Bidirectional approximate reasoning-based approach for decision support

Jin, Shangzhu; Peng, Jun; Li, Zuojin; Shen, Qiang

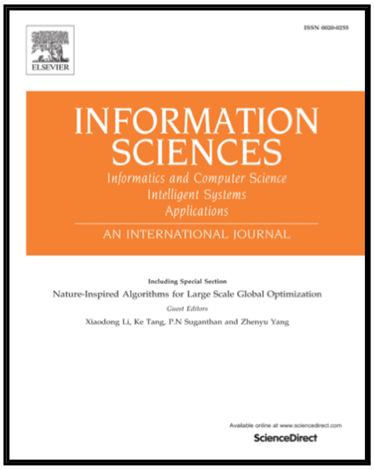
Published in:
Information Sciences

DOI:
10.1016/j.ins.2019.08.019

Publication date:
2019

Citation for published version (APA):
Jin, S., Peng, J., Li, Z., & Shen, Q. (2019). Bidirectional approximate reasoning-based approach for decision support. *Information Sciences*, *506*, 99-112. https://doi.org/10.1016/j.ins.2019.08.019

Journal Pre-proof



Bidirectional Approximate Reasoning-Based Approach for Decision Support

Shangzhu Jin, Jun Peng, Zuojin Li, Qiang Shen

PII: S0020-0255(19)30753-4
DOI: https://doi.org/10.1016/j.ins.2019.08.019
Reference: INS 14768

To appear in: *Information Sciences*

Received date: 5 December 2018
Revised date: 2 August 2019
Accepted date: 4 August 2019

Please cite this article as: Shangzhu Jin, Jun Peng, Zuojin Li, Qiang Shen, Bidirectional Approximate Reasoning-Based Approach for Decision Support, *Information Sciences* (2019), doi: https://doi.org/10.1016/j.ins.2019.08.019

Bidirectional Approximate Reasoning-Based Approach for Decision Support

Shangzhu Jin*, Jun Peng*, Zuojin Li* and Qiang Shen†

*College of Intelligent Technology and Engineering,
Chongqing University of Science and Technology,
Chongqing 401331, China
Email: gogcq@163.com, cqustlzj@sina.cn, pengjun70@126.com

†Department of Computer Science, Faculty of Business and Physical Sciences,
Aberystwyth University,
Aberystwyth SY23 3DB, UK
Email: qqs@aber.ac.uk

Abstract

Fuzzy rule-based systems are widely applied for real-world decision support, such as policy formation, public health analysis, medical diagnosis, and risk assessment. However, they face significant challenges when the application problem at hand suffers from the *"curse of dimensionality"* or *"sparse knowledge base"*. Combination of hierarchical fuzzy rule models and fuzzy rule interpolation offers a potentially efficient and effective approach to dealing with both of these issues simultaneously. In particular, backward fuzzy rule interpolation (B-FRI) facilitates approximate reasoning to be performed given a sparse rule base where rules do not fully cover all observations or the observations are not complete, missing antecedent values in certain available rules. This paper presents a hierarchical bidirectional fuzzy reasoning mechanism by integrating hierarchical rule structures and forward/backward rule interpolation. A computational method is proposed, building on the resulting hierarchical bidirectional fuzzy interpolation to maintain consistency in sparse fuzzy rule bases. The proposed techniques are utilised to address a range of decision support problems, successfully demonstrating their efficacy.

Index Terms

Hierarchical systems, Fuzzy rule interpolation, Bidirectional interpolation, Rule base refinement, Decision support.

I. INTRODUCTION

It is well known that uncertainty virtually exists in all complex situations which require intelligent decision-making [25]. Fuzzy logic based approximate reasoning offers a practically applicable approach to dealing with uncertainty and uncertainty management in decision support [7]. For instance, a fuzzy logic based expert system has been developed to diagnose common diseases such as pneumonia and jaundice, which is capable of returning an identified disease as output given observed symptoms [22]. Also for medical decision support, there have been many hybrid systems which work by integrating fuzzy logic and other computational intelligence mechanisms. These include: a fuzzy neural network approach for treating the diabetes and heart diseases [13]; an adaptive neural fuzzy rough inference system for facilitating diagnosis of tuberculosis disease [24]; and an interval-valued fuzzy rule-based classification tool for performing diagnosis of cardiovascular diseases [23]. All such work collectively demonstrates the success of fuzzy rule-based decision systems.

However, the "curse of dimensionality", namely the number of rules required to perform approximate reasoning increasing exponentially along with the number of input features [21], causes a major challenge for many automated decision-making systems, including fuzzy logic-based ones. This is because for a fuzzy rule model containing K variables, with each variable partitioned into M fuzzy values, the order of the number of the rules required in a conventional rule base is $O(M^K)$. The effort to address this difficulty has led to the development of hierarchical fuzzy systems [20]. The number of rules in a typical hierarchical fuzzy system as shown in Fig. 1 only increases linearly with an increasing number of input variables, if a K-input hierarchical fuzzy system comprises K-1 low-dimensional fuzzy systems with each sub-system taking just two inputs. In this case, given M fuzzy sets for each variable, the total number of possible rules is $(K-1)M^2$ which is a linear function of the number of the input variables.

Hierarchical fuzzy systems help reduce the modelling complexity, but they work by generally assuming that there are dense fuzzy rule bases to cover most if not all of the problem space. Unfortunately, for problems that typically involve decision-making in novel situations (e.g., diarrhoeal disease diagnosis in newly built-up regions [30], organised crime investigation [3], and counter-terrorism surveillance deployment [12]), there does not normally exist sufficient historical data to entail the generation of a dense rule base covering the complete underlying problem domain. Indeed, a method for fuzzy rule interpolation with self error-correction and adaptation mechanisms has been proposed to assist in decision-making regarding the risk of regional diarrhoea development and spread [30]. This application will also serve as a focus for comparison and discussion in this paper. As with this particular problem, there are many others where only a "sparse rule base" is available, i.e., rules given in the rule base cover only small proportions

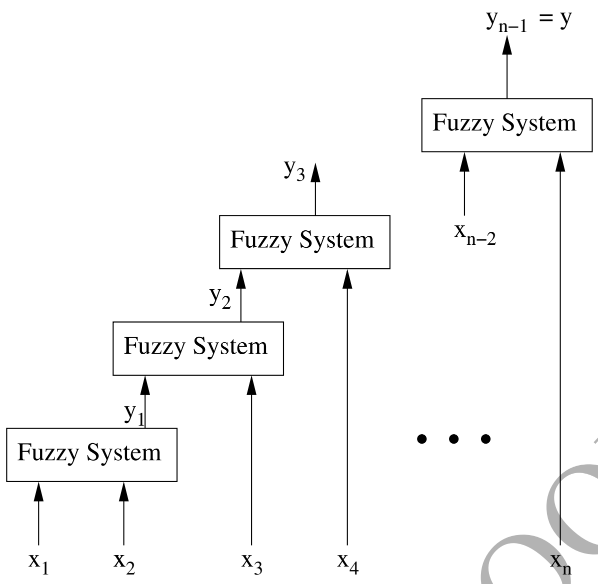


Fig. 1. A typical hierarchical fuzzy system

of the problem domain. This makes conventional rule-based reasoning unworkable for the uncovered areas. Of course, the problem may significantly worsen when there are too many input variables or values to consider [14]. Worse still is the situation in which the given observations are themselves incomplete, that is, certain features required to support decision-making may not be directly observable. Such missing antecedents in the decision rules may cause the breakdown of the inference procedure using classical fuzzy rule interpolation (FRI) that is successful in addressing normal problem cases, where all input features are assumed to be observable but there are no rules to match a certain observation.

The first attempt to address the aforementioned challenges facing the existing FRI work is the proposal for backward fuzzy rule interpolation (B-FRI) [12]. It is of great potential to enable rule interpolation to be implemented when certain rule antecedents are missing from the observation, through exploiting those given antecedents and any relevant consequent derivable during the approximate reasoning process. However, B-FRI does not explicitly tackle the problem where rules provided in the sparse rule base are arranged hierarchically. This remarkably restricts its effectiveness in a variety of possible applications where domain knowledge for decision-making is organised in a hierarchical manner. Use of hierarchical rule bases is of intuitive appeal, as they mimic more closely human expertise in handling inferences that involve many domain features. This research is therefore conducted to further develop the theoretical framework of B-FRI, supporting approximate interpolative reasoning through a hierarchy of intertwined fuzzy rules where each rule may contain multiple antecedent features that are not necessarily all observable.

The improved method is hereafter referred to as hierarchical bidirectional fuzzy rule interpolation

(HB-FRI). Following its original work as per B-FRI it is implemented by the use of the popular scale and move transformation-based fuzzy interpolative reasoning (T-FRI) mechanism [10]. Nonetheless, the underlying design principle appears to be sufficiently general. As such, the same idea may be applicable to adapting alternative FRI techniques if preferred, though the proof of this conjuncture requires further research.

Another important point is that the rule base in an intelligent decision-making system is often open. That is, the user can modify, add and delete the rules in the knowledge base. Inconsistent rules and redundant rules may be introduced in the process of such changes. This may increase the rule base complexity while decreasing rule-firing and interpolation efficacy. Inconsistencies may also arise from other aspects, such as incorrect selection of rules to fire or to interpolate with respect to inaccurate observations [30]. For real-world decision-making applications, where rules are not always involve the same antecedent features, such situations may deteriorate. To aid in handling this type of problem, a further contribution of this paper is to introduce a novel rule base refinement mechanism that allows for the removal of inconsistent rules in the rule base utilised by HB-FRI. This method works irrespective of whether any rule antecedent features are missing.

The remainder of this paper is structured as follows. Section II presents the framework of HB-FRI. For academic completeness, Section II-A first briefly introduces a method for data-driven derivation of hierarchical fuzzy rule bases that is used to form the basis upon which to represent the domain knowledge in a given application. Section II-B gives an overview of scale and move transformation based FRI and B-FRI, as the foundation to implement HB-FRI. Section III describes the algorithm that resolves inconsistencies in a fuzzy rule base via exploiting hierarchical FRI. Section IV reports on a range of experimental investigations and discusses the results. In particular, Section IV-A shows an illustrative numerical example of the proposed approach in action. Section IV-B offers a comparative analysis between the proposed work and the standard approach to FRI, while dealing with a set of prediction problems involving inconsistent rules. Section IV-C provides a real-world decision-making application using the implemented HB-FRI system. Section V concludes the paper and points out open issues for further research.

II. Hierarchical Bidirectional Fuzzy Rule Interpolation

A. Generation of Hierarchical Rule Bases

For a typical fuzzy rule-based decision-making system, especially for systems developed on the basis of historical data, data-driven learning is often applied to generate the required rule base. In particular, supervised learning is normally used to obtain an optimal rule base via heating search through a given labelled dataset. Yet, direct application of such a method to learn a hierarchical fuzzy rule base from data is not an easy task, especially for the problem domains where there are missing features. This is because in such cases, the intermediate features within the hierarchy do not normally possess any physical meaning or they may not be observable. In fact, the representation and retention of the physical meaning of intermediate output values is one of the most difficult problems to resolve in hierarchical fuzzy system modelling [26]. Fully addressing the learning of hierarchical fuzzy models is beyond the scope of this research, but gradient-descent techniques [28] are herein utilised to learn and optimise the parameters of the hierarchical fuzzy rules. That is, the updating of the parameters involved in the (hierarchically arranged) rules at a lower level can be estimated from the errors propagated back from the layer above it. Such a error backpropagation process is recursively performed within the hierarchy, ultimately through exploiting the error measured and back-propagated from the final system output. Alg. 1 summarises the gradient-descent-based learning process for producing a hierarchical rule base.

The complexity of this learning process can be assessed through the following analysis. Suppose that a conventional flat fuzzy interpolative reasoning system employs *K* input features and *M* membership functions per feature to describe each feature. Then, M^K rules are required in order to construct a rule base that will fully cover the problem domain. This means that a complexity of $O(M^K)$ is to be incurred in an effort to generate the flat rule base. For a *K*-input hierarchical fuzzy system, consider the worst scenario where it comprises *K*-1 low-dimensional sub-fuzzy systems, with each sub-system having two input features. Also, suppose that each input feature may take any of *M* values. From this, generating all the rules that is able to provide a full coverage of the domain involves a computational complexity of O($(K-1)M^2$). With FRI, it is not necessary to have a rule base that is so dense, but the present analysis is to assume the worse computation cost and hence, the full coverage. On top of this, additional runtime expense is needed to implement the backpropagation processes between the *K*-1 layers, denoting this by $O(bp(K-1))$. Together, the learning process of a hierarchical fuzzy rule base requires a maximum total runtime of $O((K-1)M^2) + O(bp(K-1))$.

B. Framework for Bidirectional Fuzzy Rule Interpolation

In this work, for generality, the process of forward fuzzy rule interpolation is represented by

$$B^* = f_{FRI}((A_1^*, \cdots, A_l^*, \cdots, A_M^*), (R_i, \cdots, R_t)) \tag{1}$$

where f_{FRI} expresses the interpolative reasoning process from M observed feature values, using N rules named R_i, R_t, etc. that are the closest to A_l^*, $l \in \{1, \cdots, M\}$ within a given sparse rule base, and B^* denotes the interpolated outcome.

Algorithm 1: Learning of Hierarchical Fuzzy Rule Base

1 Specify membership functions for each input feature, and normalise value domain of each intermediate feature to $[0, 1]$;

2 Denote:

- k: Number of iterations,
- K: Maximum number of iterations,
- $y(k)$: Actual output in iteration k,
- $y^{'}(k)$: Hierarchical system output in iteration k,
- $e(k)$: Error between actual output and hierarchical system output in iteration k,
- $F_{l,p}$: Sub-fuzzy system $p, p \geq 1$ at level $l, l \geq 1$,
- $e_p(k)$: Error of sub-fuzzy system $p, p \geq 1$,
- T: Number of training data,
- r: Index of training data,
- E: Accumulated error,
- η: Learning rate,
- $Q_{l-1,p}$: Number of original input features to $F_{l,p}$,
- $P_{l-1,p}$: Number of outputs from $(l-1)^{th}$ layer to $F_{l,p}$,
- $j_{i,p}$: i^{th} input for p^{th} sub-fuzzy system at l^{th} level,
- $\mu_{l,p,k}^{jk}(y_{l-1,p,k})$: Fuzzy membership function for $y_{l-1,p,k}$,
- $\upsilon_{l,p,k}^{jk}(x_{l,p,k})$: Fuzzy membership function for $x_{l,p,k}$,
- $y_{l,p}$: Output of p^{th} fuzzy subsystem at l^{th} layer,
- $y_{l,p}^{j_{1,p}j_{2,p}...j_{P_{l-1,p}}i_1 i_2...i_{Q_{l,p}}}$: Consequent of $j_{1,p}j_{2,p}...j_{P_{l-1,p}}i_1 i_2...i_{Q_{l,p}}{}^{th}$ fuzzy rule,

3 $y_{1,p} \leftarrow f_{1,p}(x_{0,p,1,...}, x_{0,p,Q_0})$;

4 $U_{l,p} \leftarrow \prod_{k=1}^{P_{l-1,p}} \mu_{l,p,k}^{jk}(y_{l-1,p,k})$;

5 $V_{l,p} \leftarrow \prod_{k=1}^{Q_{l-1,p}} \upsilon_{l,p,k}^{jk}(x_{l,p,k})$;

6 $y_{l,p} \leftarrow \sum_{j_{1,p}j_{2,p}...j_{P_{l-1,p}}i_1 i_2...i_{Q_{l,p}}} U_{l,p}V_{l,p} * y_{l,p}^{j_1 j_2...j_{P_{l-1,p}}i_1 i_2...i_{Q_{l,p}}}$;

7 $e_p(k) \leftarrow e_q(k) \times \frac{\partial y_{l,q}(k)}{\partial y_{l,p}(k)}$;

8 $k \leftarrow k+1$;

9 If $r < T$, $r \leftarrow r+1$;

10 $y_{l,p}^{j_{1,p}j_{2,p}...j_{P_{q,p}}i_1 i_2...i_{Q_{q,p}}}(k+1) \leftarrow y_{l,p}^{j_{1,p}j_{2,p}...j_{P_{q,p}}i_1 i_2...i_{Q_{q,p}}}(k) - \eta \times U_q(k)V_q(k) \times e_p(k)$;

11 Update parameters $y_{l,p}^{j_{1,p}j_{2,p}...j_{P_{l-1,p}}i_1 i_2...i_{Q_{l,p}}}$ for each iteration k, for every input-output pair (x^r, y^r);

12 $E \leftarrow \frac{1}{2} \times \sum_{r=1}^{T}(\hat{y}^r - y^r)^2$;

13 If $E > \varepsilon$ and $k < K$, go to 3;

14 End.

Similarly, B-FRI computes an unknown antecedent feature value using the closest rules of the following general form:

$$A_l^* = f_{B-FRI}(B^*, A_1^*, \cdots, A_{l-1}^*, A_{l+1}^*, \cdots, A_M^*), (R_i, \cdots, R_t)) \tag{2}$$

In the above, f_{B-FRI} represents the process of B-FRI during which the N closest rules R_i, R_t, etc. and the observed or interpolated values for the $(M-1)$ antecedent features, together with their corresponding consequent B^*, are used to backward interpolate A_l^* that is missing from the observation.

To be concise, the processes of transformation-based forward FRI (T-FRI) and transformation-based backward FRI (T-B-FRI) are outlined in Alg. 2 and Alg. 3, respectively. In implementation, as with the common practice in performing T-FRI, trapezoidal fuzzy sets (which include triangular sets as their specific cases) are adopted here for computational simplicity.

By examining these two algorithms, their computational complexities can be established. In particular, the complexity of Alg. 2 is mainly caused by the number of possible parameter combinations (ω, δ, $\underline{s}$, $\overline{s}$, and m). The weight ω is computed in relation to all N closest rules, thereby having a complexity of $O(N)$. The handling of the other four parameters: δ, $\underline{s}$, $\overline{s}$, and m each incurs the same complexity as handling a single ω value. Therefore, the overall complexity of Alg. 2 is still $O(N)$. Similarly, for Alg. 3, the complexity can be estimated with regards to the number of all missing antecedent features, say L, and that of all N closest rules. Particularly, the parameter combinations (ω, δ, $\underline{s}$, $\overline{s}$, and m) of $(M-L)$ antecedent features need to be calculated. Thus, Alg. 3 incurs a considerably higher complexity: $O((M-L)NL) \cdot O(\text{FRI})$, where $O(\text{FRI})$ represents the complexity of the underlying FRI process.

III. Refinement of Hierarchical Fuzzy Rule Base through HB-FRI

Generally speaking, a fuzzy logic-based approximate reasoning system consists of an inference engine and a fuzzy rule base. It performs intelligent decision-making, typically by either firing the rules in the rule base which match a given observation or carrying out interpolation if none of the rules match the observation, not even partially. Although the original rule base may be assumed to be consistent, i.e., the same input feature values are expected to always lead to the same inferred outcome, any modification or new addition to the rule base during the rule interpolation process may introduce contradictory to certain rules. This is of particular significance to systems that involve dynamic learning [18], whilst such learning is a common requirement for dealing with novel problems, where typically only a rather sparse rule base is available initially. It is therefore, clearly desirable to avoid such inconsistencies. In this section HB-FRI is employed to implement a strategy that helps refine the rule base in an effort to remove rule inconsistency.

Algorithm 2: Transformation-based Forward Interpolation ($\mathbb{U}$, $(A_1^*, \cdots, A_l^*, \cdots, A_M^*)$)

$\mathbb{U}$: given rule base;

$R \in \mathbb{U}$ with antecedent $\{A_k, k = 1, 2, \cdots, M\}$;

$O = \{A_1^*, \cdots, A_k^*, \cdots, A_M^*\}$, being observation;

$Rep(A)$: representative value of fuzzy set A;

d_{rep}: distance between representative values of two fuzzy sets;

d_f: distance between two fuzzy sets;

$d(O, R)$: distance between observation O and rule R;

$Rep(A) \leftarrow \frac{a_0 + \frac{a_1 + a_2}{2} + a_3}{3}$;

$range_k \leftarrow sup_k - inf_k$;

$d_f(A_k, A_k^*) \leftarrow \frac{d_{rep}(Rep(A_k), Rep(A_k^*))}{range_k}$;

$d(O, R) \leftarrow \sqrt{\sum_{k=1}^{M} d(A_k, A_k^*)^2}$, omitting normalising term;

$\mathbb{R} \leftarrow \{R_i | i \in \{1, \ldots, N\}\}$, N being number of rules whose $d(O, R_i)$ are first N smallest;

R_i: IF x_1 is A_1^i, $\cdots$, and x_k is A_k^i, $\cdots$, and x_M is A_M^i, THEN y is B^i, $R_i \in \mathbb{R}$;

$\omega'_{A_k^i} \leftarrow 1/(d_f(A_k^i, A_k^*) + 1)$;

$\omega_{A_k^i} \leftarrow \frac{\omega'_{A_k^i}}{\sum_{i=1}^{N} \omega'_{A_k^i}}$, $\omega_{A_k^i}$, being normalised displacement factor;

$A_k^\dagger \leftarrow \sum_{i=1}^{N} \omega_{A_k^i} A_k^i$, $A_k^\dagger$, being intermediate fuzzy terms;

$A_k^\dagger \leftarrow \sum_{i=1}^{N} \omega_{A_k^i} A_k^i$;

$\delta_{A_k} \leftarrow d_f(A_k^*, A_k^\dagger)$;

$A_k' \leftarrow A_k^\dagger + \delta_{A_k} range_{A_k}$;

$\delta_B \leftarrow \frac{1}{M} \sum_{k=1}^{M} \delta_{A_k}$;

$\omega_{B^i} \leftarrow \frac{1}{M} \sum_{k=1}^{M} \omega_{A_k^i}$;

$B' \leftarrow \sum_{i=1}^{N} \omega_{B^i} B^i + \delta_B range_B$;

$\underline{s}_{A_k} \leftarrow \frac{a_3'' - a_0''}{a_3' - a_0'}$;

$\overline{s}_{A_k} \leftarrow \frac{a_2'' - a_1''}{a_2' - a_1'}$;

$\underline{s}_B \leftarrow \frac{1}{M} \sum_{k=1}^{M} \underline{s}_{A_k}$;

$\frac{1}{M} \sum_{k=1}^{M} \overline{s}_{A_k}$;

$$\mathbb{S} \leftarrow \begin{cases} \dfrac{\frac{a_2'-a_1'}{a_3'-a_0'} - \frac{a_2-a_1}{a_3-a_0}}{1-\frac{a_2'-a_1'}{a_3'-a_0'}} & if\ \overline{s} \geq \underline{s} \geq 0 \quad , \mathbb{S} \in [0,1] \\ \dfrac{\frac{a_2'-a_1'}{a_3'-a_0'} - \frac{a_2-a_1}{a_3-a_0}}{\frac{a_2'-a_1'}{a_3'-a_0'}} & if\ \underline{s} \geq \overline{s} \geq 0 \quad , \mathbb{S} \in [-1,0] \end{cases};$$

$$\overline{s}_B \leftarrow \begin{cases} \frac{\underline{s}_B \mathbb{S}}{\overline{s}_B} - \underline{s}_B \mathbb{S} + \underline{s}_B & if\ \overline{s}_B \geq \underline{s}_B \geq 0 \\ \underline{s}_B \mathbb{S} & if\ \underline{s}_B \geq \overline{s}_B \geq 0 \end{cases};$$

$$m_{A_k} \leftarrow \begin{cases} \frac{3(a_0 - a_0'')}{a_1'' - a_0''}, & a_0 \geq a_0'' \\ \frac{3(a_0 - a_0'')}{a_3'' - a_2''}, & \text{otherwise} \end{cases};$$

$m_B \leftarrow \frac{1}{M} \sum_{k=1}^{M} m_{A_k}$;

$B^* \leftarrow T(B', \underline{s}_B, \overline{s}_B, m_B)$.

Algorithm 3: Transformation-based Backward Interpolation ($\mathbb{U}$, $(B^*, A_1^*, \cdots, A_{l-1}^*, A_{l+1}^*, \cdots, A_M^*)$)

$\mathbb{U}$: given rule base;
$R \in \mathbb{U}$ with antecedent $\{A_k, k = 1, 2, \cdots, M\}$ and consequent B;
$O' = (B^*, A_1^*, \cdots, A_{l-1}^*, A_{l+1}^*, \cdots, A_M^*)$, given consequent B^* and observation with missing antecedent term A_l^*;
$w_B \leftarrow \sum_{k=1}^{M} w_{A_k} = 1,\ w_{A_k} = \frac{1}{M},\ k \in \{1, \ldots, M\}$;
$d(O', R) \leftarrow \sqrt{w_B d_f^2(B, B^*) + \sum_{k=1,\ k\neq l}^{M} (w_{A_k} d_f^2(A_k, A^*))}$, omitting normalising term;
$\mathbb{R} \leftarrow \{R_i | i \in \{1, \ldots, N\}\}$, N being number of rules whose $d(O', R_i)$ are first N smallest;
$\omega'_{A_k^i} \leftarrow 1/(d_f(A_k^i, A_k^*) + 1)$;
$\omega_{B^i} \leftarrow \frac{1}{M} \sum_{k=1}^{M} \omega_{A_k^i}$;
$\omega_{A_l^i} \leftarrow M\omega_{B^i} - \sum_{k=1,\ k\neq l}^{M} \omega_{A_k^i}$;
$\delta_{A_l} \leftarrow M\delta_B - \sum_{k=1,\ k\neq l}^{M} \delta_{A_k}$;
$A_l^\dagger \leftarrow \sum_{i=1}^{N} \omega_{A_l^i} A_l^i$;
$A_l' \leftarrow A_l^\dagger + \delta_{A_l} range_{A_l}$;
$\underline{s}_{A_l} \leftarrow M\underline{s}_B - \sum_{k=1,\ k\neq l}^{M} \underline{s}_{A_k}$;
$\overline{s}_{A_l} \leftarrow M\overline{s}_B - \sum_{k=1,\ k\neq l}^{M} \overline{s}_{A_k}$;
$m_{A_l} \leftarrow Mm_B - \sum_{k=1,\ k\neq l}^{M} m_{A_k}$;
$\mathbb{S}_{A_l} \leftarrow M\mathbb{S}_B - \sum_{k=1,\ k\neq l}^{M} \mathbb{S}_{A_k}$;
$\overline{s}_{A_l} \leftarrow \begin{cases} \frac{\underline{s}_{A_l} * \mathbb{S}_{A_l}}{\overline{s}_{A_l}} - \underline{s}_{A_l} * \mathbb{S}_{A_l} + \underline{s}_{A_l} & if\ \overline{s}_{A_l} \geq \underline{s}_{A_l} \geq 0 \\ \overline{s}_{A_l} * \mathbb{S}_{A_l} & if\ \underline{s}_{A_l} \geq \overline{s}_{A_l} \geq 0 \end{cases}$;
$A_l^* = T(A_l', \underline{s}_{A_l}, \overline{s}_{A_l}, m_{A_l})$.

Without losing generality, suppose that two inconsistent rules are represented in the following form:

R_i: IF x_1 is A_1^i, $\cdots$, x_k is A_k^i, $\cdots$, x_M is A_M^i,
THEN y is B^i

R_i': IF x_1 is A_1^i, $\cdots$, x_k is A_k^i, $\cdots$, x_M is A_M^i,
THEN y is $B^{i'}$

where $x_k, k = 1, 2, \cdots, M$, denote the input variables, and $B^i \neq B^{i'}$.

The basic idea of the proposed refinement process is to use the average bias (aka. deviation) of the inconsistent rules to rectify the inferred or interpolated consequent. Alg. 4 details this process. Firstly, each antecedent feature value A_k or $A_k^{'}$ is backward interpolated from R_i or $R_i{'}$ respectively, using *B-FRI*. Secondly, the bias between each antecedent value A_k or $A_k^{'}$ and its corresponding observed value A_k^* is calculated. Then, the average bias over each of the two inconsistent rules in question is computed. Finally, from the resulting biases the consequent (of each of the two otherwise inconsistent rules) is corrected, resulting in the desired consequent value for the newly constructed rule that replaces the original two, which retains the same antecedent part as its originals.

Algorithm 4: Refinement of Fuzzy Rule Base via HB-FRI

Given two inconsistent rules R_i and $R_i{'}$;

R_i: IF x_1 is A_1^i, $\cdots$, x_k is A_k^i, $\cdots$, x_M is A_M^i, $(k = 1, 2, \cdots, M)$, THEN y is B^i;

$R_i{'}$: IF x_1 is A_1^i, $\cdots$, x_k is A_k^i, $\cdots$, x_M is A_M^i, $(k = 1, 2, \cdots, M)$, THEN y is $B^{i'}$; where $B^i \neq B^{i'}$.

$A_k \leftarrow (B - FRI)\{R_i, \mathbb{R} \leftarrow \{R_i | i \in \{1, \ldots, N\}\}$ N being number of rules whose d(O,Ri) are first N smallest;

$A_k^{'} \leftarrow (B - FRI)\{R_i{'}, \mathbb{R}' \leftarrow \{R_i' | i \in \{1, \ldots, N\}\}$ N being number of rules whose d(O,Ri') are first N smallest;

$\delta_k \leftarrow \frac{|Rep(A_k) - Rep(A_k^*)|}{max_{A_k} - min_{A_k}}$, as bias between an antecedent value A_k and its corresponding observed antecedent value A_k^*;

$\delta_k{'} \leftarrow \frac{|Rep(A_k{'}) - Rep(A_k^*)|}{max_{A_k} - min_{A_k}}$, as bias between an antecedent value $A_k^{'}$ and its corresponding observed antecedent value A_k^*;

$\overline{\delta} \leftarrow \frac{1}{M} \sum_{k=1}^{M} \delta_k$, as average bias for rule R_i;

$\overline{\delta'} \leftarrow \frac{1}{M} \sum_{k=1}^{M} \delta_k$, as average bias for rule $R_i{'}$;

$y \leftarrow Rep(B^i) \times \frac{\overline{\delta'}}{\overline{\delta} + \overline{\delta'}} + Rep(B^{i'}) \times \frac{\overline{\delta}}{\overline{\delta} + \overline{\delta'}}$, as refined consequence.

This algorithm presents the procedures that are required to implement both inconsistent rule removal and repair, without explicitly showing the running of HB-FRI. Nonetheless, HB-FRI is implicitly utilised within its specification and implementation since it produces the interpolated rules using the identified closest rules from the rule base, given an observation. As such, the computational complexity of this algorithm mainly depends on M, the number of antecedent features. The run-time of calculating δ_k and $\delta_k{'}$ is $O(M) \cdot O(\text{B-FRI})$, where $O(\text{B-FRI})$ denotes the complexity of the B-FRI process itself. Thus, for the entire inconsistency-removal process, the total run-time complexity is $O(M) \cdot O(\text{B-FRI}) + 2 \cdot O(\text{FRI})$.

IV. Experimental Results and Discussions

A. Evaluation with Numerical Function Approximation

1) Problem case: As the first example to demonstrate the efficacy of the proposed work, a function approximation problem is utilised here, focussing on the problem of rule inconsistency removal. The problem has three input features and one decision (output) attribute:

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

Each input variable can take one of six fuzzy sets defined using trapezoidal membership functions. The rule base required to perform the experiments is generated, assuming a uniform distribution to cover the problem domain $U = [1, 6]^3$.

Note that in real applications, the rule base may be provided by the domain experts or learned from historical data, or created from a mixture of both [5]. Here, the assumption for rule generation using prescribed fuzzy values is made purely for illustration simplicity. This does not affect the explanation of the underlying ideas. What is important is to show the potential of interpolative reasoning with a hierarchical sparse rule base. This is assured by employing only a small part of the fully constructed rule set as the sparse rule base in the example, deliberately leaving out many of the initially learned rules. Those rules which are left out are used as the ground truth to evaluate the accuracy of any interpolated result since they are learned using sufficient training samples read off the function (and hence, the associated heating search outcomes can cover the full problem space). Obviously, for any real-world application, had there been such a full rule base, there would not be a need to utilise rule interpolation.

To reveal the effectiveness of the proposed technique of rule inconsistency removal, intermediate rules that are produced during the transformation-based FRI process are collated and subsequently promoted for inclusion into the original rule base, following the advanced method of [18].

2) Rule refinement: Table I lists a sub-rule base to support the illustration of solving the present problem, where for simplicity, all fuzzy values have been denoted as a shorthand using their corresponding representative values (which approximately reflects the geometrical properties of the original fuzzy sets) [10]. In the following discussion, without causing confusion and unless otherwise stated, the representative value of a fuzzy set associated with a given feature is simply referred to as a value of that feature.

As can be observed, in Table I, $Rule_6$ and $Rule_{11}$ involve identical antecedent values but different consequent values. As such, these two form a pair of inconsistent rules and this inconsistency needs to be removed to ensure subsequent inference consistency. The proposed method works well in this case,

TABLE I

TWO INCONSISTENT RULES IN A SUB-RULE BASE

No. of Rules	x_1	x_2	x_3	y
Rule 1	3.811486	5.183552	5.419737	10.3972497
Rule 2	5.858488	1.448528	4.154679	17.88335981
Rule 3	3.990305	2.566741	5.598877	11.9899926
Rule 4	0.369646	3.392652	5.194505	3.948987083
Rule 5	0.438950	0.396712	5.820496	18.1005066
Rule 6	5.656367	5.368069	5.674704	13.8392177
Rule 7	0.512235	3.007569	5.667789	4.50420492
Rule 8	1.991630	0.445464	4.040010	22.84119918
Rule 9	4.823559	2.682018	2.387356	14.74721597
Rule 10	3.253371	3.556458	4.797134	10.11281312
Rule 11	5.656367	5.368069	5.674704	10.565988

as described below.

Applying B-FRI (or lines 4 and 5 in Alg. l4) leads to the determination of the antecedent values $Rep(A_k)$ and $Rep(A'_k)$ $(k = 1, 2, 3)$ from R_6 and R_{11}. Then, running lines 6 and 7 of the algorithm results in the estimated values for the biases between each corresponding pair of antecedent values in these two rules. Following this step, using lines 8 and 9 gives the average biases $\overline{\delta}$ and $\overline{\delta'}$, as shown in Table II and Table III, respectively. From these results, application of line 10 of Alg. 4 returns the representative value (13.26370597) of the output of the refined rule. Remarkably, this outcome is very close to the underlying ground truth value (13.2392175). This implies that the two inconsistent rules $Rule_6$ and $Rule_{11}$ are now replaced by a new rule, as presented Table IV. Such an accurate result shows the significant potential of the proposed approach in refining the rule base that would otherwise contain inconsistent rules.

TABLE II

ANTECEDENT DEVIATION FOR $Rule_6$

δ_1	δ_2	δ_3	$\overline{\delta}$
0.132227319	0.061004378	0.035891982	0.07637456

TABLE III

ANTECEDENT DEVIATION FOR $Rule_{11}$

δ'_1	δ'_2	δ'_3	$\overline{\delta'}$
0.55324718	0.27989492	0.24087783	0.35800664

TABLE IV

REFINED SUB-RULE BASE

No. of Rules	x_1	x_2	x_3	y
Rule 1	3.811486	5.183552	5.419737	10.3972497
Rule 2	5.858488	1.448528	4.154679	17.88335981
Rule 3	3.990305	2.566741	5.598877	11.9899926
Rule 4	0.369646	3.392652	5.194505	3.948987083
Rule 5	0.438950	0.396712	5.820496	18.1005066
Rule 6	5.656367	5.368069	5.674704	13.277335
Rule 7	0.512235	3.007569	5.667789	4.50420492
Rule 8	1.991630	0.445464	4.040010	22.84119918
Rule 9	4.823559	2.682018	2.387356	14.74721597
Rule 10	3.253371	3.556458	4.797134	10.11281312

B. Comparative Analysis against Standard T-FRI

1) Experimental setting: The proposed HB-FRI is in this subsection applied to four benchmark problems of time series prediction [4], [11], in order to further evaluate its performance through comparison with the use of the standard T-FRI method. Table V summarises the features of these datasets. For simplicity, the fuzzy values of all input features addressed within this experimental study are represented by triangular membership functions. For consistency, the number of membership functions is set as with the previous practice, that is, six triangular functions are defined for each input feature across all data sets. As different features have their own underlying value domains in reality, they are normalised to the common scale of 0 to 1 to ease implementation and comparison.

Note that there exists an underlying difference in the representation of rule base structures between the two methods compared, with T-FRI employing a rule base consisting of flat rules only whilst HB-FRIuses hierarchical rules. Reflecting such a fundamental difference, the rule bases are learned from each given dataset using two distinct learning mechanisms. In particular, the rules used for running T-FRI are learned with the popular method of [27] while those used by HB-FRI are generated using Alg. 1 in Section II-A. Nonetheless, both forms of rule bases are produced from the same given dataset per problem, using the

TABLE V
DATASETS USED FOR PREDICTION

Dataset	Number of Features	Number of Instances
Chemical Process Concentration Readings Prediction [4]	3	194
Chemical Process Temperature Readings Prediction [4]	3	233
Gas Furnace Prediction [4]	6	293
Mackey-Glass Chaotic Time Series Prediction [11]	4	3000

aforementioned partitions.

To support this investigation, ensuring that the learned rule bases are sparse and contain inconsistent rules, random changes to each rule base returned by the learning methods are made. Particularly, regarding rulebase sparsity, for T-FRI each learned rule base has 30% of the originally learned rules removed, and for HB-FRI each rule base is learned with 30% of the original raw data removed. Regarding rulebase inconsistency, each learned rule base is set to includes a fixed percentage of inconsistent rules, which are artificially added so that the same antecedents may have different consequences. To have a wider range of comparison, three sets of experiments are carried out, involving the containment of 5%, 10% and 20% of inconsistent rules, respectively. Any bias between the consequent of an artificially introduced rule and that of its original counterpart is randomly set to be within 10%. Both the removal of the originally learned rules or that of the original data, and the addition of inconsistent rules are randomly implemented also, with a uniform distribution across each problem domain.

For fair comparison, both the standard T-FRI and the proposed HB-FRI are herein assisted by the use of the compositional rule of inference [31], in an effort to gain reasoning efficiency for those matched observations. The performance on prediction accuracy is measured by the conventional root mean square error (RMSE), defined by

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(y_i^* - y_i^{'})^2}{N}} \tag{3}$$

where y_i^* is the predicted value of the i^{th} testing sample and $y_i^{'}$ is the i^{th} original outcome in the dataset concerned. To avoid any potential influence of noise on the forecasting quality, the results of experiments presented below are average values verified by ten times fivefold cross validation per dataset.

2) Experimental results: The average prediction RMSEs are shown in Table VI. It can be seen that for each given amount of inconstant rules involved, the accuracy of the proposed method is systematically greater than that of T-FRI across all datasets. Whilst the improvements gained by HB-FRI over T-FRI is relatively small when the rule base contains just 5% inconstant rules, the improvements are much more

TABLE VI

AVERAGE RMSE IN 10 × 5-FOLD CROSS VALIDATION WITH 5%, 10% OR 20% INCONSISTENT RULES

	T-FRI			HB-FRI		
Dataset	5%	10%	20%	5%	10%	20%
Chemical Process Concentration Readings Prediction	0.389	0.495	0.844	0.323	0.324	0.333
Chemical Process Temperature Readings Prediction	0.498	0.550	0.746	0.404	0.406	0.420
Gas Furnace Prediction	0.736	0.799	0.964	0.623	0.628	0.627
Mackey-Glass Chaotic Time Series Prediction	0.128	0.259	0.634	0.041	0.044	0.066
Average	0.438	0.526	0.797	0.348	0.351	0.362

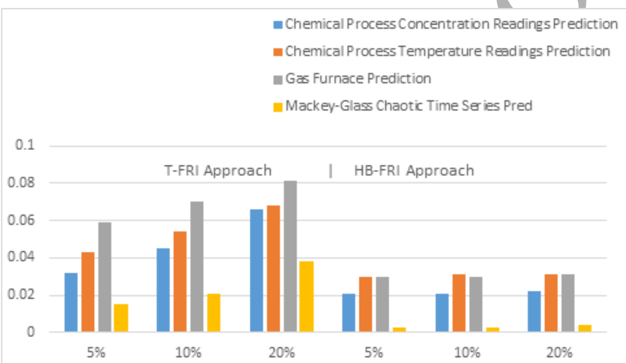


Fig. 2. Standard deviation of obtained average RMSEs

remarkable as the percentage of inconstant rules increases. The overall average of the performances measured across all four datasets as per the bottom line of Table VI further reflects the significant improvement brought forward by the proposed method. This positively shows the potential of HB-FRI in performing interpolative reasoning with the ability of correcting inconsistent rules.

Figure 2 shows the standard deviation (SD) of the RMSE measures, indicating how each method's performance varies in response to different percentages of inconsistent rules. The smaller the standard deviation, the more robust the corresponding method. It can be seen from this figure that the higher the proportion of inconsistent rules is, the better gain is achieved using the proposed HB-FRI as compared to the standard T-FRI. This sufficiently shows the effectiveness of HB-FRI in dealing with prediction problems with sparse and yet inconsistent knowledge.

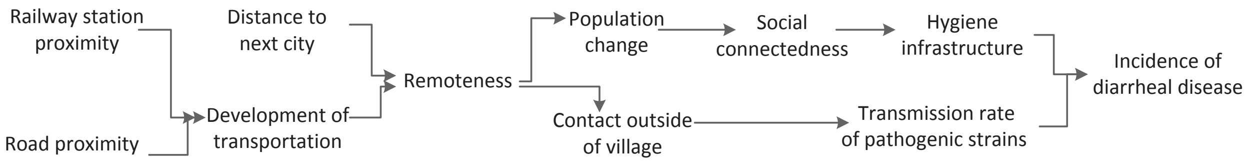


Fig. 3. Causal network model used for decision-making

C. Application to Diarrheal Disease Prediction

1) Problem specification: Environmental changes and their potential impacts upon the society, especially upon the public health, are major concerns for both governments and the public, globally. For example, much effort has been made to encourage the development of cause-effect relation models concerning environmental change events and their influences upon various diseases and disease propagation. In particular, much attention has recently been drawn towards carefully addressing the issue of decision-making and policy formation for problems such as diarrheal disease prediction and prevention [17], [19]. Having taken notice of this, and by following previous studies in this area [8], [30], the present application-oriented experimental investigation looks into the particular problem of predicting diarrheal disease rate in a remote countryside village, through approximate reasoning that utilises HB-FRI.

Taking the northern coastal region of Ecuador as an example, building a new road or railway in previously inaccessible areas may affect the epidemiology of diarrheal diseases [1], [9]. Close proximity of newly constructed roads can lead to the increase of the contact between the residents of the village and those outside the village. This in turn, can raise the rate of introduction of pathogens, which can then cause the diarrheal disease rate to increase. As a demonstration case, this study focusses on part of a much larger problem, where the occurrence of diarrheal disease is dependent upon village remoteness (which is influenced by its distance to the closest city) and the village's connectivity level to public transportation systems (which is determined by the connectivity situation to the nearest railway station and road) [8]. The complete causal relation model considered herein is shown in Fig. 3.

2) Application of hierarchical bidirectional fuzzy interpolation: The nature of the above causal model is hierarchical. Based on human interpretation of this observation and supported by Alg. 1, a hierarchical sparse rule base can be obtained (by supervised learning as outlined in Section II-A, whilst as indicated previously, the exact learning process for the derivation of this rule base is not a concern of the present

TABLE VII

SUB-RULE BASE 1

Variable	x_1	x_2	x_3
Meaning	Railway station proximity	Road proximity	Development of transportation
rule 1	(0.02,0.04,0.06,0.08)	(0.18,0.20,0.22,0.24)	(0.46,0.48,0.50,0.52)
rule 2	(0.28,0.30,0.32,0.34)	(0.39,0.41,0.43,0.45)	(0.62,0.64,0.66,0.68)

work). The result is shown in Fig. 4, where x_1, x_2, x_4 are observable input features and x_{11} is the output attribute of the overall system. All the other features are regarded as internal variables and are denoted by I_* with $* \in \{1, \ldots, 7\}$ for easy cross-referencing. From this, eight sub-rule bases are constructed as given in Tables VII-XIV, including those rules which flank an observation or a previously interpolated outcome. As with the common practice in the relevant literature, all fuzzy sets used to encode rule antecedent feature values are represented using trapezoidal membership functions.

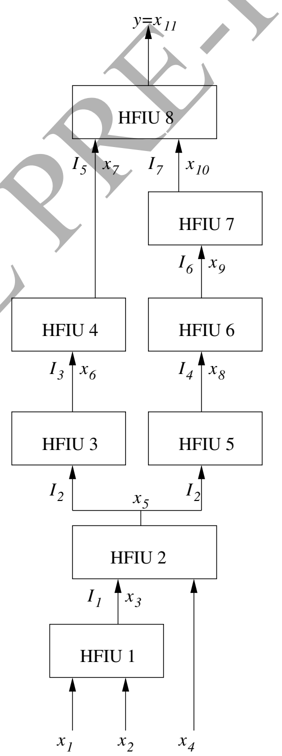


Fig. 4. Hierarchical fuzzy rule model

The task here is to predict the diarrheal disease rate of a certain village from several pieces of information obtained by different agencies. Such information is regarded as observations, respectively represented by $x_1 = (0.16, 0.18, 0.20, 0.22), x_2 = (0.34, 0.36, 0.38, 0.40), x_4 = (0.65, 0.67, 0.69, 0.71),$

TABLE VIII

SUB-RULE BASE 2

Variable	x_3	x_4	x_5
Meaning	Development of transportation	Distance to next city	Remoteness
rule 1	(0.52,0.54,0.56,0.58)	(0.52,0.54,0.56,0.58)	(0.41,0.43,0.45,0.47)
rule 2	(0.85,0.87,0.89,0.91)	(0.82,0.84,0.86,0.88)	(0.72,0.74,0.76,0.78)

TABLE IX

SUB-RULE BASE 3

Variable	x_5	x_6
Meaning	Remoteness	Contact outside of village
rule 1	(0.27,0.29,0.31,0.33)	(0.62,0.64,0.66,0.68)
rule 2	(0.58,0.60,0.62,0.64)	(0.30,0.32,0.34,0.36)

TABLE X

SUB-RULE BASE 4

Variable	x_6	x_7
Meaning	Contact outside of village	Transmission rate of pathogenic strains
rule 1	(0.38,0.40,0.42,0.44)	(0.46,0.48,0.50,0.52)
rule 2	(0.70,0.72,0.74,0.76)	(0.65,0.67,0.69,0.71)

TABLE XI

SUB-RULE BASE 5

Variable	x_5	x_8
Meaning	Remoteness	Population change
rule 1	(0.39,0.41,0.43,0.45)	(0.60,0.62,0.64,0.66)
rule 2	(0.62,0.64,0.66,0.68)	(0.30,0.32,0.34,0.36)

TABLE XII

SUB-RULE BASE 6

Variable	x_8	x_9
Meaning	Population change	Social connectedness
rule 1	(0.46,0.48,0.50,0.52)	(0.52,0.54,0.56,0.58)
rule 2	(0.68,0.70,0.72,0.74)	(0.20,0.22,0.24,0.26)

TABLE XIII

SUB-RULE BASE 7

Variable	x_9	x_{10}
Meaning	Social connectedness	Hygiene infrastructure
rule 1	(0.28,0.30,0.32,0.34)	(0.26,0.28,0.30,0.32)
rule 2	(0.55,0.57,0.59,0.61)	(0.61,0.63,0.65,0.67)

TABLE XIV

SUB-RULE BASE 8

Variable	x_7	x_{10}	x_{11}
Meaning	Transmission rate of pathogenic strains	Hygiene infrastructure	Incidence of diarrheal disease
rule 1	(0.30,0.32,0.34,0.36)	(0.36,0.38,0.40,0.42)	(0.18,0.20,0.22,0.24)
rule 2	(0.60,0.62,0.64,0.66)	(0.58,0.60,0.62,0.64)	(0.68,0.70,0.72,0.74)

and $x_8 = I_4 = (0.54, 0.56, 0.58, 0.60)$, as summarised in Table XV.

For the sparse rule base available to the problem at hand, none of the rules match the above observations. Therefore, it is unable to resolve the problem by ordinary approximate reasoning techniques (such as via applying the compositional rule of inference alone). Thankfully, FRI can help: Table XVI lists the intermediate values and the final output running HB-FRI over the hierarchical model. Note that the intermediate value $x_8 = I_4$ is provided as part of the observations and is not produced by the use of HB-FRI. However, if the intermediate value I_4 were not known, it could be approximately computed by HB-FRI. Using the computed intermediate value (instead of the given observation) the final output of the hierarchical fuzzy model is presented in Table XVII. Both of these two versions of the final result (which is the predicted fuzzy value for the diarrheal disease rate) are depicted in Fig. 5.

The original value domain defining the variable x_{11} is $[0\%, 10\%]$. Mapping the inferred value back to this domain gives the predicted diarrheal disease rate being 4.1% or 5.8% for the studied village.

TABLE XV

OBSERVATIONS GIVEN

Variable	Fuzzy set
x_1	(0.16, 0.18, 0.20, 0.22)
x_2	(0.34, 0.36, 0.38, 0.40)
x_4	(0.65, 0.67, 0.69, 0.71)
$x_8 = I_4$	(0.54, 0.56, 0.58, 0.60)

TABLE XVI

INTERMEDIATE VALUES AND FINAL OUTPUT WITH x_8 (INTERMEDIATE VARIABLE I_4) IMPORTED EXTERNALLY

Intermediate variables	Fuzzy set	Representative value
I_1	(0.46,0.58,0.62,0.64)	0.57
I_2	(0.49,0.51,0.55,0.57)	0.53
I_3	(0.38,0.40,0.44,0.46)	0.42
I_4 (as observed)	(0.54,0.56,0.58,0.60)	0.57
I_5	(0.46,0.48,0.52,0.54)	0.50
I_6	(0.40,0.42,0.46,0.48)	0.44
I_7	(0.41,0.43,0.47,0.49)	0.45
Final output x_{11}	(0.37,0.39,0.43,0.45)	0.41

TABLE XVII

INTERMEDIATE VALUES AND FINAL OUTPUT WITH x_8 (INTERMEDIATE VARIABLE I_4) INFERRED WITHIN *HB-FRI* SYSTEM

Intermediate variables	Fuzzy set	Representative value
I_1	(0.46,0.58,0.62,0.64)	0.57
I_2	(0.49,0.51,0.55,0.57)	0.53
I_3	(0.38,0.40,0.44,0.46)	0.42
I_4	(0.45,0.47,0.51,0.53)	0.49
I_5	(0.46,0.48,0.52,0.54)	0.50
I_6	(0.51,0.53,0.57,0.59)	0.55
I_7	(0.56,0.58,0.62,0.64)	0.60
Final output x_{11}	(0.54,0.56,0.60,0.62)	0.58

Which of these two outcomes holds depends on whether the observed x_8 is directly used during the prediction process or the inferred intermediate variable I_4 is used (without disrupting the internal inference mechanism). This result compares well with that which is achievable by the most advanced (and complicated) adaptive fuzzy rule interpolation mechanism in the literature [29], [30], where a result of approximately 5.5% is returned for the same problem. Yet, the existing work requires substantially more computational effort to reach this comparable outcome as it relies on the employment of an expensive model-based diagnostic engine [6] to correct the prediction errors.

3) Significance of HB-FRI for decision-making: The above practical application demonstrates that HB-FRI can be of great potential to provide valuable suggestions in decision support, for both governments and the general public. In general, it is difficult to track the transmission rate of pathogenic strains (x_7) in real time, and it is also difficult to accurately estimate the population change (x_8). Since the area

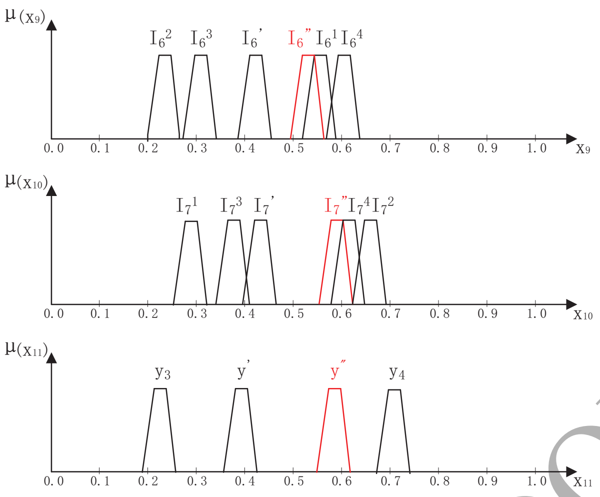


Fig. 5. Results by hierarchical bidirectional fuzzy interpolative reasoning

concerned is geographically remote and economically backward the employment of any sophisticated technologies (e.g., big data analysis) that would require the acquisition of a substantial amount of data to estimate such key values is not easy. With the use of a hierarchical model that is supported by bidirectional fuzzy rule interpolation, through examining the incidence of diarrheal disease (x_{11}, which can be measured timely in a given location), the transmission rate of pathogenic strains and population change can now be assessed. The resulting estimated changes can be utilised to alert the local government in a timely manner, thereby significantly increasing the effectiveness of any subsequent decision making process. In particular, transmission rate of pathogenic strains and population change are both controllable parameters which may be adjusted in order to model and deploy practical mechanisms that help minimise the occurrence of diarrheal diseases.

More concretely speaking, aided by the proposed bidirectional fuzzy rule interpolation, useful information can indeed be provided to the relevant government in support of its decision-making. For example, in order to reduce the incidence of diarrheal disease (x_{11}) from say, *Middle High* $(0.72, 0.74, 0.76, 0.78)$ to *Middle Low* $(0.38, 0.40, 0.50, 0.52)$, according to the proposed HB-FRI approach, the value of transmission rate of pathogenic strains (x_7) needs to be changed from the current *Middle High* $(0.62, 0.64, 0.66, 0.68)$ to *Middle Low* $(0.33, 0.35, 0.37, 0.39)$. This means that the local government needs to take medical or administrative measures to control the spread of pathogenic strains. At the same time, the population change rate (x_8) needs to be regulated to go from the current *Middle High* $(0.69, 0.71, 0.73, 0.75)$ down to *Middle* $(0.53, 0.55, 0.57, 0.59)$, which in turn, implies that population changes due to immigration need to be stopped. This shows that the proposed HB-FRI approach is very helpful to interpolate the knowledge that is useful to estimate unobservable domain features that are crucial for decision-making (say, at the time of planning). Such variables can then be subsequently controlled and adjusted through government administration. As such, the risk of diarrheal disease in the area concerned may be significantly reduced, benefitting the general public.

V. Conclusion

This paper has presented a theoretical framework for hierarchical bidirectional fuzzy rule interpolation (HB-FRI), including computational complexity analyses for the algorithms introduced. The work enables unknown antecedent feature values to be inferred in a manner involving both forward and backward rule interpolation. It facilitates an effective way of coping with insufficient information or sparse knowledge that may appear in automated decision-making. More importantly, this paper has proposed an automated method for restoring consistency in a sparse rule base through the use of HB-FRI. The work has been verified with a range of problems, including: numerical function approximation, time series prediction, and real-world decision-making application. The particular application investigation has presented a clear case for the potential benefits of utilising HB-FRI to aid in decision-making when only limited knowledge is available.

Whilst very promising, the proposed approach does not give due attention to problems that require consideration of automated dynamic update of the rule base. Also, the issue of its general scalability remains open, requiring further assessment, both theoretically and empirically. There indeed exist significant opportunities for further development. For instance, whether this approach could benefit from a full integration with the most recent advance in dynamic fuzzy rule interpolation [18], in order to cope with dynamic rule changes, appears to be a natural next step to conducting further research. In addition, how flexibly a hierarchical fuzzy model may run in response to the use of different intermediate features requires experimental investigation. Another important piece of work is to introduce weights onto antecedent features as per the most recent work of [15], thereby allowing for the selection of least number of the nearest neighbouring rules for interpolation [16], increasing the overall reasoning efficacy. Last but not least, instead of encoding all fuzzy sets with pre-specified trapezoidal form it is interesting to examine whether the employment of fuzzy values learned by data clustering tools (e.g., those introduced in [2]) would entail more accurate interpolation.

Acknowledgements

The work described in this study was supported by the National Natural Science Foundation of China (No. 61873043), the Natural Science Foundation of Chongqing, China,(No. cstc2019jcyjmsxmX0355, No. cstc2018jcyjAX0048), the Science and Technology Research Program of Chongqing Municipal Education Commission (No. KJ1713329), the Foundation of Doctor and Professor's Research Project of Chongqing University of Science and Technology (No. CK2016B02), and the Education Teaching Reform Research Key Project of Chongqing University of Science and Technology (No. 201614).

References

[1] A. Bebbington and T. Perreault. Social capital, developmentand access to resources in highland ecuador*. *Economic geography*, 75(4):395–418, 1999.

[2] T. Boongoen, C. Shang, N. Iam-On, and Q. Shen. Extending data reliability measure to a filter approach for soft subspace clustering. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 41(6):1705–1714, 2011.

[3] T. Boongoen, Q. Shen, and C. Price. Disclosing false identity through hybrid link analysis. *Artificial Intelligence and Law*, 18(1):77–102, 2010.

[4] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung. Time series analysis: Forecasting and control, 5th edition. *Journal of the Operational Research Society*, 22(2):199–201, 2015.

[5] T. Chen, C. Shang, P. Su, and Q. Shen. Induction of accurate and interpretable fuzzy rules from preliminary crisp representation. *Knowledge-Based Systems*, 146:152–166, 2018.

[6] J. deKleer and B. Williams. Diagnosing multiple faults. *Artificial Intelligence*, 32(1):97–130, 1987.

[7] D. Dubois, J. Lang, and H. Prade. Fuzzy sets in approximate reasoning, part 2: logical approaches. *Fuzzy Sets & Systems*, 40(1):203–244, 1991.

[8] J.N.S. Eisenberg, W. Cevallos, K. Ponce, K. Levy, S.J. Bates, J.C. Scott, A. Hubbard, N. Vieira, P. Endara, M. Espinel, et al. Environmental change and infectious disease: how new roads affect the transmission of diarrheal pathogens in rural ecuador. *Proceedings of the National Academy of Sciences*, 103(51):19460–19465, 2006.

[9] C. Grootaert and T. Van Bastelaer. *Understanding and measuring social capital: A multidisciplinary tool for practitioners*, volume 1. World Bank Publications, 2002.

[10] Z. Huang and Q. Shen. Fuzzy interpolation and extrapolation: A practical approach. *IEEE Trans. Fuzzy Syst.*, 16(1):13–28, 2008.

[11] R. Scott Crowder Iii. Predicting the mackey-glass timeseries with cascade-correlation learning. *Connectionist Models*, pages 117–123, 1990.

[12] S. Jin, R. Diao, C. Quek, and Q. Shen. Backward fuzzy rule interpolation. *IEEE Trans. Fuzzy Syst.*, 22(6):1682–1698, 2014.

[13] Humar Kahramanli and Novruz Allahverdi. Design of a hybrid system for the diabetes and heart diseases. *Expert Systems with Applications*, 35(1):82–89, 2008.

[14] L. T. Kóczy, K. Hirota, and L. Muresan. Interpolation in hierarchical fuzzy rule bases. In *Proceedings of International Conference on Fuzzy Systems*, pages 471–477, 2000.

[15] F. Li, C. Shang, Y. Li, J. Yang, and Q. Shen. Fuzzy rule-based interpolative reasoning supported by attribute ranking. *IEEE Trans. Fuzzy Syst.*, 26(5):2758–2773, 2018.

[16] F. Li, C. Shang, Y. Li, J. Yang, and Q. Shen. Interpolation with just two nearest neighbouring weighted fuzzy rules. *IEEE Trans. Fuzzy Syst.*, 2019.

[17] S.S. Morse. Factors in the emergence of infectious diseases. *Emerging infectious diseases*, 1(1):7, 1995.

[18] N. Naik, R. Diao, and Q. Shen. Dynamic fuzzy rule interpolation and its application to intrusion detection. *IEEE Transactions on Fuzzy Systems*, 26(4):1878–1892, 2018.

[19] J.A. Patz, D. Campbell-Lendrum, T. Holloway, and J.A. Foley. Impact of regional climate change on human health. *Nature*, 438(7066):310–317, 2005.

[20] G.V.S. Raju and J. Zhou. Adaptive hierarchical fuzzy controller. *IEEE Trans. Syst., Man, Cybern.*, 23(4):973–980, 1993.

[21] G.V.S. Raju, J. Zhou, and A.K. ROGER. Hierarchical fuzzy control. *International journal of control*, 54(5):1201–1216, 1991.

[22] A. Roychowdhury, D.K. Pratihar, N. Bose, K.P. Sankaranarayanan, and N. Sudhahar. Diagnosis of the diseases: using a ga-fuzzy approach. *Information Sciences*, 162(2):105–120, 2004.

[23] J.A. Sanz, M. Galar, A. Jurio, A. Brugos, M. Pagola, and H. Bustince. Medical diagnosis of cardiovascular diseases using an interval-valued fuzzy rule-based classification system. *Applied Soft Computing Journal*, 20(20):103–111, 2014.

[24] T. Uar, A. Karahoca, and D. Karahoca. Tuberculosis disease diagnosis by using adaptive neuro fuzzy inference system and rough sets. *Neural Computing & Applications*, 23(2):471–483, 2013.

[25] W.E. Walker, P. Harremos, J. Rotmans, J.P. van der Sluijs, M.B.A. van Asselt, P. Janssen, and M.P. Krayer von Krauss. Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment*, 4(1):5–17, 2003.

[26] D. Wang, X.J. Zeng, and J.A. Keane. Intermediate variable normalization for gradient descent learning for hierarchical fuzzy system. *IEEE Trans. Fuzzy Syst.*, 17(2):468–476, 2009.

[27] L.-X. Wang and J.M. Mendel. Generating fuzzy rules by learning from examples. *IEEE Transactions on Systems Man & Cybernetics*, 22(6):1414–1427, 2002.

[28] L.X. Wang. Analysis and design of hierarchical fuzzy systems. *IEEE Trans. Fuzzy Syst.*, 7(5):617–624, 1999.

[29] L. Yang, F. Chao, and Q. Shen. Generalized adaptive fuzzy rule interpolation. *IEEE Trans. Fuzzy Syst.*, 25(4):839–853, 2017.

[30] L. Yang and Q. Shen. Adaptive fuzzy interpolation. *IEEE Trans. Fuzzy Syst.*, 19(6):1107–1126, 2011.

[31] L. A. Zadeh. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans. Syst., Man, Cybern.*, 3:28–44, 1973.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: