-convex fuzzy sets may form a piece of interesting investigation.

3.4 Comparison of representative FRI methods over common criteria

From the above, it can be seen that in common, running an FRI algorithm, be it α -cut based or intermediate rule based, results in an inference consequent. This is in response to an unmatched observation via interpolating the fuzzy rules in a given rule base, thereby achieving the goal of interpolative reasoning. However, there are significant differences between these approaches. This subsection presents a comparison of a number of representative FRI methods reviewed previously, with respect to criteria that are commonly utilised in the literature for interpolation performance evaluation.

Theoretically, FRI is essentially a mapping (Tikk et al. 2011; Jenei 2001) that relates the input space \mathcal{A} and the output space \mathcal{Z} , where fuzzy subsets in the domain \mathcal{A} and \mathcal{Z} (denoted by $\mathcal{F}(\mathcal{A})$ and $\mathcal{F}(\mathcal{Z})$, respectively) indicate the values of rule antecedent attributes and the value of rule consequent [as defined in Eq. (1)]. That is, given a rule base R , $\forall r^i \in R$, the values $\{A_1^i, A_2^i, \dots, A_m^i\} \in \mathcal{F}(\mathcal{A})$, of the m antecedent variables, and the rule consequent value $B^i \in \mathcal{F}(\mathcal{Z})$. FRI pursues to define the correlation $I : \mathcal{F}(\mathcal{A}) \rightarrow \mathcal{F}(\mathcal{Z})$, which associates to an observation $A^*(= \{A_1^*, A_2^*, \dots, A_m^*\}) \in \mathcal{F}(\mathcal{A})$ an interpolated conclusion $I(A^*) = B^*$ where $B^* \in \mathcal{F}(\mathcal{Z})$. Thus, FRI methods are required to satisfy certain common properties as a mapping function, which also form the general criteria facilitating the comparison amongst them.

Table 6 summarises the most commonly used criteria for FRI evaluation in the literature. Any given FRI method is expected to meet or qualify at least a certain number of such properties to be effective in performing interpolative reasoning. Over the history of FRI development, a number of approaches that have been reported at the early stages have been analysed and compared against these criteria in the previous work of Johanyák and Kovács (2006), Tikk et al. (2011), especially for the α -cut based FRI methods including the seminal linear interpolation mechanism introduced in Kóczy and Hirota (1993a, b) and its derivatives. Such comparative discussion is therefore, not comprehensively included in the present review to avoid redundancy. Instead, particular attention is drawn for more recently developed FRI approaches, including many popular transformation-based techniques. As a summary, Table 7 presents the results of evaluating the representatives of such FRI methods, over the common criteria.

In general, it is not necessary that all such criteria are fulfilled in developing an FRI method. However, it is expected that most of the property should be satisfied, with other problem specific parameters to fulfill given a certain application. This also points out the trend in the development of FRI techniques, that is to amend the drawbacks of the existing FRI methods and to satisfy more criteria. For example, the very first proposal for FRI, KH linear interpolation (Kóczy and Hirota 1993a), is well-known that it cannot always guarantee the convexity of the derived fuzzy sets (i.e., C2 in Table 6) although they may be normal. This has led to much attention being paid to building FRI mechanisms that ensure not only normality but also convexity of inferred consequences. This in turn, has led many advanced variations of KH method. A number of recently developed FRI approaches are able to accomplish many key criteria successfully, including the listed C1–C7 as shown in Table 7. Also, criteria C9, C10 and C11 have increasingly become more demanded as more sophisticated fuzzy systems are constructed that enjoy more significant interpolative reasoning power.

4 Weighted FRI methodologies

In conventional fuzzy interpolative reasoning systems, multiple rules are generally involved with each concerning multiple rule antecedent attributes. However, these antecedent attributes are assumed to have equal significance when they are working together within the rule interpolation process. Recently, a number of methods have been proposed to weight the rule antecedents and to integrate the weights into the traditional algorithms where attributes are not weighted.

This section reviews such advances in the development of weighted fuzzy interpolative reasoning systems. Table 8 lists the titles of the methods being reviewed, and an acronym is assigned to each to act as the short name after its inventors while reflecting the year of the relevant publication. The rest of this section is organised by first summarising five typical approaches, followed by presenting a brief comparison amongst them, and finished with a qualitative, overall comparison between the weighted FRI methods and the original non-weighted approaches.

4.1 Typical weighted FRI approaches

This section reviews five representative fuzzy interpolative reasoning mechanisms which are achieved by weighted FRI. As indicated previously, two key issues, namely weight

Table 6 Commonly adopted criteria for FRI evaluation

No.	Criterion	Interpretation
C1	Validity of conclusion	Interpolated conclusion always leads to a valid fuzzy set
C2	Preservation of convexity and normality	If an observation is normal and convex so should the interpolated conclusion also be
C3	Compatibility with rule base	For all rules $r^i \in R$ and all $A^* \in \mathcal{F}(\mathcal{A})$: If $A^* = A^i$, then $\mathcal{J}(\mathcal{A}^*) = B^* = B^i$
C4	Continuity condition	For $\epsilon > 0$ there exists $\delta > 0$ s.t. if $A, A^* \in \mathcal{F}(\mathcal{A})$, and $d(A, A^*) \leq \delta$ then $d(\mathcal{J}(A), \mathcal{J}(A^*)) \leq \epsilon$, where $d(\cdot, \cdot)$ denotes a certain distance metric
C5	Preservation of piece-wise linearity	If fuzzy sets used for interpolation are piece-wise linear, so should interpolated conclusion be
C6	Preservation of “in between”	If A^* is in between A' and A'' , then interpolated conclusion $\mathcal{J}(A^*)$ should be in between $\mathcal{J}(A')$ and $\mathcal{J}(A'')$
C7	Use of arbitrary membership function	Mapping \mathcal{J} is applicable to any convex and normal form of membership functions
C8	Shape invariance	If all fuzzy sets in rule antecedents are of same type of membership function, so should interpolated $\mathcal{J}(A^*)$ be
C9	Applicability for multiple rules	Interpolation mapping \mathcal{J} can handle fuzzy interpolative reasoning with any number of rules
C10	Applicability for multidimensional input	Interpolation mapping \mathcal{J} is applicable to any finite number of input variables.
C11	Capability of extension to extrapolation	Interpolation mapping \mathcal{J} is extrapolatively extensible

Table 7 Evaluation of representative FRI methods over common criteria

FRI method	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
KH FRI	×	×	✓	×	✓	✓	×	✓ (†)	✓	✓	✓
	(† only if C1 is satisfied)										
CCL FRI	✓	✓	–	–	–	–	✓	✓	✓	✓	–
T-FRI	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adaptive T-FRI	✓	✓	✓	✓	✓ (†)	✓	✓	✓	×	✓	×
	(† first-order piecewise linear)										
Backward T-FRI	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
CK FRI	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	–
Dynamic T-FRI	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Generalised function-based FRI	✓	✓	✓	✓	✓ (†)	✓	✓ (★)	–	–	✓	✓
(† when involving triangular fuzzy sets)											
(★ for certain combinations of methods in this family)											

✓Denotes satisfaction of certain criterion

×Denotes dissatisfaction of certain criterion

–Denotes information unavailable to decide on certain criterion

learning and integration of weights in FRI, are the main concerns in implementing any weighted FRI approach. The following subsections are therefore composed of three parts for each method, by first reporting the development regarding these two issues and then drawing summarising remarks. To facilitate better understanding, all weighted FRI methods are outlined by the use of unified pseudo code for the main procedural steps.

4.1.1 Center of gravity-based weighted FRI (LHTZ2005)

A weighted FRI is presented in Li et al. (2005) as the original approach for fuzzy interpolative reasoning supported with antecedent weights. This is referred to as the LHTZ2005 method in Table 8 and hereafter. All weights in this approach are implemented by the use of trapezoidal fuzzy sets.

4.1.1.1 Learning weights There is little learning involved in LHTZ2005, but the weights of the individual rule antecedents are predefined by domain experts. Each antecedent attribute within different rules of the rule base is assigned a different weight. For instance, two weighted fuzzy rules used for weighted interpolation are therefore, represented such that

$$\begin{aligned}
 r^1 &: \text{if } a_1 \text{ is } A_1^1 (AW_1^1) \text{ and } a_2 \text{ is } A_2^1 (AW_2^1) \text{ and } \dots \text{ and } a_m \text{ is } A_m^1 (AW_m^1), \text{ then } z \text{ is } B^1(C_1) \\
 r^2 &: \text{if } a_1 \text{ is } A_1^2 (AW_1^2) \text{ and } a_2 \text{ is } A_2^2 (AW_2^2) \text{ and } \dots \text{ and } a_m \text{ is } A_m^2 (AW_m^2), \text{ then } z \text{ is } B^2(C_2)
 \end{aligned}
 \tag{39}$$

where AW_j^i stands for the weight of the j th antecedent variable ($j = 1, 2, \dots, m$) in the rule r^i , $i = 1, 2$, and C_i , $i = 1, 2$ is the certainty factor of r^i . Note that all of the AW_j^i and C_i are specified as trapezoidal fuzzy numbers. As such, the computational effort involved may generally increase significantly.

Table 8 Weighted fuzzy interpolative reasoning schemes with short names

Short Name	Title	Reference
LHTZ2005	Weighted fuzzy interpolative reasoning method	Li et al. (2005)
CC2008	A new method for multiple fuzzy rules interpolation with weighted antecedent variables	Chang and Chen (2008)
CKCP2009	Weighted fuzzy interpolative reasoning based on weighted increment transformation and weighted ratio transformation techniques	Chen et al. (2009)
CC2011a	Weighted fuzzy rule interpolation based on GA-based weight-learning techniques	Chen and Chang (2011b)
CC2011b	Weighted fuzzy interpolative reasoning for sparse fuzzy rule-based systems	Chen and Chang (2011a)
CLS2013	Weighted fuzzy interpolative reasoning systems based on interval type-2 fuzzy sets	Chen et al. (2013b)
CH2014	Weighted fuzzy interpolative reasoning based on the slopes of fuzzy sets and particle swarm optimization techniques	Chen and Hsin (2014)
DJS2014	Antecedent selection in fuzzy rule interpolation using feature selection techniques	Diao et al. (2014)
CC2016	Weighted fuzzy interpolative reasoning for sparse fuzzy rule-based systems based on piecewise fuzzy entropies of fuzzy sets	Chen and Chen (2016)
CA2018	Weighted fuzzy interpolated reasoning based on ranking values of polygonal fuzzy sets and new scale and move transformation techniques	Chen and Adam (2018)
LSLYS18	Fuzzy rule based interpolative reasoning supported by attribute ranking	Li et al. (2018b)
LSLYS19	Interpolation with just two nearest neighbouring weighted fuzzy rules	Li et al. (2020b)

4.1.1.2 Weighting FRI This weighted fuzzy interpolative reasoning process essentially extends the FRI mechanism of Huang and Shen (2003) that uses only triangular fuzzy sets. In this work, the center of gravity (COG) of a fuzzy set is used to represent the fuzzy set for simplicity. In particular, the COG of a trapezoidal fuzzy set $A = (a, b, c, d)$ is defined as a pair (h_L, h_R) :

$$h_L = \frac{1}{3}(a + b + d) \quad h_R = \frac{1}{3}(a + c + d) \quad (40)$$

where a, b, c, d denote the characteristic points of A with a and d having a membership of 0, and b and c being the odd normal points (i.e., the two extrema points of the nuclear of the trapezoidal with a membership of 1).

The distance between two trapezoidal fuzzy numbers A_1 and A_2 are defined using their COG pairs [namely, (h_{1L}, h_{1R}) and (h_{2L}, h_{2R})] as follows:

$$d(A_1, A_2) = \frac{1}{2}(h_{2R} + h_{2L} - h_{1R} - h_{1L}) \quad (41)$$

From this, LHTZ2005 can be summarised in Algorithm 1, showing the main execution steps. Note that the weights of rule antecedents are typefaced in bold wherever they are integrated within FRI in order to highlight the weighting mechanism.

Algorithm 1 Weighted FRI in LHTZ2005

Input: – Rules r^1, r^2 in form of Eqn. (39)

- Observation $o^* = \{A_1^*, \dots, A_m^*\}$, over m conditional attributes
- Individual weight, of rule antecedent variable $AW_j^i, i = 1, 2, j = 1, 2, \dots, m$
- Certainty factors $C_i, i = 1, 2$ of rules $r^i, i = 1, 2$

Output: – Interpolated consequent B^*

- 1: Construct a new inference rule by manipulating two given adjacent rules r^1, r^2 , in form of

$$r' : \text{if } a_1 \text{ is } A'_1 \text{ and } a_2 \text{ is } A'_2 \text{ and } \dots \text{ and } a_m \text{ is } A'_m, \text{ then } z \text{ is } B'$$

where $A'_j (j = 1, 2, \dots, m)$ is obtained by

$$A'_j = (1 - \lambda_j)A_j^1 + \lambda_j A_j^2$$

with λ_j calculated using distance between fuzzy sets as per Eqn. (41):

$$\lambda_j = \frac{d(A_j^1 \otimes \mathbf{AW}_j^1 \otimes C_1, A_j^*)}{d(A_j^1 \otimes \mathbf{AW}_j^1 \otimes C_1, A_j^2 \otimes \mathbf{AW}_j^2 \otimes C_2)}$$

where \otimes denotes multiplication operator between trapezoidal fuzzy sets. B' is calculated similarly, such that

$$B' = (1 - \lambda_c)B^1 + \lambda_c B^2, \quad \lambda_c = \frac{1}{m} \sum_{j=1}^m \lambda_j$$

- 2: Calculate scale rate s_j and move rate l_j ($j = 1, 2, \dots, m$) for each rule antecedent to assess difference between A'_j and A_j^* .
- 3: Aggregate scale and move rate for consequent variable by

$$s_c = \frac{1}{m} \sum_{j=1}^m s_j, \quad l_c = \frac{1}{m} \sum_{j=1}^m l_j$$

- 4: Calculate consequent fuzzy set B^* by transforming B' using s_c, l_c . Note that antecedent weights are not involved here (more details can be referred to original work of Li et al. (2005)).
 - 5: **Return** Interpolated consequent B^*
-

4.1.1.3 Remarks

1. The weights of rule antecedent variables in this approach are assumed to be predefined, which requires human intervention and hence, adversely reduces the flexibility of the resulting fuzzy interpolative reasoning system.
2. As reflected in Algorithm 1, the individual weights for antecedent variables are only involved in the calculation of the aggregation factors $\lambda_j, j = 1, 2$ while constructing the new inference rule r' (as shown in Line 1). The aggregation over the scale and move rates to compute the consequent variable is simply implemented by an algebraic average of the corresponding antecedent items (Line 3), which are externally assigned, to signify their individual significance levels in influencing the consequent.
3. Computational complexity may be increased significantly due to the use of trapezoidal fuzzy sets to represent the weights, but the interpretability may be improved if these weights are associated with domain-specific linguistic terms (which is possible given they are defined by the domain experts).

4.1.2 GA-based weighted FRI (CC2011a)

The method of genetic algorithm (GA)-based weighted fuzzy interpolative reasoning integrates the weights of rule antecedents within the underlying FRI procedure it adopts. This work is referred to as CC2011a with details given in Chen and Chang (2011b). In this method, the weights of the antecedent variables are automatically learned by the use of a GA-based weight-learning algorithm. The fuzzy sets are represented with polygonal or bell-shaped membership functions.

4.1.2.1 Learning weights The learning method for generating the optimal weights of the rule antecedent variables used for this weighted FRI is based on the CHC algorithm (Eshelman 1991), which is a specialisation of traditional GAs (Holland 1975). This GA-based learning mechanism encodes the weights of individual antecedent attributes into a chromosome, on which each gene represents an individual attribute weight.

An initial population is randomly generated, forming the start point of the evolutionary weight learning process. For each chromosome in the current population, it decodes a certain weight value, which is to be employed in the proposed weighted FRI. The weighted interpolative scheme is then triggered for a set of training samples, with the interpolated outcomes recorded. The selection of “good” chromosomes depends on a predefined fitness function by comparing the values between the inferred outputs and the target outputs of the training samples. From this, a crossover operation is carried out among the selected chromosomes, forming the next generation. Once the number of evolutions reaches a predefined maximum number of evolutions, this iterative weight learning procedure terminates and the chromosome with the largest fitness value is deemed the optimal. The final learned weights for the rule antecedent variables are obtained by decoding the optimal chromosome. In so doing, this GA-based weight learning scheme follows a so-called “wrapper” approach, which mixes up weight learning and weighted FRI procedures. The weights obtained from the current generation are required to be integrated within FRI, to enable the evaluation of the fitness values.

4.1.2.2 Weighting FRI A key concept employed for facilitating this weighted interpolative reasoning is the ratio of fuzziness (RF) (Chen and Chang 2011b). For polygonal fuzzy sets, the degree of fuzziness is computed in relation to the areas of the fuzzy sets. Let A and B be two polygonal fuzzy sets, the ratio of fuzziness $RF(A, B)$ of A to B is defined as follows:

$$RF(A, B) = \frac{S(A)}{S(B)} \quad (42)$$

where $S(A)$ and $S(B)$ denote the area of the membership function of A and that of B , respectively.

The central idea for the computation of the interpolated consequent in response to an (unmatched) observation is the following: The algorithm attempts to keep the RF of the fuzzy set of each attribute in the observation over that of the corresponding antecedent of a selected rule for interpolation the same as the RF of the fuzzy set of the required consequent (to be computed) over that of an artificially constructed intermediate rule consequent. The intermediate rule consequent is herein generated by aggregating the consequents of rules that are involved for interpolation. This idea in effect reflects fuzzy or generalised modus ponens, a keen to the approach taken by T-FRI.

The weights of the rule antecedent variables are integrated within FRI by following the routine which is summarised in Algorithm 2. In order to emphasise the role that those antecedent weights play in the entire weighted FRI procedure, the individual weights are highlighted in bold in this algorithm description.

Algorithm 2 Weighted FRI in CC2011a

Input: – Rule base $R = \{r^1, \dots, r^N\}$ consisting of N rules in form of Eqn. (1)

– Observation $o^* = \{A_1^*, \dots, A_m^*\}$, over m conditional attributes

– Number of closest rules n

Output: – Interpolated consequent B^*

- 1: Initialise randomly individual weights of rule antecedent variables $AW_j, j = 1, 2, \dots, m$;
- 2: Calculate representative values for antecedent fuzzy sets $A_j, j = 1, \dots, m$ of rules $r^i, i = 1, \dots, N$ and o^* , namely $Rep(A_j^i)$ and $Rep(A_j^*)$;
- 3: Select n nearest neighbouring fuzzy rules with respect to o^* from rule base;
- 4: Calculate weight for each of selected rules $W_i, i = 1, \dots, n$, by aggregating **individual antecedent weights** $AW_j, j = 1, 2, \dots, m$ such that

$$W_i = \frac{Min_{1 \leq j \leq m} AW_j \lambda_{ij}}{\sum_{q=1}^n Min_{1 \leq j \leq m} AW_j \lambda_{qj}}$$

where λ_{ij} measures similarity between j th antecedent value of rule r^i and its corresponding value in o^* , using their representative values;

- 5: Construct (triangular) intermediate consequence fuzzy set $B' = (b'_1, b'_2, b'_3)$ (though any other polygonal MF may be used);
- 6: Divide membership function of each $A_j^i, i = 1, \dots, n$ and $A_j^*, j = 1, \dots, m$ into two (triangular) areas, namely $S_L(A)$ and $S_R(A)$, expressing such area on left hand side and that on right hand side of normal point, respectively;
- 7: Calculate RF of composite fuzziness of observed fuzzy sets $S(A_1^*, A_2^*, \dots, A_m^*)$ over composite fuzziness of n weighted neighbouring antecedent fuzzy sets for each left and right area, such that

$$RF = \frac{S(A_1^*, A_2^*, \dots, A_m^*)}{\sum_{i=1}^n W_i S(A_1^i, A_2^i, \dots, A_m^i)}$$

$$S(A_1^*, A_2^*, \dots, A_m^*) = \sum_{j=1}^m S(A_j^*) = \sum_{j=1}^m (S_L(A_j^*) + S_R(A_j^*))$$

$$S(A_1^i, A_2^i, \dots, A_m^i) = \sum_{j=1}^m S(A_j^i) = \sum_{j=1}^m (S_L(A_j^i) + S_R(A_j^i))$$

- 8: Calculate characteristic points of interpolated consequent fuzzy set $B^* = (b_1^*, b_2^*, b_3^*)$, such that

$$b_2^* = b'_2$$

$$b_1^* = b_2^* - RF \times |b'_2 - b'_1|$$

$$b_3^* = b_2^* + RF \times |b'_3 - b'_2|$$

- 9: **Return** Interpolated consequent $B^* = (b_1^*, b_2^*, b_3^*)$
-

4.1.2.3 Remarks

1. The GA-based weight learning scheme requires many predefined parameters, such as fitness function and the maximum iteration number.
2. In the evolutionary learning process, the updating of weights requires repeated runs of weighted FRI to compute the consequent using the current weights, in order to evaluate their fitness. This means the weight learning process is affected by the implementation of the underlying FRI process itself.
3. The individual weights of rule antecedent attributes are only involved in the aggregation to obtain the rule weights, as shown in Line 3. They are not integrated with the underlying FRI.

4.1.3 Piecewise fuzzy entropies-based weighted FRI (CC2016)

In Chen and Chen (2016), another method for weighted fuzzy interpolative reasoning is proposed through the exploitation of the concept of piecewise fuzzy entropies. This is referred to as CC2016 hereafter, which can handle fuzzy sets defined by polygonal and bell-shaped membership functions. In this method, weights are assigned differently to each antecedent variable when dealing with the same variable that is involved in different rules used for interpolation.

4.1.3.1 Learning weights In CC2016 a fuzzy rule is generally represented in the following:

$$r^i : \text{if } a_1 \text{ is } A_1^i (AW_1^i) \text{ and } a_2 \text{ is } A_2^i (AW_2^i) \text{ and } \dots \text{ and } a_m \text{ is } A_m^i (AW_m^i), \text{ then } z \text{ is } B^i$$

where AW_j^i stands for the weight for j th antecedent variable in the rule r^i . As indicated above, for a certain antecedent variable, its weight is allowed to be different when it is involved in different fuzzy rules. Such weights of rule antecedent attributes are generated during the weighted FRI process itself, which is explained next.

4.1.3.2 Weighting FRI The fuzzy sets in this work are assumed to be polygonal, which are represented by their characteristic points (CPs), paired with the corresponding membership degrees. Let A be a polygonal fuzzy set in the universe of discourse and the number of CPs for characterising A be n , then

$$A = (a_0, a_1, \dots, a_l, a_c, a_r, \dots, a_{n-1}; \mu_0, \mu_1, \dots, \mu_{n-1})$$

where a_l and a_r are called the “left normal point” and the “right normal point”, and $a_c = \frac{a_l + a_r}{2}$ the “central point”, with $\mu_0 = \mu_{n-1} = 0$, $\mu_l = \mu_r = 1$.

The basic idea is to construct an interpolated consequent fuzzy set B^* with regard to an input observation, by estimating its n CPs and the corresponding membership values such that

$$B^* = (b_0^*, b_1^*, \dots, b_{n-1}^*; \mu_0^*, \mu_1^*, \dots, \mu_{n-1}^*)$$

The key step to perform the estimation of the membership degrees is carried out through computing the piecewise fuzzy entropies of the fuzzy sets involved. The concept of piecewise fuzzy entropy is defined via the notion of non-probability fuzzy entropy of a fuzzy set (see Al-Sharhan et al. 2001; De Luca and Termini 1972). In particular, the piecewise fuzzy entropy $H_{t-1,t}(A)$ between the $(t-1)$ th CP and the t th CP of a polygonal fuzzy set A is specified as below:

$$H_{t-1,t}(A) = -K \sum_{s=t-1}^t [\mu_s \log_{10}(\mu_s) + (1 - \mu_s) \log_{10}(1 - \mu_s)] \quad (43)$$

where $K = 1/n$, $1 \leq t \leq n-1$, and μ_s denotes the degree of membership of the characteristic point a_s .

The weighted interpolative approach is summarised in Algorithm 3, with the individual weights highlighted in bold where they are learned and used.

Algorithm 3 Weighted FRI in CC2016

Input: — Rule base $R = \{r^1, \dots, r^N\}$ consisting of N rules in form of Eqn. (1)

— Observation $o^* = \{A_1^*, \dots, A_m^*\}$, over m conditional attributes

Output: — Interpolated consequent B^*

— **Estimation of CPs of B^***

- 1: Calculate representative values for antecedent fuzzy sets A_j , $j = 1, \dots, m$ of rules r^i , $i = 1, \dots, N$ and observation o^* , namely $Rep(A_j^i)$ and $Rep(A_j^*)$;
- 2: Calculate relative placement factor γ_j for A_j^* with respect to antecedent fuzzy sets of its neighbouring rules using their corresponding Rep values, and obtain mean value of these:

$$\gamma = \frac{\sum_{j=1}^m \gamma_j}{m}, \quad j = 1, 2, \dots, m$$

and

$$\gamma_j = \frac{Rep(A_j^*) - Rep(A_{lj})}{Rep(A_{rj}) - Rep(A_{lj})}$$

where A_{lj} and A_{rj} are left and right closest fuzzy set of A_j^* , respectively, in ascending order of A_j^i , $i = 1, \dots, N, *$.

- 3: Calculate CPs of B^* such that

$$b_t^* = (1 - \gamma) \times b_t^l + \gamma \times b_t^r$$

where b_t^* , b_t^l , b_t^r are t th CP of B^* , B^l , B^r , respectively, $1 \leq t \leq n-1$, and n denotes number of CPs used for characterising a polygonal fuzzy set, $l = \lfloor \frac{(n-1)}{2} \rfloor$ and $r = \lceil \frac{(n-1)}{2} \rceil$.

— **Estimation of membership values with respect to CPs of B^***

- 4: Calculate piecewise fuzzy entropy between $(t-1)$ th CP and t th CP for each A_j^i , A_j^* and B^i following Eqn. (43), namely $H_{t-1,t}(A_j^i)$, $H_{t-1,t}(A_j^*)$ and $H_{t-1,t}(B^i)$ where $i = 1, 2, \dots, N, j = 1, 2, \dots, m$ and $1 \leq t \leq n-1$.
- 5: Calculate central point $a_{j,c}^i$, $a_{j,c}^*$, $b_{j,c}$ of each A_j^i , A_j^* and B^i , respectively where $i = 1, 2, \dots, N, j = 1, 2, \dots, m$.
- 6: Calculate **weight \mathbf{AW}_j^i** for j th antecedent variable in rule r^i as follows:

$$\mathbf{AW}_j^i = 1 - \frac{d(A_j^*, A_j^i)}{\max_{1 \leq h, k \leq N} d(A_j^h, A_j^k)}$$

where $d(A_j^*, A_j^i)$ and $d(A_j^h, A_j^k)$ are defined by

$$d(A_j^*, A_j^i) = |a_{j,c}^* - a_{j,c}^i|, \quad d(A_j^h, A_j^k) = |a_{j,c}^h - a_{j,c}^k|$$

with $1 \leq j \leq m, i, h, k \in 1, 2, \dots, N$.

- 7: Calculate weight for each rule W_i , $i = 1, \dots, N$, by aggregating **antecedent weights \mathbf{AW}_j^i** , $j = 1, 2, \dots, m$ such that

$$W_i = \frac{\min_{j=1, \dots, m} \mathbf{AW}_j^i}{\sum_{i=1}^N \min_{j=1, \dots, m} \mathbf{AW}_j^i}$$

- 8: Calculate piecewise fuzzy entropy $H_{t-1,t}(B^*)$ between $(t-1)$ th CP and t th CP of B^* :

$$H_{t-1,t}(B^*) = \begin{cases} \sum_{j=1}^m \left\{ H_{t-1,t}(A_j^*) \times \left[\sum_{i=1}^N W_i \times \frac{H_{t-1,t}(B^i)}{\sum_{j=1}^m H_{t-1,t}(A_j^i)} \right] \right\} & \text{if } \exists i, j, H_{t-1,t}(A_j^i) > 0 \\ \sum_{j=1}^m H_{t-1,t}(A_j^*) & \text{if } \forall i, j, H_{t-1,t}(A_j^i) = 0 \end{cases}$$

where $1 \leq t \leq n-1$.

- 9: Calculate membership degrees $\mu_0^*, \mu_1^*, \dots, \mu_{n-1}^*$ with regards CPs of interpolated result B^* , using piecewise fuzzy entropy $H_{t-1,t}(B^*)$, $1 \leq t \leq n-1$ as per method of Chapra and Canale (1998) (more details of which can be found in Chen and Chen (2016)).
- 10: Obtain interpolated fuzzy set B^* in terms of CPs (from Line 3) and corresponding membership values (from Line 9), resulting in

$$B^* = (b_0^*, b_1^*, \dots, b_{n-1}^*; \mu_0^*, \mu_1^*, \dots, \mu_{n-1}^*)$$

- 11: **Return** Interpolated consequent B^*
-

4.1.3.3 Remarks

1. As indicated above, in this method, the weights for individual rule antecedent variables are assigned differently when different rules involving them are taken into consideration.
2. The generation of the antecedent attribute weights is achieved during the weighted FRI process, as shown in Line 5 of the algorithm.
3. The individual weights of rule antecedent attributes are only involved in the aggregation stage to obtain the overall rule weights, as shown in Line 6. Unfortunately, such useful information is not integrated within the rest of the fuzzy interpolative reasoning process.

4.1.4 Weighted increment and ratio transformation-based weighted FRI (CKCP2009)

Another weighted FRI method is presented in Chen et al. (2009), referred to as CKCP2009 hereafter. It uses weighted increment transformation and weighted ratio transformation to enable weighted fuzzy interpolative reasoning. A “wrapper” algorithm is implemented for automatically tuning the optimal weights of the antecedent variables appearing in a fuzzy rule, capable of dealing with polygonal, Gaussian and bell-shaped membership functions.

4.1.4.1 Learning weights The weights on individual rule antecedent variables are automatically learned with a “wrapper” mechanism. The weighted interpolation process is required to be iteratively triggered in order to update the current weights. Particularly, the weight learning procedure within the proposed weighted FRI is tailored for a certain system control problem, where one input may lead to several states indicating the current values of the observation. The weight learning process is summarised below.

Individual weights are initialised with the same value to start the first iteration. A set of training samples as rule antecedent attribute values are then employed as the input to the FRI system, together with the current weights, resulting in the next states of these variables. To adjust the weighting of each rule antecedent attribute, the gradient-descent training method is utilised, where a predefined fitness function over the rule antecedent variable is evaluated using the value recorded in its final state. The fitness function generates the prediction error for each antecedent variable in the current iteration and the weights are then modified with the aim to minimise the error, which are subsequently employed to run the next iteration. The entire iterative weight updating process is terminated when a preset maximum number of iterations is reached.

4.1.4.2 Weighting FRI As with many FRI methods reviewed previously using *Rep* values, a unique real value is also defined herein and associated with a certain fuzzy set for reflecting the key information on the overall location in its domain. In CKCP2009, for a polygonal fuzzy set $A = (a_1, a_2, \dots, a_{n-1})$, the characteristic value $CV(A)$ (or *Rep* as termed elsewhere) is defined as follows:

$$CV(A) = \frac{a_0 + a_1 + \dots + a_{n-1}}{n} \quad (44)$$

The distance between two polygonal fuzzy sets P and Q is then specified by the use of their CV values such that

$$d(P, Q) = |CV(P) - CV(Q)| \quad (45)$$

Given these notions, the weighted FRI method can be summarised as shown in Algorithm 4 (Chen et al. 2009), where the multiplication operation between a polygonal fuzzy set A and a real value w ($w \in [0, 1]$) is defined by

$$A \otimes w = (a_1, a_2, \dots, a_{n-1}) \otimes w = (wa_1, wa_2, \dots, wa_{n-1}) \quad (46)$$

As with the other weighted FRI approaches, the individual weights are highlighted in bold within the algorithm description.

Algorithm 4 Weighted FRI in CKCP2009

Input: – Rules $R = \{r^1, r^2\}$ in form of Eqn. (39)

– Observation $o^* = \{A_1^*, \dots, A_m^*\}$, over m conditional attributes

Output: – Interpolated consequent B^*

1: Initialise individual weights of rule antecedent variable $AW_j^i, i = 1, 2, j = 1, 2, \dots, m$;

2: Construct fuzzy set $A'_j = (a'_{j1}, a'_{j2}, a'_{j3})$ from fuzzy sets A_j^1 and A_j^2 such that
Let

$$a'_{j1} = a_{j2}^* - la'_{j,1-2}$$

$$a'_{j2} = a_{j2}^*$$

$$a'_{j3} = a_{j2}^* + la'_{j,2-3}$$

where

$$la'_{j,1-2} = (1 - \lambda_{rep}^j)la_{j,1-2}^1 + \lambda_{rep}^jla_{j,1-2}^2$$

$$la'_{j,2-3} = (1 - \lambda_{rep}^j)la_{j,2-3}^1 + \lambda_{rep}^jla_{j,2-3}^2$$

$$la_{j,1-2}^i = d(a_{j1}^i, a_{j2}^i) \quad la_{j,2-3}^i = d(a_{j2}^i, a_{j3}^i) \quad i = 1, 2. \quad j = 1, 2, \dots, m$$

and

$$\lambda_{rep}^j = \frac{d(A_j^1 \otimes \mathbf{AW}_j^1, A_j^* \otimes (\mathbf{AW}_j^1 + \mathbf{AW}_j^2)/2)}{d(A_j^1 \otimes \mathbf{AW}_j^1, A_j^2 \otimes \mathbf{AW}_j^2)}$$

3: Calculate normal point (i.e. b_2^*) of interpolated fuzzy set B^* such that

$$b_2^* = (1 - \lambda)CV(B^1) + \lambda CV(B^2)$$

where

$$\lambda = \sum_{j=1}^m \frac{\mathbf{AW}_j^1 + \mathbf{AW}_j^2}{W} \lambda_j \quad W = \sum_{j=1}^m (\mathbf{AW}_j^1 + \mathbf{AW}_j^2)$$

$$\lambda_j = \frac{d(A_j^1, a_j^*)}{d(A_j^1, A_j^2)} = \frac{|CV(A_j^1) - a_j^*|}{|CV(A_j^1) - CV(A_j^2)|}$$

4: Construct fuzzy set $B' = (b'_1, b'_2, b'_3)$ based on fuzzy sets B^1 and B^2 as with that shown in Line 2, such that

$$b'_1 = b_2^* - lb'_{1-2}$$

$$b'_2 = b_2^*$$

$$b'_3 = b_2^* + lb'_{2-3}$$

where

$$lb'_{1-2} = (1 - \lambda_{rep})lb_{1-2}^1 + \lambda_{rep}lb_{1-2}^2$$

$$lb'_{2-3} = (1 - \lambda_{rep})lb_{2-3}^1 + \lambda_{rep}lb_{2-3}^2$$

with

$$lb_{1-2}^i = d(b_1^i, b_2^i) \quad lb_{2-3}^i = d(b_2^i, b_3^i) \quad i = 1, 2$$

$$\lambda_{rep} = \sum_{j=1}^m \left(\frac{\mathbf{AW}_j^1 + \mathbf{AW}_j^2}{\sum_{j=1}^m (\mathbf{AW}_j^1 + \mathbf{AW}_j^2)} \right) \lambda_{rep}^j$$

5: Calculate extreme points (namely b_1^* and b_3^*) of interpolated fuzzy set B^* as follows: Given artificially constructed antecedent and consequent fuzzy sets $A'_j, j = 1, 2, \dots, m$ and B' , by aggregating corresponding components within two given rules, analogical reasoning is performed to derive remaining characteristic points of interpolated result B^* , i.e., b_1^* and b_3^* . This is achieved through two sub-routines, namely increment and ratio transformations, which are implemented by assessing similarity between each of A'_j and A_j^* first and then, applying similarity measure to B' in order to obtain B^* . **Antecedent weights** \mathbf{AW} are computed as shown in Line 3 (more computational details are omitted but can be found in Chen et al. (2009));

6: **Return** Interpolated consequent $B^* = (b_1^*, b_2^*, b_3^*)$

4.1.4.3 Remarks

1. This weight learning scheme is an iterative process. It is integrated within the weighted interpolation procedure, of which the outcome is required to be collected to evaluate the fitness function to update the current weights.
2. Although the weights are designed to be automatically tuned for optimisation, the approach is tailored to a specific problem, where the fitness functions for each rule antecedent attribute are predefined. This limits the generality of the underlying techniques.
3. This algorithm reflects the intuition in approximate reasoning in that “how an observation is transformed from an intermediate antecedent fuzzy set should be reflected in how the interpolated outcome is transformed from the intermediate consequent”. This is basically the same as the idea adopted by the conventional T-FRI.

4.1.5 Attribute weighted scale and move transformation-based FRI (LSLYS2018)

A weighted extension of T-FRI, the scale and move transformation-based FRI method (Huang and Shen 2006, 2008), is presented in Li et al. (2018b), which is referred to as LSLYS2018 hereafter. It enables the weights learn from the given sparse rule base only to automatically determine the relative importance of rule antecedent attributes. Also, the obtained attribute weights are thoroughly applied within each core step of T-FRI.

4.1.5.1 Learning weights Unlike many weight learning schemes reported earlier, in LSLYS2018, the individual weights of rule antecedent attributes are automatically learned by the use of the given sparse rule base only, without acquisition of any observations nor that of any other human intervention. This attribute weighting scheme is enabled by an innovative *Reverse Engineering* procedure, by generating an artificial training decision table from the given sparse rule base. The essential idea is to reformulate all rules in the rule base into a common representation, where each (possibly) missing value of any rule antecedent is replaced by one of the alternative fuzzy values from its domain. All these reformulated rules, artificial or original, are collated for evaluation of the relative significance degrees of the individual attributes.

The weights of the attributes are individually measured using a certain feature ranking method, which is implemented by modifying the feature evaluation procedure extracted from a selected feature selection (FS) technique. Five approaches are reported for such usage, including the individual feature ranking-based method [namely, Information Gain (IG), Relief-F, Laplacian Score (LS) and Local Learning-based Clustering for FS (LLCFS)] and feature subset evaluation-based method [rough set-based FS (RSFS)]. References for these FS methods can be found in Li et al. (2018b). They may be employed equally but one and just one of the five types of the weighting scheme is required at a time for implementing the weighted T-FRI. The resultant learned weights are uniquely associated with each of rule antecedent variables, no matter which rule it is involved in.

4.1.5.2 Weighting FRI The weighted T-FRI algorithm presented in LSLYS2018 is summarised in Algorithm 5. In particular, the individual attribute weights are integrated with every procedure of the underlying non-weighted algorithm. Hence, there are a total of three procedures that involve such integration: the selection of the nearest neighbouring

rules, the construction of intermediate rules, and the computation of scale and move transformation factors. All computational steps in the original T-FRI, which effectively deals with evenly calculated average of the attribute values, are improved by a weighted aggregation of the corresponding components (as highlighted in bold in Algorithm 5). As a weighted extension to the conventional T-FRI that is described in Sect. 3.3.2, the general rule interpolation process of this extended algorithm remains the same as its original, where the relative computations can be referred to.

Algorithm 5 Weighted FRI in LSLYS2018

Input: – Sparse rule base $R = \{r^1, \dots, r^N\}$ consisting of N rules in form of Eqn. (1)

– Observation $o^* = \{A_1^*, \dots, A_m^*\}$, over m conditional attributes

– Number of closest rules n

– Attribute weights $AW = (AW_1, \dots, AW_m)$

Output: – Interpolated consequent B^*

– **Step 1: Closest Rules Selection Guided by Attribute Weights:**

1: **for** $i = 1$ to $i = N$ **do**

2: Calculate weighted distance $\tilde{d}(o^*, r^i, AW)$ between o^* and r^i , such that

$$\tilde{d}(o^*, r^i, AW) = \frac{1}{\sqrt{\sum_{t=1}^m (1 - AW_t)^2}} \sqrt{\sum_{j=1}^m \left((1 - AW_j) d(A_j^i, A_j^*) \right)^2}$$

3: **end for**

4: Select n rules of shortest distance(s);

– **Step 2: Intermediate Rule (r') Construction Guided by Attribute Weights:**

5: Obtain weights $w_j^i, i = 1, \dots, n, j = 1, \dots, m$, as computed by original T-FRI to j th antecedent attribute of i th selected rule;

6: Compute antecedent attribute values of intermediate rule $A_j', j = 1, 2, \dots, m$, by linearly aggregating corresponding weighted antecedent values over selected n rules using normalised weights \hat{w}_j^i ;

7: Calculate weight \tilde{w}_z^i for each consequent per selected closest rule, by accumulating normalised weights contributed by w_j^i , such that

$$\tilde{w}_z^i = \sum_{j=1}^m AW_j w_j^i$$

8: Construct fuzzy set B' for consequent attribute of intermediate rule, by aggregating consequent values of n closest rules $B^i, i = 1, \dots, n$, which are respectively weighted by w_z^i ;

– **Step 3: Scale and Move Factors Calculation:**

9: **for** each rule antecedent attribute **do**

10: Obtain scale rate s_{A_j} that modifies A_j' into \hat{A}_j' such that it maintains same scale as corresponding component in o^* ;

11: Obtain move ratio m_{A_j} that modifies \hat{A}_j' for it to maintain same position as corresponding component in o^* ;

12: **end for**

– **Step 4: Scale and Move Transformations with Weighted Factors:**

13: Calculate overall transformation factors for B' to ensure analogy, by aggregating corresponding weighted scale and move factors, such that

$$\tilde{s}_z = \sum_{j=1}^m AW_j s_{A_j} \quad \tilde{m}_z = \sum_{j=1}^m AW_j m_{A_j}$$

14: Compute final interpolated outcome B^* by applying scale and move factors to B' , such that $B^* = T(B', \tilde{s}_z, \tilde{m}_z)$;

15: **Return** B^*

4.1.5.3 Remarks

1. The entire weight learning scheme exploits just the knowledge already available, i.e., the fuzzy rules in the given sparse rule base, without acquisition of any observations nor need for human intervention.
2. Such a learning method is independent from the underlying FRI mechanism. Note that in Table 8, two weighted FRI approaches reported in LSLYS19 (Li et al. 2020b) are the weighted extensions to the original non-weighted FRI [namely KH (Kóczy and Hirota 1993a; Wong et al. 2005) and CCL (Chang et al. 2008)]. Analogous to the weighted T-FRI herein, the weights of antecedent attributes are generated in the same manner, which are then integrated with the two conventional FRI methods throughout each computational procedure.
3. For the implementation of this weighted T-FRI method, no specification of which feature ranking mechanism to use is specified to generate the required weights. Indeed, any of the feature ranking methods available may be taken to assess the relative significance of individual antecedent attributes, offering flexibility in performing interpolative reasoning.
4. Individual attribute weights are integrated with every procedure of the conventional T-FRI. Nonetheless, this weighted T-FRI algorithm can easily degenerate to its original when all antecedent attributes are assumed to be of equal significance. This is because the weight of each attribute has been normalised over the ranking scores derived from a given feature ranking method, which results in an identical weight for each of rule antecedent being $AW_j = 1/m, j = 1, 2, \dots, m$ if all weights are assumed to be equal.

4.2 Comparison of weighted FRI methods

The above five weighted fuzzy interpolative reasoning mechanisms are typical approaches from the viewpoint of weight learning and FRI weighting. This subsection contrasts these approaches and categorises other weighted FRI methods listed in Table 8 in relation to the distinct features of these five approaches.

4.2.1 Weight learning mechanisms

As reflected by the preceding subsections, typical approaches to weighted fuzzy interpolative reasoning contain basic properties along with which other weighted FRI methods can be grouped and compared.

4.2.1.1 Predefined versus automatically learned The initial idea for obtaining weights on rule antecedent attributes is simply to predefine them with domain expertise directly acquired from the experts. This approach includes the early work as reported in LHTZ2005 and CC2008. It requires human intervention and hence, adversely reduces the flexibility of the resulting fuzzy systems. Automated weight learning schemes are obviously preferred. Indeed, all of the remaining methods in Table 8 pursue alternative ways to learn weights automatically.

4.2.1.2 Unique weight versus multiple weights for an antecedent attribute In general, weighted FRI works with fuzzy rule bases that involve multiple rule antecedent variables. Different significance levels are associated with different variables to indicate their different contributions towards the conclusion. In the literature, for a given rule antecedent attribute, certain methods learn a unique weight for each variable independent of what rules that variable appears in, whilst others assign different weights to one common attribute in different rules. The former includes work in CC2011a, CC2008, CH2014, DJS2014, CC2011b, LSLYS18 and LSLYS19, and the latter includes LHTZ2005, CC2016, CKCP2009 and CA2018. When a rule antecedent attribute may be assigned with multiple weights, depending upon which rules it may appear in, the overall rule model becomes more complicated and harder to interpret. Moreover, more specific information regarding the antecedent variables of observations may become necessary, in order to compute the characteristic points of the corresponding fuzzy sets. This may include for example, information on central points (Chen and Chen 2016) or that on ranking values (Cheng et al. 2015) of the fuzzy sets, thereby at the expense of involving more computation to produce the weights than otherwise. Besides, in so doing, the weights are only measurable during the running of the weighted FRI system when an observation is provided.

4.2.1.3 Filter schemes versus wrapper schemes The terms *filter* and *wrapper* are used to group the weight learning schemes, based on their dependence upon whether a weighted FRI method will be recursively called on during the process of weight generation. That is, those weight learning methods following the filter scheme are independent of the weighted FRI process, whereas the wrapper methods need to exploit the outcome of the weighted FRI in order to evaluate the “goodness” or quality of the current weights. The filter approach is taken by CC2016, DJS2014, CA2018, LSLYS18 and LSLYS19, and the wrapper approach by CC2011a, CKCP2009, CH2014 and CC2011b. Since methods belonging to the wrapper group employ interpolated results for constructing the fitness functions (in an effort to update the required weights in the current iteration), their performance in terms of accuracy may be very high, but the computational overheads is relatively costly at the same time.

4.2.2 Weighting FRI procedures

This issue is concerned with how the generated weights of rule antecedent attributes are integrated within the underlying FRI, for revealing the relative significance level of each individual attribute in contributing to the derivation of the interpolated results. As can be seen from the typical weighted FRI mechanisms reviewed previously, the following observations can be drawn:

1. Most existing techniques generally work by artificially creating an overall weight to each of the rules before running the weighted rules in FRI. Such weights are normally computed through aggregating the weights calculated for individual rule antecedent variables, thereby involving additional weight aggregation procedures. Weighted FRI approaches are the most recent developments in the literature. Established examples include: weighted T-FRI, weighted KH and weighted CCL (see Li et al. 2018b, 2020b for details), all of which exploit the individual antecedent weights to improve the original unweighted methods, signifying the importance of each attribute in influencing the conclusion given an unmatched observation.

2. Learned weights are seldom systematically integrated within all major components of the weighted FRI algorithm, but just involved in certain computational subroutines. As such, information regarding domain attribute significance is not exploited to its full potential. Fortunately, the recent developed weighted FRI offers a possibility of a general weighting scheme that enables different unweighted FRI methods to be supported with antecedent weights in a common manner. In so doing, it helps facilitate transplanting a developed weighting scheme from one FRI mechanism to another once the weights of rule antecedent attributes are available.

4.3 Weighted versus non-weighted FRI

In Sect. 3 and the above of the present section, the conventional (unweighted or flat) FRI techniques and the recently advanced approaches for weighted FRI have been systematically reviewed, and the relative pros and cons of the individual methods have been pointed out within each of these two categories. Detailed, quantitative comparisons between them are beyond the scope of this review, but such results can be found in the relevant references that report the advancements of individual FRI methods (e.g., Li et al. 2018b, 2020b). Nonetheless, it is helpful in completing this review of the existing weighted FRI mechanisms to qualitatively contrast the performances between these two categories of approaches.

As indicated in Sect. 4.1.5, three major approaches to weighted FRI have been proposed in the literature, each being a weighted extension to their original (namely, T-FRI, KH and CCL), by computing and taking into account of the rule antecedent attribute weights. The resultant weighted T-FRI, weighted KH and weighted CCL have been comprehensively compared against their corresponding non-weighted originals, when applied to solving a wide range of benchmark classification problems, from the viewpoints of both effectiveness and efficiency. First of all, all empirical results available in the literature have demonstrated that the weighted fuzzy interpolative reasoning mechanisms outperform their non-weighted FRI counterparts, in terms of the accuracy of interpolated results and hence, of the effectiveness of the corresponding FRI methods, especially when the rule bases are considerably sparse.

As the weights are integrated within the computational progress of FRI, the time complexity of a weighted algorithm may become a natural concern. Thus, investigations have also been carried out to reveal how much extra computation effort may be required by running a weighted approach. Such a comparison will only make sense if it is conducted between those weighted methods and their underlying non-weighted originals, just as with the case when their relative performances in terms of accuracy have been assessed. Importantly, note that the attribute weights can also be exploited to help modify the selection of the nearest neighbouring rules (as shown in Algorithm 5), which are to be utilised to implement weighted rule interpolation. Comparative experimental investigations have therefore, been accomplished in the literature with regard to various cases where different numbers of closest rules (in the range of 2-6) are selected to conduct interpolation. The results of running time have convincingly shown that there is no significant increase in the time cost by a weighted FRI as compared to that by its original where no weights are involved, while using the same number of rules for any of three FRI methods. This positively differs from the initial expectation due to the first glance at the extended algorithms.

Whilst there is indeed an increase in time consumption (though generally rather small) when exploiting more closest rules for all FRI methods, which is independent of using a

weighted FRI method or not, a more important conclusion has been drawn from observing the change of classification accuracy by varying the number of neighbouring rules being used. That is, in general, the weighted FRI methods (no matter which one) only require the least number (i.e., 2) of the nearest neighbouring rules to be taken to perform interpolation while achieving the best performance. This finding significantly enhances the algorithm efficiency through avoiding the involvement of more rules in the implementation of the interpolative process.

Finally, it is worth pointing out that, in terms of the criteria commonly adopted in the literature to evaluate the FRI methods (as per the discussions of Sect. 3.4), weighted approaches naturally inherit those properties possessed by their respective originals. However, it should be noted that there exists limited work on the development of weighted fuzzy interpolative reasoning schemes which differ from what is focussed on in this review. For example, a particular piece of research as reported in Chen et al. (2013b) constructs a weighted FRI method based on interval type-2 fuzzy sets. Such work however, involves higher-order representation and hence, substantially more complex computation than those reviewed above. Being significantly more sophisticated in their underlying mathematical representation and computational implementation, and empirically less studied so far, their details are omitted in this paper.

5 Conclusion

This survey has reviewed a range of important techniques for approximate reasoning that work with incomplete and imprecise knowledge. It has demonstrated that fuzzy rule interpolation (FRI) is able to perform approximate inference when traditional rule-based methodologies fail, for situations where no existing rules in a given sparse rule base match a novel observation. The paper has reviewed the general FRI methodologies, highlighting the strengths and limitations of classical approaches. It has also comprehensively analysed a family of most recent FRI algorithms that successfully address the common and important problem shared by many conventional FRI mechanisms, where all antecedent attributes are forcefully assumed to be of equal significance. In particular, this survey has compared different FRI techniques, qualitatively outlining the advantages of running a certain FRI method. This work therefore, enables the readers to have an informed choice of what may be the potentially suitable FRI technique(s) to apply given their specific domain problems.

Whilst it has been shown that the FRI techniques in general and the weighted FRI schemes in particular are able to strengthen the power of approximate reasoning, there is much room for further improvement. As such, this survey also helps provide a number of insightful suggestions for future research in this important area. For instance, the current weighted FRI works on a static rule base. Yet, a volume of intermediate fuzzy rules are typically generated while executing the transformation-based rule interpolation. From this, the ideas of a dynamic FRI can be exploited to enrich the rule base by refining and promoting these intermediate rules, gaining efficiency by allowing for more direct rule-firing without running the interpolation procedure. Also, all mainstream work carried out so far in the FRI literature has to do with fuzzy rules of Mamdani type. Most recently, there has been research which reports on extending conventional T-FRI methods to building FRI mechanisms for Takagi Sugeno Kang (TSK) fuzzy models. These initial attempts all follow the conventional unweighted approach. It would therefore, be very interesting to consider

further extending such work within the weighted FRI framework. A natural start point for this would be to directly mirror the weight-learning techniques developed within the Mamdani model-based T-FRI to estimate the antecedent weights of the otherwise unweighted conditional attributes in a TSK model. How the resulting weights could be integrated with TSK model-based FRI requires much investigation.

Furthermore, the curse of dimensionality remains a challenging problem in all fuzzy rule-based reasoning systems, including those supported with an FRI procedure. The performance of an FRI system might be strengthened by increasing the number of input variables and that of the linguistic values which each input may take. However, this would inevitably result in the enormously exponential raise of the number of the fuzzy rules required, thereby impractically increasing the system complexity. Even ignoring the issue of computational complexity, the original, fundamental problem that FRI faces is the lack of fuzzy rules in the first place, how to acquire a much large rule base is itself a significant challenge. Fuzzy hierarchical inference models may provide an effective way to alleviate such difficulties, by introducing multiple low-dimensional input fuzzy sub-systems. Thus, another important future study is to investigate how a hierarchical fuzzy model may be integrated with a given FRI working procedure.

Acknowledgements This work is partly supported by the Strategic Partner Acceleration Award (80761-AU201) under the Sêr Cymru II programme, UK.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Al-Sharhan S, Karray F, Gueaieb W, Basir O (2001) Fuzzy entropy: a brief survey. In: 10th IEEE international conference on fuzzy systems (Cat. No. 01CH37297), vol 3. IEEE, pp 1135–1139
- Baranyi P, Kóczy L (1996a) Multidimensional fuzzy rule inter-and extrapolation based on geometric solution. In: Proceedings of the 7th international power electronics and motion control conference (PEMC'96), vol 3. pp 443–447
- Baranyi PZ, Kóczy L (1996b) A general and specialised solid cutting method for fuzzy rule interpolation. Journal BUSEFAL, URA-CNRS, Université Paul Sabatier, pp 13–22
- Baranyi P, Gedeon T, Kóczy L (1995) A general method for fuzzy rule interpolation: specialized for crisp triangular and trapezoidal rules. In: EUFIT'95. pp 99–102
- Baranyi P, Gedeon T, Kóczy L (1996a) Rule interpolation by spatial geometric representation. Proceedings of IPMU 96:483–488
- Baranyi P, Gedeon TD, Kóczy LT (1996b) A general interpolation technique in fuzzy rule bases with arbitrary membership functions. In: 1996 IEEE international conference on systems, man and cybernetics. Information intelligence and systems (Cat. No. 96CH35929), vol 1. IEEE, pp 510–515
- Baranyi P, Mizik S, Kóczy LT, Gedeon TD, Nagy I (1998) Fuzzy rule base interpolation based on semantic revision. In: SMC'98 Conference proceedings. 1998 IEEE international conference on systems, man, and cybernetics (Cat. No. 98CH36218), vol 2. IEEE, pp 1306–1311
- Baranyi PZ, Tikk D, Yam Y, Kóczy L, Nádaí L, (1999) A new method for avoiding abnormal conclusion for a-cut based rule interpolation. In: 1999 IEEE international fuzzy systems conference, FUZZ-IEEE '99. Seoul, IEEE, pp 1383–1388

- Baranyi P, Kóczy L, Gedeon T (2004) A generalized concept for fuzzy rule interpolation. *IEEE Trans Fuzzy Syst* 12(6):820–837
- Bede B (2013) *Mathematics of fuzzy sets and fuzzy logic*. Springer, Berlin Heidelberg
- Berecz A (2009) Fuzzy rule interpolation based tool life modeling using RBE-SI and FRIPOC. In: 2009 5th international symposium on applied computational intelligence and informatics. IEEE, pp 11–16
- Bouchon-Meunier B, Valverde L (1999) A fuzzy approach to analogical reasoning. *Soft Comput* 3(3):141–147
- Bouchon-Meunier B, Delechamp J, Marsala C, Mellouli N, Rifqi M, Zerrouki L (1999) Analogy and fuzzy interpolation in the case of sparse rules. In: *Proceedings of joint conference eurofuse-soft and intelligent computing*, pp 132–136
- Bouchon-Meunier B, Marsala C, Rifqi M (2000) Interpolative reasoning based on graduality. In: Ninth IEEE international conference on fuzzy systems. FUZZ-IEEE 2000 (Cat. No. 00CH37063), vol 1. IEEE, pp 483–487
- Bouchon-Meunier B, Dubois D, Marsala C, Prade H, Ughetto L (2001) A comparative view of interpolation methods between sparse fuzzy rules. In: *Proceedings joint 9th IFSA world congress and 20th NAFIPS international conference* (Cat. No. 01TH8569). IEEE, pp 2499–2504
- Burkhardt DG, Bonissone PP (1992) Automated fuzzy knowledge base generation and tuning. In: [1992 Proceedings] IEEE international conference on fuzzy systems. IEEE, pp 179–188
- Chang YC, Chen SM (2008) A new method for multiple fuzzy rules interpolation with weighted antecedent variables. In: 2008 IEEE international conference on systems, man and cybernetics. IEEE, pp 76–81
- Chang YC, Chen SM, Liao CJ (2008) Fuzzy interpolative reasoning for sparse fuzzy-rule-based systems based on the areas of fuzzy sets. *IEEE Trans Fuzzy Syst* 16(5):1285–1301
- Chapra SC, Canale RP (1998) *Numerical methods for engineers*, vol 2. McGraw-Hill, New York
- Chen SM, Adam SI (2017) Adaptive fuzzy interpolation based on general representative values of polygonal fuzzy sets and the shift and modification techniques. *Inf Sci* 414:147–157
- Chen SM, Adam SI (2018) Weighted fuzzy interpolated reasoning based on ranking values of polygonal fuzzy sets and new scale and move transformation techniques. *Inf Sci* 435:184–202
- Chen SM, Chang YC (2011a) Weighted fuzzy interpolative reasoning for sparse fuzzy rule-based systems. *Expert Syst Appl* 38(8):9564–9572
- Chen SM, Chang YC (2011b) Weighted fuzzy rule interpolation based on GA-based weight-learning techniques. *IEEE Trans Fuzzy Syst* 19(4):729–744
- Chen SM, Chen ZJ (2016) Weighted fuzzy interpolative reasoning for sparse fuzzy rule-based systems based on piecewise fuzzy entropies of fuzzy sets. *Inf Sci* 329:503–523
- Chen SM, Hsin WC (2014) Weighted fuzzy interpolative reasoning based on the slopes of fuzzy sets and particle swarm optimization techniques. *IEEE Trans Cybern* 45(7):1250–1261
- Chen SM, Ko YK (2008) Fuzzy interpolative reasoning for sparse fuzzy rule-based systems based on α -cuts and transformations techniques. *IEEE Trans Fuzzy Syst* 16(6):1626–1648
- Chen SM, Lee LW (2011) Fuzzy interpolative reasoning for sparse fuzzy rule-based systems based on interval type-2 fuzzy sets. *Expert Syst Appl* 38(8):9947–9957
- Chen C, Shen Q (2017) Transformation-based fuzzy rule interpolation using interval type-2 fuzzy sets. *Algorithms* 10(3):91
- Chen SM, Ko YK, Chang YC, Pan JS (2009) Weighted fuzzy interpolative reasoning based on weighted increment transformation and weighted ratio transformation techniques. *IEEE Trans Fuzzy Syst* 17(6):1412–1427
- Chen SM, Chang YC, Chen ZJ, Chen CL (2013a) Multiple fuzzy rules interpolation with weighted antecedent variables in sparse fuzzy rule-based systems. *Int J Pattern Recognit Artif Intell* 27(05):1359002
- Chen SM, Lee LW, Shen VR (2013b) Weighted fuzzy interpolative reasoning systems based on interval type-2 fuzzy sets. *Inf Sci* 248:15–30
- Chen SM, Cheng SH, Chen ZJ (2015) Fuzzy interpolative reasoning based on the ratio of fuzziness of rough-fuzzy sets. *Inf Sci* 299:394–411
- Chen C, Parthaláin NM, Li Y, Price C, Quek C, Shen Q (2016) Rough-fuzzy rule interpolation. *Inf Sci* 351:1–17
- Chen T, Shang C, Yang J, Li F, Shen Q (2020) A new approach for transformation-based fuzzy rule interpolation. *IEEE Trans Fuzzy Syst* 28(12):3330–3344
- Cheng SH, Chen SM, Chen CL (2015) Fuzzy interpolative reasoning based on ranking values of polygonal fuzzy sets and automatically generated weights of fuzzy rules. *Inf Sci* 325:521–540
- Cheng SH, Chen SM, Chen CL (2016) Adaptive fuzzy interpolation based on ranking values of polygonal fuzzy sets and similarity measures between polygonal fuzzy sets. *Inf Sci* 342:176–190
- Das S, Chakraborty D, Koczy LT (2019) Linear fuzzy rule base interpolation using fuzzy geometry. *Int J Approx Reason* 112(SEP.):105–118

- De Luca A, Termini S (1972) A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory. *Inf Control* 20(4):301–312
- deKleer J (1986) An assumption-based TMS. *Artif Intell* 28(2):127–162
- Diao R, Jin S, Shen Q (2014) Antecedent selection in fuzzy rule interpolation using feature selection techniques. In: 2014 IEEE international conference on fuzzy systems (FUZZ-IEEE). IEEE, pp 2206–2213
- Ding L, Shen Z, Mukaidono M (1989) A new method for approximate reasoning. In: Proceedings. The nineteenth international symposium on multiple-valued logic. IEEE, pp 179–185
- Ding L, Shen Z, Mukaidono M (1992) Revision principle for approximate reasoning, based on linear revising method. In: Proceedings of the second international conference on fuzzy logic and neural networks, Iizuka'92. pp 305–308
- Eshelman LJ (1991) The CHC adaptive search algorithm: How to have safe search when engaging in non-traditional genetic recombination. In: Foundations of genetic algorithms, vol 1. Elsevier, pp 265–283
- Fu X, Shen Q (2010) Fuzzy compositional modeling. *IEEE Trans Fuzzy Syst* 18(4):823–840
- Holland J (1975) Adaptation in natural and artificial systems. MIT Press, Cambridge
- Hsiao WH, Chen SM, Lee CH (1998) A new interpolative reasoning method in sparse rule-based systems. *Fuzzy Sets Syst* 93(1):17–22
- Huang Z, Shen Q (2003) A new fuzzy interpolative reasoning method based on center of gravity. In: The 12th IEEE international conference on fuzzy systems, 2003. FUZZ'03, vol 1. IEEE, pp 25–30
- Huang Z, Shen Q (2006) Fuzzy interpolative reasoning via scale and move transformations. *IEEE Trans Fuzzy Syst* 14(2):340–359
- Huang Z, Shen Q (2008) Fuzzy interpolation and extrapolation: a practical approach. *IEEE Trans Fuzzy Syst* 16(1):13–28
- Huang DM, Tsang E, Yeung D (2004) A fuzzy interpolative reasoning method. In: Proceedings of 2004 international conference on machine learning and cybernetics (IEEE Cat. No. 04EX826), vol 3. IEEE, pp 1826–1830
- Jenei S (2001) Interpolation and extrapolation of fuzzy quantities revisited-an axiomatic approach. *Soft Comput* 5(3):179–193
- Jenei S, Klement EP, Konzel R (2002) Interpolation and extrapolation of fuzzy quantities-the multiple-dimensional case. *Soft Comput* 6(3–4):258–270
- Jin S, Diao R, Quek C, Shen Q (2014) Backward fuzzy rule interpolation. *IEEE Trans Fuzzy Syst* 22(6):1682–1698
- Jin S, Shen Q, Peng J (2019) Transformation based backward fuzzy rule interpolation with multiple missing antecedent values. In: Backward fuzzy rule interpolation. Springer, pp 75–89
- Jin S, Peng J, Li Z, Shen Q (2020) Backward fuzzy rule interpolation. *Inf Sci* 506:99–112
- Johanyák ZC (2010) Student evaluation based on fuzzy rule interpolation. *Int J Artif Intell* 5(A10):37–55
- Johanyák ZC, Kovács S (2006) Survey on various interpolation based fuzzy reasoning methods. *Prod Syst Inf Eng* 3(1):39–56
- Kawaguchi MF, Miyakoshi M (2000a) A fuzzy rule interpolation technique based on bi-splines in multiple input systems. In: Ninth IEEE international conference on fuzzy systems. FUZZ-IEEE 2000 (Cat. No. 00CH37063), vol 1. IEEE, pp 488–492
- Kawaguchi MF, Miyakoshi M (2000b) Fuzzy spline functions and nonlinear interpolation of if-then rules. In: Proceedings of the international conference on soft computing. pp 785–792
- Kawaguchi MF, Miyakoshi M (2001) Fuzzy spline interpolation in sparse fuzzy rule bases. In: A new paradigm of knowledge engineering by soft computing. World Scientific, pp 95–120
- Kawaguchi MF, Miyakoshi M, Kawaguchi M (1997) Linear interpolation with triangular rules in sparse fuzzy rule bases. In: Proceedings of the IFSA world congr. pp 138–143
- Kóczy L, Hirota K (1991) Rule interpolation by α -level sets in fuzzy approximate reasoning. *J BUSE-FAL, Automne, URA-CNRS* 46:115–123
- Kóczy L, Hirota K (1993a) Approximate reasoning by linear rule interpolation and general approximation. *Int J Approx Reason* 9(3):197–225
- Kóczy L, Hirota K (1993b) Interpolative reasoning with insufficient evidence in sparse fuzzy rule bases. *Inf Sci* 71(1):169–201
- Kóczy L, Hirota K (1993c) Ordering, distance and closeness of fuzzy sets. *Fuzzy Sets Syst* 59(3):281–293
- Koczy LT, Hirota K (1997) Size reduction by interpolation in fuzzy rule bases. *IEEE Trans Syst Man Cybern Part B (Cybern)* 27(1):14–25
- Kóczy LT, Hirota K, Gedeon TD (2000a) Fuzzy rule interpolation by the conservation of relative fuzziness. *JACIII* 4(1):95–101

- Kóczy LT, Hirota K, Muresan L (2000b) Interpolation in hierarchical fuzzy rule bases. In: Ninth IEEE international conference on fuzzy systems. FUZZ-IEEE 2000 (Cat. No. 00CH37063), vol 1. IEEE, pp 471–477
- Kovács S (2006) Extending the fuzzy rule interpolation five by fuzzy observation. In: Computational intelligence, theory and applications. Springer, pp 485–497
- Kovács S, Kóczy LT (1997a) Approximate fuzzy reasoning based on interpolation in the vague environment of the fuzzy rule base as a practical alternative of the classical cri. In: Proceedings of the 7th International fuzzy systems association world congress. Prague, Czech Republic, pp 144–149
- Kovács S, Kóczy LT (1997b) Approximate fuzzy reasoning based on interpolation in the vague environment of the fuzzy rulebase. In: Proceedings of IEEE international conference on intelligent engineering systems. IEEE, pp 63–68
- Kovács S, Kóczy LT (1997c) The use of the concept of vague environment in approximate fuzzy reasoning. In: Fuzzy set theory and applications, vol 12. Tatra Mountains Mathematical Publications, Mathematical Institute Slovak Academy of Sciences, Bratislava, Slovak Republic, pp 169–181
- Kovács S, Kóczy LT (1999) Application of an approximate fuzzy logic controller in an agv steering system, path tracking and collision avoidance strategy. In: Fuzzy set theory and applications, vol 16. In Tatra Mountains Mathematical Publications, Mathematical Institute Slovak Academy of Sciences, pp 456–467
- Kovács S, Kóczy LT (2004) Application of interpolation-based fuzzy logic reasoning in behaviour-based control structures. In: 2004 IEEE international conference on fuzzy systems (IEEE Cat. No. 04CH37542), vol 3. IEEE, pp 1543–1548
- Li YM, Huang DM, Zhang LN (2005) Weighted fuzzy interpolative reasoning method. In: 2005 Proceedings of 2005 international conference on machine learning and cybernetics, vol 5. IEEE, pp 3104–3108
- Li F, Shang C, Li Y, Shen Q (2017) Feature ranking-guided fuzzy rule interpolation. In: 2017 IEEE international conference on fuzzy systems (FUZZ-IEEE). IEEE, pp 1–6
- Li F, Li Y, Shang C, Shen Q (2018a) Improving fuzzy rule interpolation performance with information gain-guided antecedent weighting. *Soft Comput* 22(10):3125–3139
- Li F, Shang C, Li Y, Yang J, Shen Q (2018b) Fuzzy rule based interpolative reasoning supported by attribute ranking. *IEEE Trans Fuzzy Syst* 26(5):2758–2773
- Li F, Shang C, Li Y, Shen Q (2019) Interpretable mammographic mass classification with fuzzy interpolative reasoning. *Knowl Based Syst* 191:105279
- Li F, Li Y, Shang C, Shen Q (2020a) Fuzzy knowledge-based prediction through weighted rule interpolation. *IEEE Trans Cybern* 50(10):4508–4517
- Li F, Shang C, Li Y, Yang J, Shen Q (2020b) Interpolation with just two nearest neighbouring weighted fuzzy rules. *IEEE Trans Fuzzy Syst* 28(9):2255–2262
- Mamdani E, Assilian S (1999) An experiment in linguistic synthesis with a fuzzy logic controller. *Int J Hum Comput Stud* 51(2):135–147
- Mendel JM, John RI, Liu F (2006) Interval type-2 fuzzy logic systems made simple. *IEEE Trans Fuzzy Syst* 14(6):808–821
- Mukaidono M, Ding L, Shen Z (1990) Approximate reasoning based on revision principle. In: Proceedings of NAFIPS'90, vol 1. pp 94–97
- Nagy S, Sziová B, Kóczy LT (2018) The effect of image feature qualifiers on fuzzy colorectal polyp detection schemes using kh interpolation-towards hierarchical fuzzy classification of coloscopic still images. In: 2018 IEEE international conference on fuzzy systems (FUZZ-IEEE). IEEE, pp 1–7
- Naik N, Diao R, Shang C, Shen Q, Jenkins P (2017a) D-fri-winfirewall: dynamic fuzzy rule interpolation for windows firewall. In: 2017 IEEE international conference on fuzzy systems (FUZZ-IEEE). IEEE, pp 1–6
- Naik N, Diao R, Shen Q (2017b) Dynamic fuzzy rule interpolation and its application to intrusion detection. *IEEE Trans Fuzzy Syst* 26(4):1878–1892
- Ross TJ (2005) Fuzzy logic with engineering applications. Wiley, Hoboken
- Shen Q, Yang L (2011) Generalisation of scale and move transformation-based fuzzy interpolation. *J Adv Comput Intell Inform* 15(3):288–298
- Shen Z, Ding L, Mukaidono M, (1988) Fuzzy resolution principle. In: Proceedings. The eighteenth international symposium on multiple-valued logic. IEEE, pp 210–215
- Shen Z, Ding L, Mukaidono M (1993) Methods of revision principle. In: Proceedings of the 5th IFSA World Congr. pp 246–249
- Terano T, Asai K, Sugeno M (2014) Applied fuzzy systems. Academic Press, Cambridge

- Tikk D (1999) Investigation of fuzzy rule interpolation and the universal approximator property of fuzzy controllers. PhD thesis, PhD thesis, Department of Telecommunications and Telematics
- Tikk D, Baranyi P (2000) Comprehensive analysis of a new fuzzy rule interpolation method. *IEEE Trans Fuzzy Syst* 8(3):281–296
- Tikk D, Kóczy L, Moser B (1997) Stability of interpolative fuzzy kh-controllers. multi-dimensional case. In: Proceedings of the tempus JEP modify symposium. p 8
- Tikk D, Baranyi P, Yam Y, Kóczy LT (1999) Stability of a new interpolation method. In: IEEE SMC'99 conference proceedings. 1999 IEEE international conference on systems, man, and cybernetics (Cat. No. 99CH37028), vol 3. IEEE, pp 7–9
- Tikk D, Jóó I, Kóczy L, Várlaki P, Moser B, Gedeon TD (2002) Stability of interpolative fuzzy KH controllers. *Fuzzy Sets Syst* 125(1):105–119
- Tikk D, Csaba Johanyák Z, Kovács S, Wong KW (2011) Fuzzy rule interpolation and extrapolation techniques: criteria and evaluation guidelines. *J Adv Comput Intell Inform* 15(3):254–263
- Turksen I, Zhong Z (1988) An approximate analogical reasoning approach based on similarity measures. *IEEE Trans Syst Man Cybern* 18(6):1049–1056
- Ughetto L, Dubois D, Prade H (2000) Fuzzy interpolation by convex completion of sparse rule bases. In: Ninth IEEE international conference on fuzzy systems. FUZZ-IEEE 2000 (Cat. No. 00CH37063), vol 1. IEEE, pp 465–470
- Vass G, Kalmár L, Kóczy L (1992) Extension of the fuzzy rule interpolation method. In: Proceedings of the international conference fuzzy sets theory applications. pp 1–6
- Vincze D, Kovács S (2008) Using fuzzy rule interpolation based automata for controlling navigation and collision avoidance behaviour of a robot. In: 2008 IEEE international conference on computational cybernetics. IEEE, pp 79–84
- Wong K, Gedeon T (2000) Petrophysical properties prediction using self-generating fuzzy rules inference system with modified alpha-cut based fuzzy interpolation. In: Proceedings of the seventh international conference of neural information processing ICONIP (2000), Korea, (November 2000). pp 1088–1092
- Wong KW, Gedeon T, Tikk D (2000) An improved multidimensional α -cut based fuzzy interpolation technique. In: International conference on artificial intelligence in science and technology (AISAT 2000). University of Tasmania, pp 16–20
- Wong KW, Tikk D, Gedeon TD, Kóczy LT (2005) Fuzzy rule interpolation for multidimensional input spaces with applications: a case study. *IEEE Trans Fuzzy Syst* 13(6):809–819
- Wu ZQ, Masaharu M, Shi Y (1996) An improvement to kóczy and Hirota's interpolative reasoning in sparse fuzzy rule bases. *Int J Approx Reason* 15(3):185–201
- Yam Y, Kóczy L (1998) Cartesian representation for fuzzy interpolation. In: Proceedings of the 37th IEEE conference on decision and control (Cat. No. 98CH36171), vol 3. IEEE, pp 2936–2937
- Yam Y, Kóczy LT (2000) Representing membership functions as points in high-dimensional spaces for fuzzy interpolation and extrapolation. *IEEE Trans Fuzzy Syst* 8(6):761–772
- Yam Y, Kóczy L (2001) Fuzzy interpolation with cartesian representation and extensibility functions. In: Proceedings joint 9th IFSA world congress and 20th NAFIPS international conference (Cat. No. 01TH8569). IEEE, pp 2852–2857
- Yam Y, Baranyi P, Tikk D, Kóczy L (1999) Eliminating the abnormality problem of α -cut based interpolation. In: Proceedings of 8th IFSA World congress, vol 2. pp 726–766
- Yam Y, Kreinovich V, Nguyen HT (2000a) Extracting fuzzy sparse rules by cartesian representation and clustering. In: SMC 2000 conference proceedings. 2000 IEEE international conference on systems, man and cybernetics. 'cybernetics evolving to systems, humans, organizations, and their complex interactions' (cat. no. 0), vol 5. IEEE, pp 3778–3783
- Yam Y, Wong M, Baranyi PZ (2000b) Interpolation as mappings between cartesian spaces. *Proc Int Conf Intell Technol* 3:459–466
- Yang L, Shen Q (2011) Adaptive fuzzy interpolation. *IEEE Trans Fuzzy Syst* 19(6):1107–1126
- Yang L, Shen Q (2013) Closed form fuzzy interpolation. *Fuzzy Sets Syst* 225:1–22
- Yang L, Chao F, Shen Q (2017) Generalized adaptive fuzzy rule interpolation. *IEEE Trans Fuzzy Syst* 25(4):839–853
- Yang J, Shang C, Li Y, Li F, Shen Q (2019) ANFIS construction with sparse data via group rule interpolation. *IEEE Trans Cybern* 51(5):2773–2786
- Yang J, Shang C, Li Y, Li F, Shen L, Shen Q (2021) Constructing ANFIS with sparse data through group-based rule interpolation: an evolutionary approach. *IEEE Trans Fuzzy Syst*. <https://doi.org/10.1109/TFUZZ.2021.3049949>

Zadeh LA (1965) Fuzzy sets. *Inf Control* 8(3):338–353

Zadeh LA (1973) Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans Syst Man Cybern* 1:28–44

Zimmermann HJ (2011) Fuzzy set theory-and its applications. Springer, Berlin

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Fangyi Li^{1,2} · Changjing Shang² · Ying Li³ · Jing Yang^{2,3} · Qiang Shen² 

Fangyi Li
fal2@aber.ac.uk

Changjing Shang
cns@aber.ac.uk

Ying Li
lybyp@nwpu.edu.cn

Jing Yang
jy6@aber.ac.uk

¹ School of Artificial Intelligence, Beijing Normal University, Beijing, China

² Department of Computer Science, Aberystwyth University, Aberystwyth, UK

³ School of Computer Science, Northwestern Polytechnical University, Xi'an, China