Hierarchical Bidirectional Fuzzy Rule Interpolation

Jin, Shangzhu; Jiang, Yanling ; Peng, Jun; Shen, Qiang

Published in:
2018 IEEE 17th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC)

DOI:
10.1109/ICCI-CC.2018.8482077

Publication date:
2018

Citation for published version (APA):
Jin, S., Jiang, Y., Peng, J., & Shen, Q. (2018). Hierarchical Bidirectional Fuzzy Rule Interpolation. In N. Howard, S. Kwong, Y. Wang, J. Feldman, B. Widrow, & P. Sheu (Eds.), *2018 IEEE 17th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC)* (pp. 351-357). Article 8482077 (International Conference on Cognitive Informatics and Cognitive Computing). IEEE Press. https://doi.org/10.1109/ICCI-CC.2018.8482077

Hierarchical Bidirectional Fuzzy Rule Interpolation

Shangzhu Jin*, Yanling Jiang†, Jun Peng* and Qiang Shen‡

*College of Electrical and Information Engineering, Chongqing University of Science and Technology, Chongqing 401331, China
Email: gogcq@163.com, pengjun70@126.com

†College of Business Administration, Chongqing University of Science and Technology, Chongqing 401331, China
Email: jyl511621@126.com

‡Department of Computer Science, IMPACS, Aberystwyth University, Aberystwyth, UK, SY23 3DB
Email: qqs@aber.ac.uk

***Abstract*—The *"curse of dimensionality"* and *"sparse rule base"* are two common and important problems in conventional fuzzy systems. Using hierarchical fuzzy systems is an effective way to deal with the *"curse of dimensionality"* problem, whilst fuzzy rule interpolation offers a useful means for enhancing the robustness of fuzzy models, making inference possible in systems containing only a sparse rule base. In particular, backward fuzzy interpolation can be employed to allow interpolation to be carried out when certain antecedents of observation variables are absent, whereas conventional methods do not work. In order to deal with both *"curse of dimensionality"* and *"sparse rule base"* simultaneously, an initial idea of hierarchical bidirectional fuzzy interpolation is presented in this paper, combining hierarchical fuzzy systems and forward/backward fuzzy rule interpolation. Hierarchical bidirectional fuzzy interpolation is applicable to situations where a multiple multi-antecedent rules system needs to be reconstructed to a multi-layer fuzzy system and any sub-layer rule base is sparse. The implementation of this approach is based on fuzzy rule interpolative reasoning that utilities scale and move transformation. An illustrative example and application scenario are provided to demonstrate the efficacy of this proposed approach.**

***Index Terms*—Curse of dimensionality, Hierarchical system, Sparse rule base, Backward fuzzy rule interpolation**

I. INTRODUCTION

An effective way to deal with the "curse of dimensionality" is to use hierarchical fuzzy systems [1], [2]. Suppose that there are K input variables and M membership functions for each variable, then M^K rules are required in order to construct a system that fully covers the underlying problem domain. This often leads to the rule-explosion challenge facing systems modelling, usually referred to as the curse of dimensionality. The rule-explosion problem can be addressed in two ways. The first is to reduce the number of fuzzy partitions M, which usually results in significant reduction of model accuracy [3], [4], [5]. Besides, for many practical applications, there may not be sufficient historical data to support the creation of the needed rules that would cover the entire problem space, but a sparse rule base. Fortunately, fuzzy rule interpolation can be employed for dealing with this group of problems. The other way is to reduce the dimensionality K of the sub-rule bases using meta-levels or hierarchical fuzzy rule bases. A potentially more powerful case is the combination of both, which could improve the computational complexity dramatically [6]. In such a combined hierarchical interpolation system, however, situations may become even more complicated where certain crucial antecedents may be absent from given observations. This is because missing antecedents may well be involved in the subsequent (sub-system) inference process, causing the final conclusion not deducible.

To address the underlying problem of performing interpolation for certain antecedent variables, an original technique for backward fuzzy rule interpolation (B-FRI) has been proposed [7]. It is a branch of FRI and extends the existing FRI techniques. B-FRI can be employed to allow interpolation to be carried out when certain antecedents of observation variables are absent, whereas conventional methods do not work. The missing antecedents may be inferred or interpolated using the known antecedents and given conclusion during the interpolative reasoning process. B-FRI can also be employed to calculate certain antecedents for testing purposes, no mater whether the antecedents are known or unknown. It supports both interpolation and extrapolation which involve multiple intertwined fuzzy rules, with each having multiple antecedents. This allows missing observations which are directly related to the conclusion to be inferred or interpolated from the other known antecedents and the given conclusion. In addressing real-world problems, the rules adopted are typically irregular in nature (i.e., they may not always address the same antecedents). Indeed, rules may be arranged in an inter-connected mesh, where observations and conclusions in different subsets of rules may overlap, and yet may not be directly related throughout the entire rule base. For such complex systems, any missing values in a given set of observations may lead to failure if only unidirectional interpolation is employed.

In this paper, the initial theoretical work of hierarchical bidirectional fuzzy rule interpolation (HB-FRI) is proposed to meet the aforementioned challenges. Based on previous research work [7], [10], [11], hierarchical bidirectional fuzzy rule interpolation based on T-FRI can be proposed as outlined in the flowchart given in Figure 1. HB-FRI is herein implemented using scale and move transformation-based fuzzy interpolative reasoning (T-FRI) [8], [9], owing to their

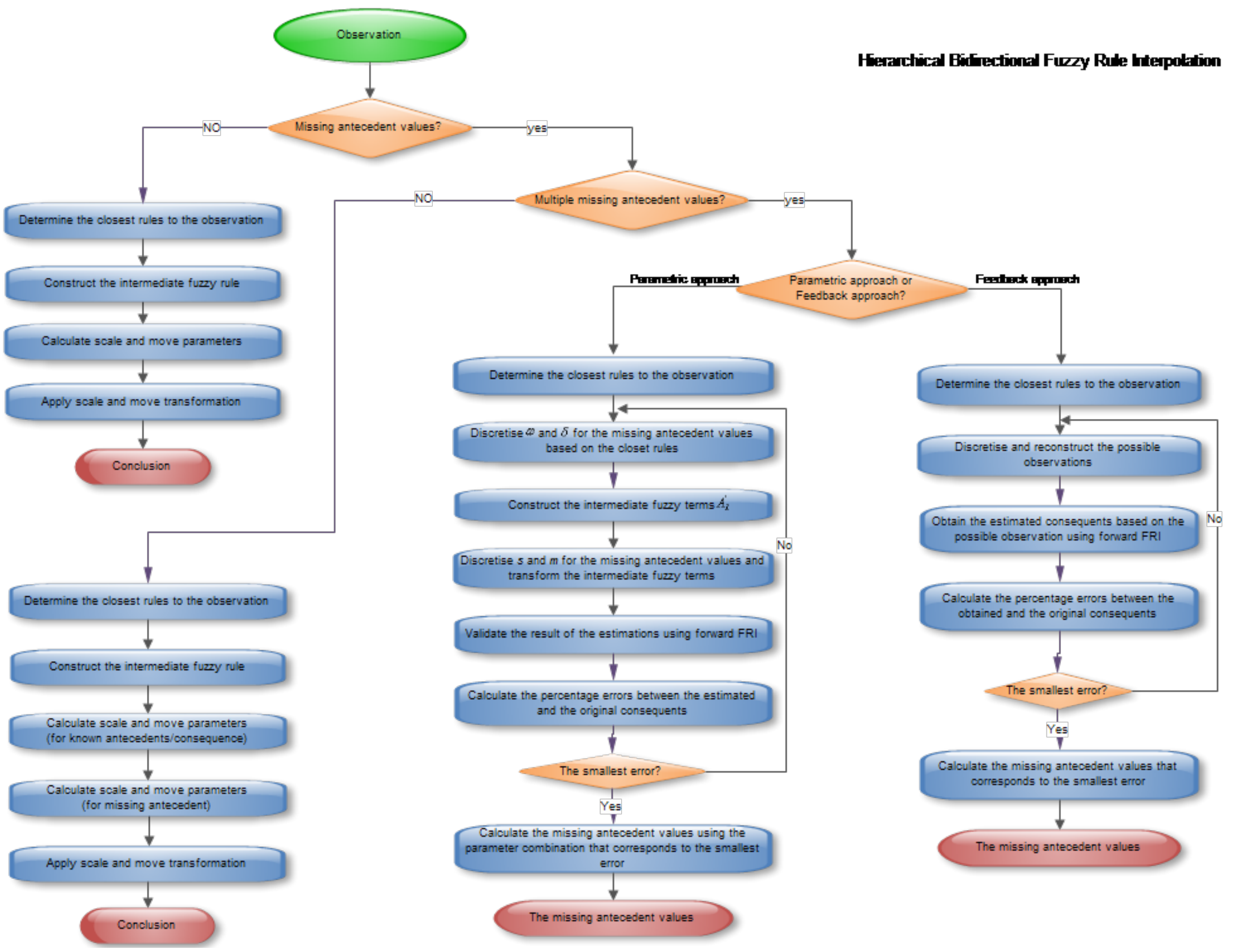

Fig. 1. Structure of hierarchical bidirectional fuzzy rule interpolation

popularity and availability (although other FRI methods may be adapted to serve as the alternative if preferred). In particular, T-FRI offers a flexible means to handle both interpolation and extrapolation involving multiple, multi-antecedent fuzzy rules. It guarantees the uniqueness, normality, and convexity of the resulting fuzzy sets. T-FRI is also able to handle various fuzzy set representations, including polygonal and bell-shaped fuzzy membership functions.

The rest of this paper is organised as follows. Section II introduces the general concepts of HB-FRI, along with detailed descriptions of a method for generating hierarchical fuzzy rule bases. Section III presents the scale and move transformation based fuzzy rule interpolation and backward fuzzy interpolation. The algorithm of hierarchical fuzzy interpolation approach is given in Section IV. An illustrative example is presented in Section V to explain the effectiveness of the approach. Section VI concludes the paper.

II. Hierarchical Rule Base Generation

A. Representation of Intermediate Variables

Without losing generality, the output variable of each layer within a certain *HFS* is represented by [12]:

$$y_{l,p} = \sum_{j_1 j_2 \ldots j_{P_{l-1,p}} i_1 i_2 \ldots i_{Q_{l,p}}} U_{l,p} V_{l,p} * y_{l,p}^{j_1 j_2 \ldots j_{P_{l-1,p}} i_1 i_2 \ldots i_{Q_{l,p}}} \tag{1}$$

where

$$U_{l,p} = \prod_{k=1}^{P_{l-1,p}} \mu_{l,p,k}^{jk}(y_{l-1,p,k}) \tag{2}$$

and

$$V_{l,p} = \prod_{k=1}^{Q_{l-1,p}} v_{l,p,k}^{jk}(x_{l,p,k}) \tag{3}$$

with $y_{l,p}$ representing the output of the p^{th} fuzzy subsystem in the l^{th} layer; $y_{l,p}^{j_1 j_2 \ldots j_{P_{l-1,p}} i_1 i_2 \ldots i_{Q_{l,p}}}$ being the THEN part of $j_1 j_2 \ldots j_{P_{l-1,p}} i_1 i_2 \ldots i_{Q_{l,p}}\ {}^{th}$ fuzzy rule; $P_{l-1,p}$ being the total number of outputs from the $(l-1)^{th}$ layer to $F_{l,p}$; $Q_{l-1,p}$ being the total number of original input variables to $F_{l,p}$;

and $\mu_{l,p,k}^{jk}(y_{l-1,p,k})$ and $\upsilon_{l,p,k}^{jk}(x_{l,p,k})$ are fuzzy membership functions for $y_{l-1,p,k}$ and $x_{l,p,k}$, respectively.

B. Learning Algorithm

For a standard fuzzy system the Least Square Method (LSM) is usually used to gain an optimal modelling result. However, it is not easy to apply LSM when developing a hierarchical fuzzy system because in many cases the intermediate variables have no physical meaning. A possible solution is to use gradient-descent techniques [12], [14], such as the error backpropagation algorithm, which is a popular method to optimise the parameters in hierarchical fuzzy systems. The parameter updating of the lower levels is based on the errors propagated back from the upper fuzzy layer (which are ultimately based on the exploitation of the error back-propagated from the the final output). The gradient-descent learning algorithm is given as follows.

First, let $e(k)$ be the error between the actual output $y(k)$ and the hierarchical system output $y^{'}(k)$ at time k:

$$e(k) = y^{'}(k) - y(k) \tag{4}$$

and the errors propagated back are defined as:

$$e_p(k) = e_q(k) \times \frac{\partial y_q(k)}{\partial y_p(k)} \tag{5}$$

where $e_p(k)$ denotes the error of sub-fuzzy system $p, p \geq 1$, which is propagated from its immediate adjacent upper sub-layer fuzzy system $q, q = p-1$, and $e_q(k)$ is defined in the same manner.

The parameters $y_p^{j_1 j_2 \dots j_{P_{q,p}} i_1 i_2 \dots i_{Q_{q,p}}}(k)$ in gradient descent learning are computed by

$$\begin{aligned} & y_p^{j_1 j_2 \dots j_{P_{q,p}} i_1 i_2 \dots i_{Q_{q,p}}}(k+1) \\ & = y_p^{j_1 j_2 \dots j_{P_{q,p}} i_1 i_2 \dots i_{Q_{q,p}}}(k) - \eta \times U_q(k) V_q(k) \times e_p \end{aligned} \tag{6}$$

where η is the learning rate, $U_q(k) = \prod_{i=1}^{P_q} \mu_{q,i}^{j_i}(k)$, and $V_q(k) = \prod_{i=1}^{Q_q} \mu_{q,i}^{j_i}(x_p, i)$.

Based on the above, the algorithm for learning a hierarchical rule base can be summarised below:

1) Choose the membership functions for each input variable, including both the original, real input variables of the system being modelled and the intermediate variables. An even partition for each input variable is assigned on their corresponding definition domain. The definition domain for the original input variables can be directly gained from the training data set. However, the definition domain for the intermediate variables may not have any actual meaning and hence, may not be associated with any explicit definition domain. Thus, the definition domain for the intermediate variables can be assumed (and normalised) as $[0, 1]$.
2) Choose initial parameters $y_{l,p}^{j_1 j_2 \dots j_{P_{l-1,p}} i_1 i_2 \dots i_{Q_{l,p}}}$ for the p^{th} sub-fuzzy system of the l^{th} level randomly. These parameters will be adjusted in the following steps.
3) Update the parameters $y_{l,p}^{j_1 j_2 \dots j_{P_{l-1,p}} i_1 i_2 \dots i_{Q_{l,p}}}$ with respect to each learning iteration k, for each given input-output pair (x^r, y^r), where r denotes the index of the training data.
4) Go to Step 3 with $r = r + 1$ if $r < T$, where T is the total number of training data in the training set.
5) Compute the accumulated error $E = \frac{1}{2} \times \sum_{r=1}^{T} (\hat{y}^r - y^r)^2$ and check if E is less than a prespecified small value ϵ, or if k is larger than a prespecified maximal iteration number K: if so, end the training process; else, go to Step 3 with $k = k + 1$.

III. Bidirectional Fuzzy Rule Interpolation

A. Transformation-based Forward Fuzzy Rule Interpolation (T-FRI)

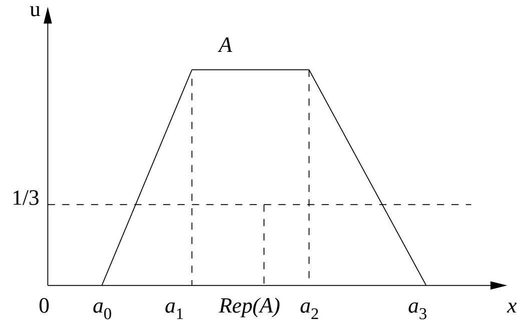

Fig. 2. Representative value of trapezoidal fuzzy set

Trapezoidal fuzzy sets are adopted in the present work, as with the common practice in the T-FRI literature. A key concept used in T-FRI is the representative value $Rep(A)$ of a given fuzzy set A. When trapezoidal representation is used, $Rep(A)$ is defined as the centre of gravity of its four points (a_0, a_1, a_2, a_3):

$$Rep(A) = \frac{a_0 + \frac{a_1 + a_2}{2} + a_3}{3} \tag{7}$$

where a_0, a_3 represent the left and right extremities (with membership values 0), and a_1, a_2 denote the normal points (with membership value 1) over the support of the fuzzy set, as shown in Fig. 2. Based on this representation, T-FRI can be summarised as follows:

1) Determination of the Closest Rules: Given a rule base $\mathbb{U}$, a fuzzy rule $R \in \mathbb{U}$ with M antecedents $A_k, k = 1, 2, \cdots, M$, and an observation O are expressed in the following format:

R: IF x_1 is A_1, $\cdots$, and x_k is A_k, $\cdots$, and x_M is A_M, THEN y is B
O: A_1^*, $\cdots$, A_k^*, $\cdots$, A_M^*

The distance d between a rule and an observation is determined by computing the aggregated distance of all the antecedent variables:

$$d = \sqrt{\sum_{k=1}^{M} d(A_k, A_k^*)^2} \tag{8}$$

with

$$d(A_k, A_k^*) = \frac{d(Rep(A_k), Rep(A_k^*))}{range_k} \tag{9}$$

where $range_k = sup_k - inf_k$ is the domain range of the variable x_k. As such, $d(A_k, A_k^*) \in [0, 1]$ is the normalised result of the otherwise absolute distance measure, so that distances are compatible with each other over different variable domains. The N $(N \geq 2)$ rules which have the least distance measurements with regard to the observed values A_k^* are then chosen to be used in the later steps.

2) Construction of the Intermediate Rule: The intermediate fuzzy terms $A_k^\dagger$ that are to be used to build the required intermediate rule are constructed from the antecedents of the N closest rules. These are then shifted to $A_k^{'}$ such that they have the same representative values as those of A_k^*. The shifted intermediate consequence $B^{'}$ can be computed, with the parameters ω_{B^i} and δ_B being aggregated from the corresponding values of $A_k^{'}$.

3) Scale Transformation: For each antecedent variable of the N chosen rules, the scale transformation works by calculating two scale rates $\overline{s}_{A_k}$ and $\underline{s}_{A_k}$. The support (a_0', a_3') of the corresponding shifted fuzzy set $A^{'}$ is transformed into a new support (a_0'', a_3''), and the core (a_1', a_2') is transformed into another (a_1'', a_2''). This leads to a scaled fuzzy set $A_k^{''} = (a_0'', a_1'', a_2'', a_3'')$. The corresponding parameters $\underline{s}_B$ and $\overline{s}_B$ of fuzzy set B^* can be calculated as follows:

$$\underline{s}_B = \frac{1}{M}\sum_{k=1}^{M}\underline{s}_{A_k} \qquad \overline{s}_B = \frac{1}{M}\sum_{k=1}^{M}\overline{s}_{A_k} \tag{10}$$

4) Move Transformation: In general, for multiple antecedent rules, each variable dimension has its own move rate m_{A_k}, in order to move each of the scaled fuzzy sets $A_k^{''}$ to new locations, that coincide with those of the originally observed values. This allows the initially constructed intermediate fuzzy terms to be completely transformed so that they become the same as the given observation. That is, the final transformed fuzzy sets then match the exact shapes of the observed values A_k^*. Without losing generality, for a given scaled intermediate fuzzy term: $A_k^{''} = (a_0'', a_1'', a_2'', a_3'')$, its current support (a_0'', a_3''), and core (a_1'', a_2'') can be moved to (a_0, a_3) and (a_1, a_2), using a move rate m_{A_k} calculated as follows:

$$\begin{cases} m_{A_k} = \frac{3(a_0 - a_0'')}{a_1'' - a_0''}, \ a_0 \geq a_0'' \\ m_{A_k} = \frac{3(a_0 - a_0'')}{a_3'' - a_2''}, \ \text{otherwise} \end{cases} \tag{11}$$

Similar to the scale transformation, the move rate m_B for the consequent dimension can be calculated by obtaining the arithmetic average of those of the antecedent variables, such that:

$$m_B = \frac{1}{M}\sum_{k=1}^{M} m_{A_k} \tag{12}$$

The final interpolated result B^* that corresponds to the observation A^* can now be computed by applying the scale and move transformation to $B^{'}$, using the resulting parameters $\underline{s}_B$, $\overline{s}_B$, and m_B.

B. Transformation-based Backward Fuzzy Rule Interpolation (BFRI)

The BFRI algorithm that works reversely to T-FRI but follows a similar underlying approach, in order to infer a missing antecedent value, can be summarised below.

1) Determination of the Closest Rules: In reference to the earlier definition of the T-FRI process in Eqn. 8, when $B^*, (A_1^*, \cdots, A_{l-1}^*, A_{l+1}^*, \cdots, A_M^*)$ are given, in order to interpolate/extrapolate the unknown antecedent A_l^*, the discovery of the closest rules $R_i, i = 1, \cdots, N$, is required. Mirroring the use of the distance measure in T-FRI, the following scheme has been proposed in order to reflect the biased consideration towards the consequent variable:

$$\hat{d} = \sqrt{{d_B}^2 + \sum_{k=1,\ k\neq l}^{M} {d_{A_k}}^2} \tag{13}$$

2) Construction of the Intermediate Fuzzy Terms: To help explanation, suppose that a certain set of closest rules $R_i, i = 1, \cdots, N, R_i \in \mathbb{U}$ are identified, which are returned by applying the previous distance metric. Following the original T-FRI algorithm, in order to create the intermediate (shifted) fuzzy terms for the known antecedent variables: A_k', $k = 1, \cdots, M, k \neq l$, the following parameters $w_{A_k^i}$, $i = 1, \cdots, N$, and δ_{A_k} need to be computed first. The parameter values for the intermediate (shifted) consequent fuzzy term B': w_{B^i}, $i = 1, \cdots, N$, and δ_B can be computed using exactly the same formulae as those of A_k, since its value B^* is also directly observed. Both w_{B^i} and δ_B are algebraic averages of the parameter values from individual antecedent terms. The parameter values for the missing antecedent such as $\omega_{A_l^i}$, are then calculated by subtracting those of the known antecedents from that of the consequent. The acquisition of these parameter values entails the construction of the intermediate (shifted) fuzzy term A_l' for the missing antecedent dimension.

3) Scale and Move Transformation: Having obtained the intermediate (shifted) fuzzy terms, the essential parameters $\underline{s}_{A_l}$, $\overline{s}_{A_l}$, and m_{A_l} involved in the transformation process can be derived. Following the same intuition and computational steps as those for $w_{A_k^i}$, $i = 1, \cdots, N$, and δ_{A_l}, by reversing the forward transformation procedure introduced in Eqns. 10 and 12, the required values can be found as follows:

$$\begin{aligned} \underline{s}_{A_l} &= M\underline{s}_B - \sum_{k=1,\ k\neq l}^{M} \underline{s}_{A_k} \\ \overline{s}_{A_l} &= M\overline{s}_B - \sum_{k=1,\ k\neq l}^{M} \overline{s}_{A_k} \\ m_{A_l} &= Mm_B - \sum_{k=1,\ k\neq l}^{M} m_{A_k} \end{aligned} \tag{14}$$

Finally with all parameters acquired, the transformation on $A_l^{'}$ can be performed, resulting in the (backward) interpolated value A_l^*.

$$T(A_l^{'}, A_l^*) = \{\underline{s}_{A_l}, \overline{s}_{A_l}, m_{A_l}\} \tag{15}$$

IV. ALGORITHM OF HB-FRI

Further to the outline as shown previously in Figure 1, the HB-FRI algorithm can be summarised in principle, as follows.

1) Determine the distances between the observed values for each input variable and the antecedents of each rule in the sub-rule base of the lowest layer, and choose the closest rules to construct the transformation parameters required for computing the interpolation through the subsequent layers.
2) Calculate the sub-consequence for each sub-layer using the multiple multi-antecedent rules interpolation method. If this sub-consequence is not the output of the final layer, then, the sub-consequence will form the fuzzy term for the relevant input variable of the next layer.
3) Compute the output at the layer above from the terms of the intermediate variables, and the values of the original input variables of course, by iterating the first and second steps.

Thus, the hierarchical interpolative approach can be detailed below:

1) *Determine Closest Rules with Respect to Input Variables* When certain variables are present to an HFIU as input, the first task of the HFIU is to determine the closest rules for them. In general, the variables here not only refer to the original input variables involved in the observation, but may also include the intermediate variables which have been introduced by previous applications of HB-FRI. Particularly, the distance d_j^*, $j = 1, ..., k$ between the fuzzy set A_j^i and the set A_k^* is calculated using Eqn. 9. The n ($n \geq 2$) rules which have the minimum distance measures are chosen as the closest rules from the observation.
2) *Construct the Intermediate Rule* The computation and representation of the intermediate variable $y_i, (i = 1, 2, ..., K-2)$ is the most important issue in any interpolation, hierarchical or not; there is no exception in HB-FRI. The process of generating the intermediate variables of the HB-FRIis however, the same as outlined in Sections III-A2 and III-B2, depending on whether forward or backward T-FRI is required.
3) *Carry out Scale Transformation* For each trapezoidal antecedent fuzzy value appearing in any of the N chosen rules, this step calculates two scale rates $\overline{s}^*_{A_j}$ and $\underline{s}^*_{A_j}$ according to Eqn. 14, which rescale the top and bottom supports of $A_j^{'}$ with respect to the observation (or previously interpolated outcome) A_j^*, resulting in $A_j^{''}$. The corresponding bottom scale rate $\underline{s}_{B_j}$, and the intermediate top scale rate $\overline{s}_{B_j}{}'$ of the intermediate conclusion $B_j^{'}$, are obtained by averaging those computed for the antecedents. The final $\overline{s}_{B_j}$ is obtained from applying Eqn. 10.
4) *Carry out Move Transformation* Using the move rate m_{A_j} as given in Eqn. 12, $A_j^{''}$ is moved so that the final transformed fuzzy set matches the exact shape of the observed (or interpolated) value A_j^*. From this, m_{B_j} for the conclusion can be calculated according to Eqn. 12. The final interpolated result B_j^* can now be calculated by applying the scale and move transformation to $B_j^{'}$, using the parameters $\underline{s}_{B_j}$, $\overline{s}_{B_j}$, and m_{B_j}.

V. FUNCTION APPROXIMATION AND EXPERIMENTAL EVALUATION

A. Experimental Setup

In this section, an illustrative example is used to demonstrate the proposed HB-FRI approach. In particular, a function approximation problem with three input variables is considered:

$$y = f(x_1, x_2, x_3) = (1 + {x_1}^{0.5} + {x_2}^{-1} + {x_3}^{-1.5})^2$$

where six fuzzy sets are evenly defined for each input variable. The hierarchical structure is shown in Fig. 3, where HFIU1 and HFIU2 are the two sub-system units, i.e., the first layer and second layer of this HB-FRI approximation model, respectively. For simulation, 341 samples are uniformly created on the 3-dimensional problem space, 216 of which are used for training and the remaining 125 for testing.

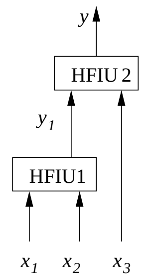

Fig. 3. Structure of illustrative hierarchical function approximation example

According to Section II-B, suppose that the learning rate is set as 0.0004, each original sub-layer fuzzy system will contain 32 rules, with the system having a total of 72 rules. In each hierarchical fuzzy interpolation unit, to demonstrate the effect of running only a sparse rule subset, it is assumed that only four fuzzy rules are chosen to construct its original sub-rule base. These four rules jointly describe the minimum and maximum of the output and also those output values in response to the four points bounding the corresponding input space. Tables I and II display the sub-fuzzy sparse rules of the lower layer and those of the upper layer, respectively.

The approximation accuracy is measured by the Average Percentage Error ($APE\%$) given as follows:

$$APE\% = (100/N)\sqrt{\sum_{i=1}^{N}(\frac{y_i^* - y_i^{'}}{y_i^{'}})} \quad (16)$$

where y_i^* is the i^{th} objective output of the underlying function and $y_i^{'}$ is the i^{th} model output.

TABLE I
FUZZY SUB-SPARSE RULE BASE OF THE LOWER LAYER

Rule	Antecedents	Consequence
Rule 1	$x_1 = A_1^1 = (0.02, 0.51, 1.52, 2.01)$, $x_2 = A_2^1 = (0.03, 0.53, 1.53, 2.03)$	$y_1 = B_1^1 = (-0.24, 0.26, 1.26, 1.76)$
Rule 2	$x_1 = A_1^2 = (0.02, 0.51, 1.51, 2.01)$, $x_2 = A_2^2 = (4.99, 5.49, 6.49, 6.99)$	$y_1 = B_1^2 = (0.00, 0.50, 1.50, 2.00)$
Rule 3	$x_1 = A_1^3 = (4.99, 5.49, 6.49, 6.99)$, $x_2 = A_2^3 = (0.03, 0.53, 1.53, 2.03)$	$y_1 = B_1^3 = (-1.00, -0.50, 0.50, 1.00)$
Rule 4	$x_1 = A_1^4 = (4.99, 5.49, 6.49, 6.99)$, $x_2 = A_2^4 = (4.99, 5.49, 6.49, 6.99)$	$y_1 = B_1^4 = (-0.65, -0.15, 0.85, 1.34)$

TABLE II
FUZZY SUB-SPARSE RULE BASE OF THE UPPER LAYER

Rule	Antecedents	Consequence
Rule 1	$y_1 = B_1^1 = (-1.00, -0.50, 0.50, 1.0)$, $x_3 = A_3^1 = (0.01, 0.50, 1.503, 2.55)$	$y = B^1 = (20.76, 21.26, 22.26, 22.76)$
Rule 2	$y_1 = B_1^2 = (-1.00, -0.50, 0.50, 1.00)$, $x_3 = A_3^2 = (4.98, 5.48, 6.48, 6.98)$	$y = B^2 = (14.90, 15.40, 16.40, 16.90)$
Rule 3	$y_1 = B_1^3 = (0.00, 0.50, 1.50, 2.00)$, $x_3 = A_3^3 = (0.00, 0.50, 1.50, 2.00)$	$y = B^3 = (6.37, 6.87, 7.87, 8.37)$
Rule 4	$y_1 = B_1^4 = (0.00, 0.50, 1.50, 2.00)$, $x_3 = A_3^4 = (4.98, 5.48, 6.48, 6.98)$	$y = B^4 = (3.20, 3.70, 4.70, 5.20)$

TABLE III
RESULT COMPARISON BETWEEN STANDARD AND HIERARCHICAL FUZZY SYSTEM

Method	Partitions	Rules	APE% testing
Standard Fuzzy System	(5,5,5)	216	6.50119
Standard Fuzzy System	(4,4,4)	125	6.37989
Hierarchical Fuzzy System	(5,5,5)	72	1.48293
Hierarchical Fuzzy System	(4,4,4)	32	1.71127
Hierarchical Interpolation System	(5,5,5)	8	4.84735
Hierarchical Interpolation System	(4,4,4)	8	5.76442

B. Analysis of Results

Table III presents a comparative summary of running different models through different sets of data, including: the Standard Fuzzy Systems that use all the rules given in a flat set; the Hierarchical Fuzzy Systems that use all a subset of rules arranged hierarchically; and the Hierarchical Interpolation System that only use a small portion of the rules employed in the Hierarchical Fuzzy Systems.

Note that when six fuzzy sets are defined for each input variable, the maximum number (N_{max}) of involved fuzzy rules is $6^3 = 216$ for the Standard Fuzzy System, and $6^2 + 6^2 = 72$ for Hierarchical Fuzzy System, whereas, the minimum number (N_{min}) of rules is $2 + 2 = 4$ for the proposed HB-FRI. In order to attain the accuracy of inference while using fewer number of rules, in this experimentation, the cardinality of each sub-layer rule base is taken to 4, so the total number of rules of the HB-FRI is 8. To enhance the comparison, different setups where fewer rules are involved due to looser partitions of the three input variables are also tested, as reflected in Table III.

Comparing the approximation results given in Table III, the following observations can be made. Firstly, the hierarchical fuzzy interpolative approximation performs relatively well in the sense that it can achieve good approximation with significantly fewer rules. The accuracy of HB-FRI is indeed between that of the Standard Fuzzy System and that of the Hierarchical Fuzzy System, but the other two utilise substantially more fuzzy rules and hence, require much more computational effort. Secondly, it can also be observed from the results on the APE% of the proposed Hierarchical Interpolation System that the more fuzzy terms are included to represent the system variables, the more accurate the inference results may be obtained without increasing the number of rules used, showing the robustness of the approach. This is not the case for the other two types of fuzzy model (though the classical hierarchical technique also leads to more accurate results if more detailed descriptions on variables are used, but this requires a much large number of rules).

In addition, due to the underlying principle taken by the proposed approach for hierarchical bidirectional fuzzy rule interpolation, the sub-layer rule bases can be constructed with less constraints over what intermediate variables to use such that the overall hierarchical model represents the original input-output problem space. This property enables HB-FRI to allow for different settings for the original input variables to flow into the sub-layers in any arbitrary order without affecting the final outcome. This may have been implicitly reflected in the results gathered in Table III, but an explicit verification of this remains active research.

VI. CONCLUSION

This paper has presented a novel technique, named hierarchical fuzzy rule interpolation, to address the problem of fuzzy modelling in high-dimensional systems. HB-FRI enables unknown antecedent values to be interpolated, given other antecedents and the conclusion. This integrated approach, of hierarchical reasoning and bidirectional interpolative inference, provides a flexible and systematic way of dealing with insufficient information or knowledge that may often appear in real-world problems. The system implementing the proposed technique is able to draw a final conclusion through the exploitation of BFRI even when it is presented with partial observations. This also helps identify hidden variables that may be useful during any subsequent intelligent decision support processes. The initial experimental results reported have demonstrated that the proposed method can retain model accuracy while significantly reducing the number of the rules required in the system model. As indicated earlier, to reveal the full potential of this work, an investigation into exactly how flexible in designing the hierarchical structure by introducing different intermediate variables requires further experimental research.

ACKNOWLEDGEMENTS

The work described in this study was supported by the Foundation of Doctor and Professor's Research Project of Chongqing University of Science and Technology (Grant No. CK2016B02), the Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJ1713329), and the Education Teaching Reform Research Key Project of Chongqing University of Science and Technology (Grant No. 201614).

REFERENCES

[1] G. Raju, J. Zhou, and A. ROGER, "Hierarchical fuzzy control," *International journal of control*, vol. 54, no. 5, pp. 1201–1216, 1991.
[2] G. Raju and J. Zhou, "Adaptive hierarchical fuzzy controller," *IEEE Trans. Syst., Man, Cybern.*, vol. 23, no. 4, pp. 973–980, 1993.
[3] M. Sugeno and G. Kang, "Fuzzy modeling and control of multilayer incinerator," *Fuzzy Sets and Systems*, vol. 18, pp. 329–346, 1986.
[4] M. Sugeno, I. Hirano, S. Nakamura, and S. Kotsu, "Development of an intelligent unmanned helicopter," in *Proceedings of International Conference on Fuzzy Systems*, vol. 5. IEEE, 1995, pp. 33–34.
[5] M. Sugeno and G. Kang, "Structure identification of fuzzy model," *Fuzzy sets and systems*, vol. 28, no. 1, pp. 15–33, 1988.
[6] L. T. Kóczy, K. Hirota, and L. Muresan, "Interpolation in hierarchical fuzzy rule bases," in *Proceedings of International Conference on Fuzzy Systems*, 2000, pp. 471–477.
[7] S. Jin, R. Diao, C. Quek, and Q. Shen, "Backward fuzzy rule interpolation," *IEEE Trans. Fuzzy Syst.*, vol. 22, no. 6, pp. 1682–1698, 2014.
[8] Z. Huang and Q. Shen, "Fuzzy interpolative reasoning via scale and move transformations," *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 2, pp. 340–359, 2006.
[9] ——, "Fuzzy interpolation and extrapolation: A practical approach," *IEEE Trans. Fuzzy Syst.*, vol. 16, no. 1, pp. 13–28, 2008.
[10] S. Jin, R. Diao, C. Quek, and Q. Shen, "Backward fuzzy rule interpolation with multiple missing values," in *Proceedings of IEEE International Conference on Fuzzy Systems*, 2013, pp. 1–8.
[11] S. Jin, R. Diao, and Q. Shen, "Backward fuzzy interpolation and extrapolation with multiple multi-antecedent rules," in *Proceedings of IEEE International Conference on Fuzzy Systems*, 2012, pp. 1170–1177.
[12] D. Wang, X. Zeng, and J. Keane, "Intermediate variable normalization for gradient descent learning for hierarchical fuzzy system," *IEEE Trans. Fuzzy Syst.*, vol. 17, no. 2, pp. 468–476, 2009.
[13] H. Wang, S. Kwong, Y. Jin, W. Wei, and K. Man, "Multi-objective hierarchical genetic algorithm for interpretable fuzzy rule-based knowledge extraction," *Fuzzy sets and systems*, vol. 149, no. 1, pp. 149–186, 2005.
[14] L. Wang, "Analysis and design of hierarchical fuzzy systems," *IEEE Trans. Fuzzy Syst.*, vol. 7, no. 5, pp. 617–624, 1999.