This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2892348, IEEE Transactions on Fuzzy Systems

Generic Disjunctive Belief Rule Base Modeling, Inferencing, and Optimization

Lei-lei Chang, Zhi-jie Zhou, Huchang Liao, *Senior Member*, *IEEE*, Yu-wang Chen, Xu Tan, and Francisco Herrera, *Senior member*, *IEEE*

Abstract**—**The combinatorial explosion problem is a great challenge for Belief Rule Base (BRB) when a complex system has over-numbered attributes and/or referenced values for the attributes. This is because BRB is conventionally constructed under the conjunctive assumption, conjunctive BRB, which requires transversally covering each possible combination of all referenced values for all attributes. To solve this challenge, this study proposes a generic modeling, inferencing, and optimization approach for BRB under the disjunctive assumption, disjunctive BRB, that can significantly reduce its size. First, a disjunctive BRB is defined based on the mathematical description of the BRB space. The minimum size requirement for a disjunctive BRB is also discussed in comparison to a conjunctive one. Building on this, the generic disjunctive BRB modeling and inferencing procedures are proposed. Furthermore, an improved optimization model with further relaxed restrictions is constructed, and an optimization algorithm is developed in which only the new rule is optimized and its referenced values range is determined by the optimal solution in the former round optimization. With fewer variables and a more concise solution space, the new optimization algorithm is more efficient. Three cases are studied to validate the efficiency of the proposed disjunctive BRB approach. The study confirms that by integrating both experts' knowledge and historic data, the modeling and inferencing processes can be well understood. Moreover, optimization can further improve the modeling accuracy while it facilitates downsizing BRB in comparison with previous studies and other approaches.

Keywords- Belief rule base, disjunctive assumption, modeling, inferencing, optimization

Manuscript received XXXX; revised XXXX, accepted XXXX. Date of publication XXXX. Date of current version XXXX.

This work was supported by the National Science Foundation of China under Grants 71601180, 71501135, 71771156, 61773388, 61751304, 61702142, and U1709215, the Post-Doctorial Science Foundation under Grants 2015M570847 and 2016T90938, and the Natural Science Foundation of Hainan Province under Grant 617120, the Ministry of Education in China Liberal Arts and Social Sciences Foundation (Nos. 17YJCZH157) and the Pengcheng Scholar Funded Scheme.

Lei-lei Chang is with School of Automation, Hangzhou Dianzi University Hangzhou 310018, China; He is also with the Department of Management, High-Tech Institute of Xi'an, Xi'an 710025, China (Email: leileichang@hotmail.com).

Zhi-jie Zhou is with the Department of Control Engineering, High-Tech Institute of Xi'an, Xi'an 710025, China (Email: zhouzj04@tsinghua.edu.cn).

Huchang Liao is with the Business School, Sichuan University, Chengdu 610064, China (e-mails: liaohuchang@163.com).

Yu-wang Chen is with the Manchester Business School, the University of Manchester, Manchester MP15 6PB UK (Email: yu-wang.chen@mbs.uk.ac).

Xu Tan is with the Shenzhen Institute of Information Technology, Shenzhen 518172, China (Email: tanxu_nudt@yahoo.com).

Francisco Herrera is with the Department of Computer Science and Artificial Intelligence at the University of Granada, Granada 18071, Spain (Email: herrera@decsai.ugr.es).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier.

I. INTRODUCTION

A Belief Rule Base (BRB) expert system has the advantages of representing and integrating multiple types of information under uncertainty, e.g., quantitative historic data and qualitative expert knowledge [9] [31] [32]. As a rule-based expert system, it is close to human knowledge presentation [12] [13] [30]. Moreover, its modeling and inferencing processes are accessible to experts and decision makers as a white box and the inferencing result can be easily understood [7] [30]–[32]. Since it was proposed, BRB has been successfully applied in solving complex system modeling problems in many research and practical fields [25]–[26] [30] [32] [33]–[38].

However, BRB must address the combinatorial explosion challenge when it is constructed under the conjunctive assumption, conjunctive BRB, which requires covering each possible combination of all referenced values for all attributes. Hence, the size of a conventional conjunctive BRB can grow exponentially with over-numbered (referenced values for) attributes; e.g., for a problem with 5/6/7 attributes where each attribute has three referenced values, the size of a conjunctive BRB would be $3^5/3^6/3^7$, respectively. For a BRB whose size is this large, it would be impossible to directly construct a complete BRB by gathering information from sensors and/or experts.

To address this challenge, many endeavors have been undertaken. These can be categorized into five types:

1) Structure learning for BRB, which selects the most representative attributes. Yang *et al* proposed using the Principle Component Analysis (PCA) to transform multiple attributes into reduced-numbered "principle components" [32]. Chang *et al.* first proposed BRB structure learning using multiple dimensionality reduction techniques [4]. Wang *et al.* further explored BRB structure learning using techniques from the rough set theory [25].

2) Parameter learning for BRB, which selects the most representative referenced values for attributes [15]. Note that this is not the same as the previous parameter learning for conventional BRB, whose goal was to improve the modeling accuracy [31] [35]–[38]. Only by reducing the number of the referenced values for the attributes can BRB be downsized (if the attributes are not reduced by structure learning).

Note that a joint approach has been proposed [6] [22] by integrating BRB structure and parameter learning with the Akaike Information Criterion (AIC) as the objective [1]. However, it does not contribute as a new type of research because it continues to be under the conjunctive assumption (it is still designed for the conjunctive BRB).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2892348, IEEE Transactions on Fuzzy Systems

3) Construct multi-level BRB. Upon a further understanding on the inner mechanism of the complex systems, a multi-level model can be constructed and thus a BRB with multiple sub-BRBs can be constructed [29] [31]. The requirement of this approach is a clear and thorough knowledge of the complex system.

4) Replace the complete BRB with an incomplete version for online assessment. Zhou *et al.* [38] proposed "statistical utility" to rate a rule to determine whether it should be maintained in the BRB; this can result in an incomplete BRB. In [37], only five rules, instead of the complete 56 rules, are required. An incomplete BRB can be acceptable for an online assessment because the input(s) in online assessment problems is/are within a time interval, which activates a limited number of rules. Therefore, a partial of the complete BRB is sufficient for an online assessment problem.

5) Construct BRB under the disjunctive assumption, disjunctive BRB. Chang *et al.* firstly applied disjunctive BRB in classification problems by proposing new (yet not generic) rule-activation and weight-calculation procedures [7]. Yang *et al.* further applied this in bridge risk assessment based on an extended disjunctive BRB [33]. Chang *et al.* also extended the joint optimization approach [6] to disjunctive BRB [5].

To summarize, the above five endeavors can facilitate downsizing BRBs to different degrees. Comparatively, the first four are conducted within the scope of conjunctive BRB. Thus, they cannot fundamentally solve the combinatorial explosion problem. Only by constructing a disjunctive BRB can this challenge be effectively addressed because the referenced values for the attributes are not transversally constructed for a disjunctive BRB.

However, in limited studies on disjunctive BRB, many questions remain unanswered. What is the mathematical basis for a disjunctive BRB? What is its minimal size requirement? How should a disjunctive BRB be constructed? What is its inferencing process? How should it be optimized?

The main work and contribution of this study focuses on addressing two aspects: providing mathematical basis for the disjunctive BRB (addressed in Section III) and proposing the generic modeling, inferencing, and optimization approach for the disjunctive BRB (addressed in Section IV).

Compared with many previous studies on BRB [9]-[11] [31] [35]-[38], this study is the first attempt to provide a mathematical basis for disjunctive BRB. Building on the concept of the BRB space, the disjunctive BRB modeling and inferencing procedures are proposed in a generic fashion. A generic optimization model is also constructed with an additional differentiation ratio v included as the decisive variable, and further relaxed restraints requirements. For the optimization algorithm, the computational efficiency is further improved. In each round of optimization, only the new rule is optimized instead of the complete rules. Moreover, the bounds of the new rule are determined by the optimization results from the last round which produces the smallest error.

The remainder of this study is organized as follows. BRB basics and its challenge are introduced in Section II. Section III defines the BRB space and disjunctive BRB. Section IV proposes the generic disjunctive BRB modeling, inferencing, and optimization approach. Three cases are studied for verification in Section V. This study is concluded in Section VI.

II. Background and Challenge of Conventional BRB construction

A. BRB system

A BRB system is an expert system that can address different types of information under uncertainty. BRB is comprised of multiple belief rules in the same belief structure [31] [32]. The *k*th rule in the BRB system is described as:

$$\begin{aligned} R_k: \text{if} \quad & \left(x_1 \text{ is } A_1^k\right) \wedge \left(x_2 \text{ is } A_2^k\right) \wedge \cdots \wedge \left(x_M \text{ is } A_M^k\right), \\ & \text{then } \{(D_1, \beta_{1,k}), \cdots, (D_N, \beta_{N,k})\} \\ & \text{with rule weight } \theta_k, \text{attribute weight } \delta_m \end{aligned} \tag{1}$$

where $x_m (m=1,\cdots M)$ denotes the *m*th attribute, $A_m^k (m=1,\cdots M; k=1,\cdots K)$ denotes the referenced values of the *m*th attribute in the *k*th rule, *M* denotes the number of attributes, $\beta_{n,k} (n=1,\cdots N)$ denotes the belief for the *n*th scale, D_n, and *N* denotes the number of the scales. "$\wedge$" denotes that the rule in (1) follows the conjunctive assumption.

Definition 1: Conjunctive assumption

A conjunctive assumption denotes that the conclusion of a rule stands when all attributes stand. In other words, a conjunctive rule (a rule which follows the conjunctive assumption) would be activated only if all of its attribute are activated. Correspondingly, a conjunctive BRB is comprised of multiple conjunctive rules.

As a rule-based expert system, it is very close to human knowledge presentation [9] [10] [30]. Moreover, its inferencing process is a white box [5] [9] [30] [31]. The nonlinearity modeling ability of BRB has been validated by successfully applied in solving multiple cases from different research and practical fields [30] [32] [33]–[38].

B. Combinatorial Explosion Challenge

However, BRB has to address the combinatorial explosion problem when there are over-numbered attributes and/or referenced values for the attributes. In the pipeline leak detection case [35]–[38] with only two attributes, 56 rules are required because the two attributes have seven and eight referenced values, respectively. In the capability assessment problem with five attributes [4], a maximum 243 (= 3^5) rules are required by assuming that each attribute has three referenced values.

This is because the conventional conjunctive BRB assumes a rule stands when all of its attributes are conjunctively activated. By doing so, at least one rule would be activated for whatever input. Therefore, it requires covering each possible combination of the referenced values for the attributes.

Definition 2: Size of BRB

The size of a BRB denotes the number of rules in the BRB. Thus, the size of a conjunctive BRB is

$$size_{BRB,con} = \prod_{m=1}^{M} p_m \tag{2}$$

where *M* denotes the number of the attributes and p_m denotes the number of the referenced values for the *m*th attribute.

Eq. (2) indicates that for problems with over-numbered attributes and/or referenced values for the attributes, the size of a BRB would grow exponentially. Therefore, traditional conjunctive BRB must address the combinatorial explosion problem [4] [10].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2892348, IEEE Transactions on Fuzzy Systems

Remark 1. The size of a conjunctive BRB is fixed once its attributes and referenced values for the attributes are fixed. This is because, for a conjunctive BRB, it is the transversal combination of the referenced values for the attributes and there is no new conjunctive rule to add. In other words, the size of a conjunctive BRB has reached its maximum and can not be added with new rules.

III. DISJUNCTIVE ASSUMPTION IN BRB SPACE

This section is to mathematically define the BRB space as an attribute product measure space using measure theory [2] [17] and prove that BRB under the disjunctive assumption (disjunctive BRB) can help solve the combinatorial explosion problem. It is the mathematical basis for further proposing the generic approach for the disjunctive BRB.

A. BRB under the disjunctive assumption

A disjunctive BRB is comprised of multiple disjunctive rules. The kth disjunctive rule is given as in (3) [31] [7].

$$\begin{aligned} R_k:&\text{ if } \left(x_1 \text{ is } A_1^k\right)\vee\left(x_2 \text{ is } A_2^k\right)\vee\cdots\vee\left(x_M \text{ is } A_M^k\right),\\ &\text{then } \{(D_1,\beta_{1,k}),\cdots,(D_N,\beta_{N,k})\}\\ &\text{with rule weight } \theta_k, \text{ attribute weight } \delta_m \end{aligned} \tag{3}$$

where "$\vee$" denotes that the rule in (3) follows the disjunctive assumption.

Definition 3: Disjunctive assumption

A disjunctive assumption denotes that the conclusion of a rule stands when either one of its attributes stands. In other words, a disjunctive rule (a rule which follows the disjunctive assumption) would be activated if at least one attribute is activated. Correspondingly, a disjunctive BRB is comprised of multiple disjunctive rules.

B. BRB space

Definition 4: Attribute measure space

Let μ_x be the attribute measures in $(\mathcal{A},\mathcal{B})$, where $\mu_x(A)$ is used to measure the degree with which element x of the object space has attributes A. $(\mathcal{A},\mathcal{B},\mu_x)$ is an attribute measure space if the following restraints are satisfied:

(1) Nonnegativity: $\mu_x(A)\ge 0,\ \ \forall A\in\mathcal{B}$;

(2) Regularity: $\mu_x(\mathcal{A})=1$;

(3) Countable additivity: $\mu_x(\bigcup_{i=m}^{\infty} A_m)=\sum_{m=1}^{\infty}\mu_x(A_m),\ \ \text{if}\ \ A_m\in \mathrm{B}$.

Definition 5: Attribute product measure space

Let X be the object space and $(\mathrm{A}_m,\mathrm{B}_m,\mu_x^{(m)}),\, m=1,2,\cdots,M$ be an attribute measure space on A_m. Then, $\left(\prod_{m=1}^M \mathrm{A}_m,\prod_{m=1}^M \mathrm{B}_m,\prod_{m=1}^M \mu_x^{(m)}\right)$ is the attribute product measure space, where

(1) the multiple attribute space $\prod_{m=1}^M \mathrm{A}_m=\left\{\left(a_1,a_2,\cdots,a_M\right)\middle| a_m\in\mathrm{A}_m,\, m=1,2,\cdots,M\right\}$;

(2) the product σ algebra $\prod_{m=1}^M \mathrm{B}_m=\sigma\left\{b_1\times b_2\times\cdots\times b_M \middle| b_m\in\mathrm{B}_m,\, m=1,2,\cdots,M\right\}$, where "$\times$" indicates that the intersected points in the BRB space are generated by the transversal combination of the referenced values of the attributes;

(3) the measure function $\prod_{m=1}^M \mu_x^{(i)}=\{(\mu_x^{(1)}(b_1),\mu_x^{(2)}(b_2),\cdots,\mu_x^{(M)}(b_M))\mid\forall b_m\in\mathrm{B}_m,\ m=1,2,\cdots,M\}$. The measure function is used to measure the value range of the referenced values of the attributes in BRB.

Furthermore, the measure function $\prod_{m=1}^M \mu_x^{(m)}$ must satisfy three properties:

(1) Nonnegative: $\mu_x(A)\ge\vec{0},\ \forall A\in\prod_{m=1}^M \mathrm{B}_m$;

(2) Regularity: $\mu_x\left(\prod_{m=1}^M \mathrm{A}_m\right)=\vec{1}$;

(3) Countable additivity: $\mu_x\left(\bigcup_{m=1}^{\infty}A_m\right)=\sum_{m=1}^{\infty}\mu_x(A_m)$, where $A_j=(a_1^{(j)},a_2^{(j)},\cdots,a_M^{(j)})\in\prod_{m=1}^M \mathrm{B}_m$, and $a_m^{(j)}\cap a_m^{(k)}\ne\varnothing,\ m=1,2,\cdots,M$.

Suppose that $\left(\prod_{m=1}^M \mathrm{A}_m,\prod_{m=1}^M \mathrm{B}_m,\prod_{m=1}^M \mu_x^{(m)}\right)$ is an attribute product measure space and $[0,1]^N$ is a vector space. Then, the following mapping relation is given:

$$R_k:\ \prod_{m=1}^M \mathrm{B}_m\to[0,1]^N.$$

The mapping relation can be regarded as the "THEN" part in a BRB model. More specifically, the beliefs for N scales can be assigned based on the mapping relation when the referenced values of the attributes are known.

Remark 2. For a BRB with M attributes and c is its maximal attributes set, then the product σ algebra $\prod_{m=1}^M \mathrm{B}_m$ is constructed by subsets of $\prod_{m=1}^M \mathrm{A}_m$, and the measure function $\mu_x^{(m)}$ is A_i^k under Rule R_k, which correlates the mth attribute to its referenced values.

C. Base of BRB space

Based on the definition of the BRB space, its base is defined as follows.

Definition 6: Base of the BRB space

Let X be the object space and $\left(\prod_{m=1}^M \mathrm{A}_m,\prod_{m=1}^M \mathrm{B}_m,\prod_{m=1}^M \mu_x^{(m)}\right)$ be the attribute product measure space, The number of referenced values of the mth attribute set A_m is $p(m)$ for $m=1,2,\cdots,M$. Note that $K=\max_{m\in\{1,2,\cdots,M\}} p(m)$, which can be called the dimension of the BRB space. Further, $\{x_1,x_2,\cdots,x_K\}$ is a selected objects set from the object space and satisfies the following conditions.

The corresponding measure of object x_k is $\mu_{x_k}(A)=\left(A_1^k,A_2^k,\cdots,A_M^k\right)$ for $k=1,2,\cdots,K$. For attribute set A_m, the rank of $\{A_m^k\}_{k=1}^N$ is $p(m)$ for $m=1,2,\cdots,M$. Then, $\left\{\left(A_1^k,A_2^k,\cdots,A_M^k\right)\right\}_{k=1}^K$ can be regarded as a base of the BRB space.

Lemma 1: The base of the BRB space guarantees the completeness of the BRB, which demands that there is/are always rule/rules being activated for any input.

Proof: For an input $I^*(I_1^*, I_2^*, \cdots, I_m^*, \cdots, I_M^*)$, there is $\min(A_m) \le I_m^* \le \max(A_m)$ for the mth attribute. By Definition 3, the base of the BRB space contains all the referenced values for the attributes. Rules k and k+1 are activated if $A_m^k < I_m^* < A_m^{k+1}$ and Rule k is activated if $A_m^k = I_m^*$ concerning the mth attribute. Finally, there are rules activated with the input concerning each attribute of the BRB if the base is included.

D. Size of conjunctive and disjunctive BRBs

Comparatively, for a disjunctive BRB, its size is not fixed. The minimal requirement for a disjunctive BRB is that its base must be included. Therefore, the minimal size of a disjunctive BRB is that of its base; the base's size is the maximal number of referenced values for the attributes, as in Eq. (4),

$$\min(size_{BRB,dis}) = \max_{m=1}^{M}(p(m)) \tag{4}$$

Especially, when $p(1) = p(2) = \cdots = p(M) = P$, there is

$$\min(size_{BRB,dis}) = P \tag{5}$$

With (4) and (5), all of the referenced values for the attributes are included in the disjunctive BRB. That is, the base is included in the disjunctive BRB. Thus, by Lemma 1, the disjunctive BRB is complete and any input can be addressed.

Furthermore, by adding new disjunctive rules different from previous rules, the size of the disjunctive BRB would increase until all possible combinations of the referenced values for the attributes are addressed, Therefore, we have,

$$\max(size_{BRB,dis}) = \prod_{m=1}^{M} p_m \tag{6}$$

By (6), the maximum size of a disjunctive BRB is the same as that of a conjunctive BRB with the same attributes and referenced values for the attributes.

Table I compares the size of a BRB with M attributes and $p(m)$ referenced values for the mth attribute under different assumptions.

TABLE I. SIZE COMPARISON FOR CONJUNCTIVE AND DISJUNCTIVE BRBS

assumption	No. attr.	No. ref. values for mth attr.	size of BRB min	size of BRB max
conjunctive	M	$p(m)$	$\prod_{m=1}^{M} p(m)$	$\prod_{m=1}^{M} p(m)$
disjunctive	M	$p(m)$	$\max_{m=1}^{M}(p(m))$	$\prod_{m=1}^{M} p(m)$

As shown in Table I, the size of a conjunctive BRB is fixed once its number of attributes and referenced values for the attributes are determined.

For a conjunctive BRB with M attributes and $p(m)$ referenced values for the mth attribute, its size is $\prod_{m=1}^{M} p(m)$. Whereas, the size of such a disjunctive BRB with the same number of attributes and referenced values for the attributes is between $\max_{m=1}^{M}(p(m))$ (the minimal size of its base) and $\prod_{m=1}^{M} p(m)$ (the size of a conjunctive BRB in the same belief structure)).

To further explore, Table I also shows that the disjunctive assumption can help downsize a BRB because the number of rules of a disjunctive BRB is only relevant to the number of the referenced values for the attributes and is irrelevant to the number of attributes. This makes a disjunctive BRB more suitable for problems with a large number of attributes compared to a conjunctive BRB.

Example 1: Suppose that there is a base for a disjunctive BRB with M attributes and each attribute has K referenced values. The following gives a possible base:

$$\begin{aligned} &R_1: \left(A_1^1, A_2^1, \cdots, A_M^1\right) \\ &R_2: \left(A_1^2, A_2^2, \cdots, A_M^2\right) \\ &\cdots \\ &R_K: \left(A_1^K, A_2^K, \cdots, A_M^K\right) \end{aligned} \tag{7}$$

For the base in (7), it is the minimal condition and its size is K for there are K rules.

A new disjunctive Rule K+1, given by a new expert or generated by a new sensor, can be added to (7). In Rule K+1, its referenced values for the first attribute x_1 is the same as that in Rule 1 in (7), A_1^1, while the rest is the same as that in Rule K in (7), $(A_2^K, \cdots, A_M^K)$. With this, a new base is then given in (8):

$$\begin{aligned} &R_1: \left(A_1^1, A_2^1, \cdots, A_M^1\right) \\ &R_2: \left(A_1^2, A_2^2, \cdots, A_M^2\right) \\ &\cdots \\ &R_K: \left(A_1^K, A_2^K, \cdots, A_M^K\right) \\ &R_{K+1}: \left(A_1^1, A_2^K, \cdots, A_M^K\right) \end{aligned} \tag{8}$$

For the new base in (8), its size is K+1 for there are K+1 rules. If more information is gathered, (8) can be further expanded until all possible combinations of the referenced values for the attributes are added. At that time, its size would be $K^M = \prod_{m=1}^{M} K$ which is the size of the conjunctive BRB with M attributes and each attribute has K referenced values.

IV. GENERIC MODELING, INFERENCING, AND OPTIMIZATION APPROACH FOR DISJUNCTIVE BRB

In this section, generic the disjunctive BRB modeling is provided in Part A, and the inferencing process is given in Part B. The optimization model and algorithm for disjunctive BRB are given in Parts C and D, respectively.

A. Generic disjunctive BRB modeling

The generic disjunctive BRB modeling procedure is presented as follows.

Step 1: Identify the attributes of the BRB, $x_m, m = 1, \cdots, M.$;

Step 2: Identify the referenced values for the attributes, $A_m^{p(m)}, m = 1, \cdots, M.$;

Step 3: Construct rules comprised of the referenced values for the attributes. Note that any rule must contain at least one attribute and only one referenced value for each attribute. Repeat Step 3 until no additional new rules are required;

Step 3.1: For the kth rule of K rules, identify θ_k, identify A_m^k for x_m where $k \in (1, \cdots, p(m)), m \in (1, \cdots, M)$ and identify $\beta_{n,k}$ for D_n where $n \in (1, \cdots, N)$ and θ_k. Repeat this step until all parameters for the total K rules are identified.

Step 3.2: Select a set of parameters to construct new disjunctive rule(s) as in (3).

Step 4: Verify if any referenced value for any attribute that has not been included in any rule. Let $A_{\text{all}} = \bigcup_{m=1}^{M} A_m^{p(m)}$, $A_{\text{BRB}} = \bigcup_{k=1}^{K}\bigcup_{m=1}^{M} A_m^k$ Go to step 6 if $A_{\text{all}} \cap A_{\text{BRB}} = \varnothing$; otherwise go to Step 5.

Step 5: Identify new parameters and go to Step 3.3.

Step 6: The construction of a disjunctive BRB is complete.

Remark 3. The minimum requirement for a disjunctive BRB is to ensure its completeness. In other words, all of the referenced values for the attributes (or the base of the BRB space) must be included. Afterwards, additional rules can be added based on experts' knowledge. If all possible combinations of the referenced values for the attributes are added, its size would be the same as that of the conjunctive BRB in the same belief structure (see Parts C/D of Section III).

B. Generic disjunctive BRB inferencing

B.1 Generic rule activation for disjunctive BRB

Unlike the rule activation for conjunctive BRB, the activation mechanism for the disjunctive BRB is more complex. A new activation factor κ is introduced. κ has two status with being "1" as the rule is activated and being "0" as not activated.

Thus, κ for a rule concerning the mth attribute is calculated by Eq. (9),

$$\kappa(I_m^*, A_m^j) = \begin{cases} 1 & j = k, k+1 (A_m^k < I_m^* < A_m^{k+1}) \\ 0 & j \neq k, k+1 (A_m^k < I_m^* < A_m^{k+1}) \end{cases} \tag{9}$$

When the input is equivalent with a referenced value in the rule, there is

$$\kappa(I_m^*, A_m^j) = \begin{cases} 1 & j = k (I_m^* = A_m^k) \\ 0 & j \neq k (I_m^* = A_m^k) \end{cases} \tag{10}$$

For a model with M attribute, the kth rule is activated if at least one attribute is activated, as calculated by Eq. (11),

$$\kappa(I^*, A^k) = \begin{cases} 1 & \text{if } \sum_{m=1}^{M} \kappa(I_m^*, A_m^k) \geq 1 \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

Specifically when there are multiple rules, namely Rule k and Rule k+1, are activated concerning the mth attribute, $\kappa(I_m^*, A_m^k) = \kappa(I_m^*, A_m^{k+1}) = 1$, and they share the same referenced values for the mth attribute, $A_m^k = A_m^{k+1}$, however, they have different activation status for another pth attribute, say $\kappa(I_p^*, A_p^k) = 0, \kappa_1(I_p^*, A_p^{k+1}) = 1$.

For the kth rule, the special condition is

$$\begin{cases} \kappa(I_m^*, A_m^k) = 1 \\ A_m^k = A_m^{k+1} \\ \kappa(I_p^*, A_p^k) = 0 \end{cases} \tag{12}$$

For the k+1th rule, the special condition is

$$\begin{cases} \kappa(I_m^*, A_m^{k+1}) = 1 \\ A_m^k = A_m^{k+1} \\ \kappa(I_p^*, A_p^{k+1}) = 1 \end{cases} \tag{13}$$

Then, κ is calculated by Eq. (14)

$$\kappa(I^*, A^j) = \begin{cases} 0 & j = k, (12) \\ 1 & j = k+1, (13) \end{cases} \tag{14}$$

Detailed illustrations on the activation mechanism (as well as matching degree and weight calculation) for the disjunctive BRB can be found in Case I.

B.2 Generic matching degree calculation and weight calculation for disjunctive BRB

The matching degree for the mth attribute in the kth rule to the input is calculated as in Eq. (15),

$$\varphi(I_m^*, A_m^j) = \begin{cases} \dfrac{A_m^{k+1} - I_m^*}{A_m^{k+1} - A_m^k} & j = k (A_m^k \leq I_m^* \leq A_m^{k+1}), \kappa(I^*, A^j) = 1 \\ \dfrac{I_m^* - A_m^k}{A_m^{k+1} - A_m^k} & j = k+1, \kappa(I^*, A^j) = 1 \\ 0 & j = 1, 2, ..., p(m), j \neq k, k+1, \kappa(I^*, A^j) = 0 \end{cases} \tag{15}$$

The integrated matching degree for the mth attribute in the kth rule is calculated as in Eq. (16),

$$\alpha_{m,k} = \frac{\varphi(I_m^*, A_m^j)\varepsilon_m}{\sum \varphi(I_m^*, A_m^j)} \tag{16}$$

where ε_m denotes the confidence of the mth attribute being assessed as I_x^*.

The activated weight for the kth rule is calculated by Eq. (17),

$$w_k = \frac{\theta_k \sum_{m=1}^{M} (\alpha_{m,k})^{\delta_m}}{\sum_{k=1}^{K} (\theta_l \sum_{m=1}^{M} (\alpha_{m,k})^{\delta_m})} \tag{17}$$

where θ_k represents the initial weight of the kth rule and $\theta_k(\delta_m)$ represents the initial weight for the kth rule (mth attribute). Therefore, $\sum_{k=1}^{K} w_k = 1$.

C. Optimization model for disjunctive BRB

The optimization objective for BRB is normally the error between the estimated outputs of BRB and the actual outputs, e. g., the mean square error (MSE). The decisive variables usually include the referenced values for the attributes, the initial rule weights, the initial attribute weights, the beliefs of the scales in the conclusion part, etc.

A new variable is introduced in the optimization model, namely the differentiation ratio v. It is used to produce an "Unknown" output when the difference between the biggest and second beliefs is not big enough.

As defined in (1) and (3), the nth scale D_n is with a belief β_n, and thus a mapping function can be defined in Eq. (18),

$$D(n) = \beta_n \tag{18}$$

Then, we have

$$n = D^{-1}(\beta_n) \tag{19}$$

where $D^{-1}(\beta_n)$ denotes to the reverse function of $D(n)$ in (19) which returns the number of scale n whose belief is β_n.

Correspondingly, the output is determined by the difference between the biggest belief $(\max_1(\beta_n))$ and the second biggest belief $(\max_2(\beta_n))$, as given in Eq. (20)

$$output = \begin{cases} \hat{n} = D^{-1}(\max_1(\beta_n)) & \text{if } \max_1(\beta_n) - \max_2(\beta_n) > v \\ \text{Unknown} & \text{if } \max_1(\beta_n) - \max_2(\beta_n) \leq v \end{cases} \tag{20}$$

The optimization model is given as follows:

$$\min \quad MSE(A_m^k, \theta_k, \delta_m, \beta_{n,k}, v) \tag{21}$$

s. t.

$$lb_m \leq A_m^k \leq ub_m \tag{21a}$$

$$A_m^p = lb_m; A_m^p = lb_m \tag{21b}$$

$$0 < \theta_k \le 1; 0 < \delta_m \le 1 \tag{21c}$$

$$0 \le \beta_{n,k} \le 1; \sum_{n=1}^{N} \beta_{n,k} \le 1 \tag{21d}$$

$$0 \le v < 1 \tag{21e}$$

where $k = 1,...,K; n = 1,...,N; m = 1,...,M; p \ne q \in [1,...,M]$. Restraints represented by Eq. (21a/c/d/e) states that the referenced values of the *m*th attribute, the initial rule weights, the initial attribute weights, the beliefs for the scales, and the differentiation ratio should be within its range. If there is no incomplete information, then $\sum_{n=1}^{N} \beta_{n,k} = 1$. If there is no "Unknown" output, then $v = 0$, which is not compulsorily required. Eq. (21b) denotes that the lower and upper bounds of the referenced values of the *m*th attribute must be included in the rules, but not limited to a specific rule.

Remark 4. Other than the new introduced differentiation ratio v, an improvement is made on restrictions represented by Eq. (21b) in the new optimization model in comparison with previous model. Compared with the previous training and learning models in [9] [30] [35]–[38], the referenced values are included in the optimization model as in (21a/b). Compared with the related studies for disjunctive BRBs [7] [33] [34], (21b) further relaxed the restrictions: no requirement for the first and last rules to be within the lower and upper bounds of the attributes (although it would be rational to derive such rules); they can be in any rule in the disjunctive BRB.

D. Optimization algorithm for disjunctive BRB

To solve the above optimization model, different algorithms are applicable as the optimization engine. In [35]-[38], the Newton approach was applied as a deterministic algorithm. In [7]-[9][33] [34], the Evolutionary Algorithms (such as the genetic algorithm (GA), the Differential Evolutionary (DE) algorithm, etc.) are applied.

Comparatively, it is argued that EAs are more suitable for problems with multiple variables [14] [20]. Among many EAs, DE is applied in this study as the optimization engine for its superior performance in multiple computation contests [20].

The new optimization algorithm is implemented in an iterative fashion. The biggest difference in comparison with previous studies is that the new individual (rule) is generated based on the previous optimal result. Moreover, in the optimization process, only the new added rule is the decisive variable instead of the complete rules, which saves more computational resources and is more efficient.

The detailed steps are given as follows:

Step 1: Initialization.

If it is the first round initialization, then go to Step 1.1; if not, go to Step 1.2;

Step 1.1: First round initialization

In the first round initialization step, there are two types of parameters to initialize. The first is the parameters of BRB, including the referenced values for the attributes, A, the initial rule weights, θ, and the beliefs of scales in the conclusion part, β. Moreover, the initial and final size of BRB (n_{ini} and n_{max}) should also be determined. Normally n_{ini}=3 and n_{max} would be set at a larger number. The second is the parameters of DE, namely the number of individuals, *np*, the number of generations, *ng*, the mutation/crossover ratios, etc. After initialization, go to Step 2.

Step 1.2: Further round initialization

For further round initialization step, the parameters to initialize include the referenced values of the attributes for the new rule, the beliefs for the new rule and the weights for all of the rules. After initialization, go to Step 2.

Step 2: Optimization

With the initialized individuals, go through the optimization operations for both the first and further rounds.

Step 2.1: Crossover

The crossover strategy states that the *j*th gene, $t'_{i,j}$, of a temporary individual, t, is selected by the probability of CR (or the *j*th gene, $c'_{i,j}$, of the current individual, c, is selected by the probability of $1-CR$) as the *j*th gene, $u'_{i,j}$, of the final individual, u, as shown in Eq. (22),

$$u'_{i,j} = \begin{cases} t'_{i,j} & \text{if } (rand \le CR) \text{ } or \text{ } (j = sn) \\ c'_{i,j} & \text{otherwise} \end{cases} \tag{22}$$

where $CR = 0.9$ is the crossover operator and $sn \in [1, 2, \ldots, n]$ is a random integer which is generated with each new individual.

Step 2.2: Mutation

The *i*th individual in the new generation, g'_i, can be obtained using Eq. (23),

$$g'_i = c_{r1} + F * (c_{r2} - c_{r3}) \tag{23}$$

where c_{r1}, c_{r2} and c_{r3} are three random individuals and $r1 \ne r2 \ne r3$, $F = 0.5$ is the mutation operator.

Step 2.3: Fitness calculation

Go to Step 2.3.2 for the first round. Go to Step 2.3.1 if not.

Step 2.3.1 Generate new BRBs (with *n*+1 rules) by combining the new rule with predetermined optimal BRBs (with *n* rules)

Replace the initial rule weights θ_{ini} with the new rule weights θ_{new} for all rules; add the new referenced values A_{new} and beliefs β_{new} of the new rule to the optimal individuals (BRBs) identified by the last round optimization. The optimal individuals are the top 25% identified in the last round.

Step 2.3.2: BRB inferencing. The rule-activation, matching-degree calculation, and weight-calculation procedures are given in Part B Section IV.

Step 2.3.3 ER inference

After the activation of certain rules, the activated L rules are integrated using ER as in Eqs. (24-25) [26],

$$\beta_n = \frac{\mu[\prod_{k=1}^{L}(w_k\beta_{n,k} + 1 - w_k\sum_{n=1}^{N}\beta_{n,k}) - \prod_{k=1}^{L}(1 - w_k\sum_{n=1}^{N}\beta_{n,k})]}{1 - \mu[\prod_{k=1}^{L}(1 - w_k)]} \tag{24}$$

$$\mu = [\sum_{n=1}^{N}\prod_{k=1}^{L}(w_k\beta_{n,k} + 1 - w_k\sum_{n=1}^{N}\beta_{n,k}) - (N-1)\prod_{k=1}^{L}(1 - w_k\sum_{n=1}^{N}\beta_{n,k})]^{-1} \tag{25}$$

where β_n represents the belief for the *n*th scale. Appendix proves $\sum_{n=1}^{N}\prod_{k=1}^{L}(w_k\beta_{n,k} + 1 - w_k\sum_{n=1}^{N}\beta_{n,k}) - (N-1)\prod_{k=1}^{L}(1 - w_k\sum_{n=1}^{N}\beta_{n,k}) \ne 0$.

Step 2.4: Selection

For the *i*th individual u_i^t, it enters the new generation when its fitness function get a higher rated value, as indicated by Eq. (26),

$$c_i^{t+1} = \begin{cases} u_i^t & \text{if } \quad f(u_i^t) \le f(c_i^t) \\ c_i^t & \text{otherwise} \end{cases} \tag{26}$$

where $f(\bullet)$ is the fitness function (MSE in this study).

Step 2.5: Check for optimization stop criterion.

If the preset criterion has been met (usually the number of generations, *ng*), then go to Step 4; if not, go to Step 3.

Step 3: Identify attribute range

This step is to identify the attribute range for generating the new rule in the next generation by locating the section with the biggest error by the optimal result from the previous generation.

Step 3.1: Compute and derive an error list.

With the optimized optimal solution, compute the absolute error for each set of data and store them in an error list, $list$;

$$error = abs(output_{estimated} - output_{actual}) \tag{27}$$

Step 3.2: Rank and select the top 25% from the error list.

Rank the data in the error list and select the top 25% from the error list, $list_{ranked}$. Decision makers can determine how many data they would like to form the error list.

Step 3.3: Map $list_{ranked}$ in the BRB space.

For each set of data in $list_{ranked}$, map them in the BRB space by projecting its values for each attribute in the corresponding dimension in the BRB space.

Step 3.4: Identify the section in the BRB space with the most selected data.

For the ranked error list, $list_{ranked}$, insert and re-rank the referenced values for the *m*th attribute, and get a new ranked list, $list_{new,ranked}$, by Eq. (28).

$$list_{new,ranked} = list_{ranked} \cup A_m \tag{28}$$

where $\cup$ denotes to add the referenced values for the *m*th attribute A_m to the ranked error list $list_{ranked}$.

Identify the order for each of the referenced values for the *m*th attribute by Eq. (29),

$$order_m = \text{find}(list_{new,ranked} == A_m) \tag{29}$$

Calculate the sets of data between neighboring referenced values for the *m*th attribute by Eq. (30),

$$num_{ref,m} = order_{m+1} - order_m \tag{30}$$

For the *m*th attribute, $[\max(num_{ref,m}), \max(num_{ref,m})+1]$ is the range of referenced values with the most sets of data.

Repeat this process for the rest *M*-1 attributes, the intersected section by *M* attributes would be the one with the most sets of data.

Step 3.5: Identify the range for each attribute for new rule generation.

With the identified section in the BRB space, the new rule would be generated in this specific section. The range of the dimensions of the identified section would be the range of the referenced values for the corresponding attribute of the new rule to generate.

After Step 3.5, go to Step 1.2.

Step 4: Check for stop criterion.

Normally, the final size of BRB is preset at a maximum number as the stop criterion. Stop if met; go to Step 1.2 if not.

Remark 5. The above optimization algorithm for this study is designed using the operators of DE. However, it is not the intension of this study to identify DE as the best candidate for BRB optimization. As long as it is a fit for the targeted problem, combination or variants of different optimization algorithms (e.g., the Genetic Algorithm (GA), the Particle Swarm Optimization (PSO), etc) are encouraged so as to produce a superior modeling performance.

E. Mechanism for the disjunctive BRB approach

The disjunctive BRB modeling, inferencing, and optimization approach provides an effective solution of using even fewer disjunctive rules to model many theoretical and practical problems with ambiguous and fuzzy characteristics.

For disjunctive BRB modeling, it is very close to human knowledge presentation and also data in varied forms which guarantees its convenience in deriving corresponding rules from either using experts' knowledge or historic records. This process should be case specific. The minimal requirement for constructing a complete disjunctive BRB is its base. However, driven by the requirements of the targeted problem, more disjunctive rules can be added with more information/data gathered (see Part C of Section III).

For disjunctive BRB inferencing, it is open and accessible for decision makers and experts. Its core principle is consistent with that of the conjunctive BRB [31] with new rule activation (see Part B.1 of Section IV) and new matching degree calculation (see Part B.2 of Section IV) procedures. The disjunctive BRB inferencing procedures specifically embody the assumption followed by different type of BRBs.

For disjunctive BRB optimization, it has direct interfaces with many optimization techniques. The construction of the optimization model is vital for disjunctive BRB optimization, specifically to determine the optimization objective and the parameters restraints. Although Part D of Section IV uses DE as the optimization engine, other EAs and their variants or deterministic techniques are also applicable. Optimization model construction and algorithm selection should be based on the requirements and characteristics of the targeted problem.

V. Case Study

A. Case I: Disjunctive BRB construction

Suppose that an object has two attributes, x_1 and x_2, and each attribute is with 2(3) referenced values, namely 5/8(/10). The minimal size of the disjunctive BRBs would be 3 based on Eq. (4). Figs. 1 and 2 display the disjunctive BRBs with the smallest size in different structures.

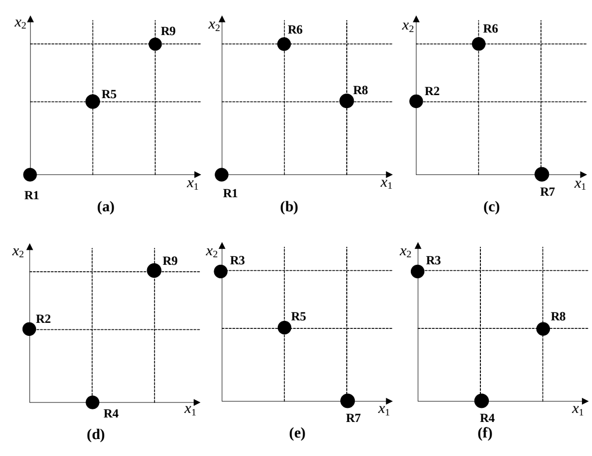

Figure 1. six possible disjunctive BRBs with two attributes (each attribute has three referenced values)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2892348, IEEE Transactions on Fuzzy Systems

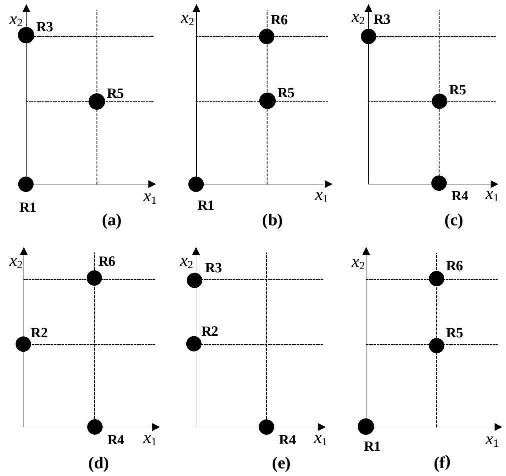

Figure 2. six possible disjunctive BRBs with two attributes (each attribute has 2/3 referenced values)

Note that when the numbers of the referenced values for different attributes are different as in Fig. 2, it is inevitable that certain rules share the same referenced values for the same attributes. In fact, even when the attributes share the same number of referenced values, different rules can be with the same referenced values for the same attribute. This situation is permitted and not against the definition of the disjunctive BRB: a disjunctive BRB is complete as long as all of the referenced values for the attributes are covered (as given in Section III).

However, in practical conditions, the size of a disjunctive BRB would normally be determined first. Then, experts are invited to provide rules. Under this circumstance, it is unlikely that different experts would produce the same referenced values for the same attribute. Thus, the condition in Fig. 1 is more likely to occur in practical conditions.

B. *Case I: Disjunctive BRB inferencing*

Continuing with this case, a subpart of Fig. 1(a) with Rules 5/6/8 is used to illustrate the disjunctive BRB inferencing process. To better demonstrate the procedures, more complex conditions are considered as in Fig. 3(a-f). The rules and matching degrees are shown in Tables II and III.

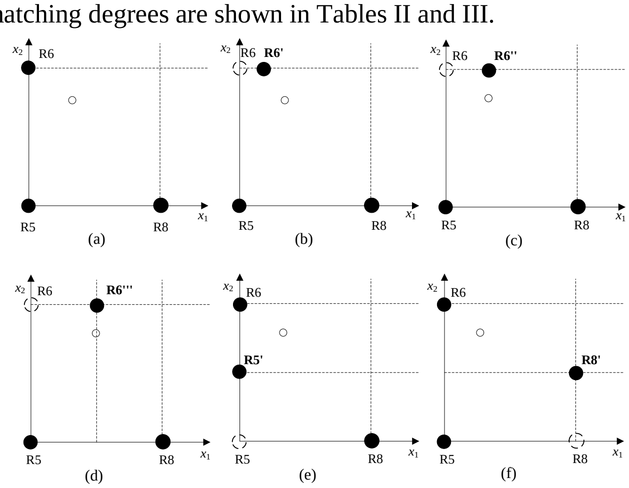

Figure 3. Comprehensive rule activation for disjunctive BRB (solid circles denote rules; hollow circle denotes input)

Compared with Fig. 3(a), Rule 6 is replaced by Rule 6’ in Fig. 3(b). In Fig. 3(c), Rule 6’’ and the input share one referenced value, yet the matching degrees of the input with Rules 5 and 8 are not equal. In Fig. 3(c), Rule 6’’’ and the input share one referenced value and the matching degrees of the input with Rules 5 and 8 are equal. In Fig. 3(e/f), Rules 5 and 8 are replaced by Rules 5’ and 8’, respectively.

Correspondingly, the input is (8.5, 9.5) in Fig. 3(a/b/c/e/f) and (9, 9.5) in Fig. 3(d), respectively.

The corresponding rules provided by the experts based on historic data are in Table II. Table III presents the matching degrees for attributes x_1 and x_2 under different conditions.

TABLE II. RULES IN FIG. 3 FOR CASE I

No.	IF		THEN	
	x_1	x_2	Good	Average
5	8	8	0.4	0.6
6	8	10	0.75	0.25
8	10	8	0.6	0.4
6’	8.1	10	0.75	0.25
6’’	8.5	10	0.65	0.35
6’’’	9	10	0.7	0.3
5’	8	9	0.55	0.45
8’	10	9	0.7	0.3

TABLE III. MATCHING DEGREES FOR RULES IN CASE I

	R5/5’		R6/6’/ 6’’ / 6’’’		R8/8’	
	x_1	x_2	x_1	x_2	x_1	x_2
Fig. 3(a)	0.75	0.25	0.75	0.75	0.25	0.25
Fig. 3(b)	0	0	0.79	0.75	0.21	0.25
Fig. 3(c)	0	0.25	1	0.75	0	0.25
Fig. 3(d)	0	0.25	1	0.75	0	0.25
Fig. 3(e)	0.75	0.5	0.75	0.5	0.25	0
Fig. 3(f)	0	0	0.75	0.5	0.25	0.5

Based on in Table III, it can be observed that:

(1) Rule 5 would not be activated in Fig. 3(b);

(2) Rules 5/8 would be activated in Fig. 3(c/d) since they are of the same distance to the input regarding on attribute x_1;

(3) Rule 8 would be activated in Fig. 3(e) since it would be activated regarding on attribute x_1;

(4) Rule 5 would not be activated in Fig. 3(f).

Remark 6. The rules in Table II are used as both disjunctive and conjunctive rules. This is for a fair comparison. Because, if another conjunctive BRB is given, the results comparison between the two BRBs (one disjunctive and the other conjunctive) would be meaningless.

C. *Case I: Comparative results*

Continuing with this case, next we calculate the assessment result and compare two indices, the MSE and mean absolute percentage error (MAPE) using Eqs. (31–32)

$$\text{mse} = \frac{1}{N}\sqrt{\sum_{n=1}^{N}(output_{\text{conj}} - output_{\text{disj},n})^2} \tag{31}$$

$$\text{mape} = \frac{1}{N}\sum_{n=1}^{N}\frac{\text{abs}(output_{\text{conj}} - output_{\text{disj},n})}{output_{\text{conj}}} \tag{32}$$

where N denotes the number of scales in the conclusion part, $output_{\text{conj}}$ and $output_{\text{disj}}$ denote the outputs of the conjunctive and disjunctive BRBs, respectively.

An additional Rule 9 is added to ensure the completeness of the conjunctive BRB.

R_9 : if $(x_1 \text{ is } 10)\vee(x_2 \text{ is } 10)$, then $\{(\text{Good}, 0.8), (\text{Average}, 0.2)\}$

Tables IV–VI display the rule weights, results (in belief structure and by utility), and MSE/MAPE of Case I.

Other than illustrating the generic modeling and inferencing procedures in different conditions, the following conclusions can be drawn:

(1) Disjunctive BRB can help downsize BRB while maintaining consistent results. With reduced number of rules, the biggest beliefs and the results by utility (in Table V) are rather close with the results by a complete conjunctive BRB.

(2) If the number of rules (NOR) is considered as input requirements, then an average of 84.14% accuracy can be achieved with only 62.5% of original input information in average (in Table VI). If the disjunctive rules are well selected, then 90.34% accuracy can be achieved with only 50% of original input information in average (Fig. 3(f) in Table VI). This is of practical importance since information gathering itself can be very expensive and consuming.

TABLE IV. RULE WEIGHTS FOR CASE I IN MULTIPLE CONDITIONS

Rule	Conj.	Disj.	Fig. 3(a)	Fig. 3(b)	Fig. 3(c)	Fig. 3(d)	Fig. 3(e)	Fig. 3(f)
5/5’	0.1875	0.2500	0.3333		0.1111	0.1111	0.4545	
6/6’/6’’/6’’’	0.5625	0.3750	0.5000	0.7697	0.7778	0.7778	0.4545	0.6250
8/8’	0.0625	0.1250	0.1667	0.2303	0.1111	0.1111	0.0909	0.3750
9	0.1875	0.2500						

Note: “Conj.” denotes the conjunctive BRB with Rules 5/6/8/9; “Disj.” denotes the disjunctive BRB with Rules 5/6/8/9; Rule 6 are Rules 6’/6’’/6’’’ in “Fig. 3(b/c/d)”, respectively; Rules 5/8 are Rules 5’/8’ in “Fig. 3(e/f)”, respectively.

TABLE V. RESULTS FOR CASE I IN MULTIPLE CONDITIONS

	Conj.	Disj. (4 rules)	Fig. 3(a)	Fig. 3(b)	Fig. 3(c)	Fig. 3(d)	Fig. 3(e)	Fig. 3(f)
belief structure	{(S, 72.74%), (A, 27.26%)}	{(S, 69.38%), (A, 30.62%)}	{(S, 64.17%), (A, 35.83%)}	{(S, 74.82%), (A, 25.18%)}	{(S, 64.08%), (A, 35.92%)}	{(S, 68.78%), (A, 31.22%)}	{(S, 67.50%), (A, 32.50%)}	{(S, 76.57%), (A, 23.43%)}
by utility	0.8637	0.8469	0.8209	0.8741	0.8204	0.8439	0.8375	0.8829

Note: “Belief structure” denotes that the results are provided in belief structure; “By utility” denotes that the results are integrated by considering the utility of the scales. The scales “S” and “A” are assigned with the utilities of “1” and “0.5”, respectively.

TABLE VI. MSE AND MAPE FOR CASE I IN MULTIPLE CONDITIONS

	Disj. (complete)	Fig. 3(a)	Fig. 3(b)	Fig. 3(c)	Fig. 3(d)	Fig. 3(e)	Fig. 3(f)	average (Fig. 3(a-f))
MSE	2.3759E-02	6.0599E-02	1.4708E-02	6.1235E-02	2.8001E-02	3.7052E-02	2.7082E-02	2.3759E-02
MAPE	4.6192E-02	1.1782E-01	2.8595E-02	1.1905E-01	5.4440E-02	7.2037E-02	5.2653E-02	4.6192E-02
NOR	4	3	2	3	3	3	2	2.5

D. Case II: Background and basic settings

The pipeline leak detection case which has become a benchmark case for verification of the BRB approach in multiple literatures [5] [29] [34] [37] [38] is studied in this study.

In the pipeline leakage detection case, a total of 2007 sets of data are gathered from 07:00:08 to 12:34:27 in which each set of data is with two attributes, namely *PressureDiff* (ranging from -9.5725 to 1.1) and *FlowDiff* (ranging from -0.015 to 0.039588).

In many of the previous studies [29] [37] [38], the two attributes were assumed to be with 7 and 8 referenced values, respectively. A conjunctive BRB with 56 rules were initialized and then optimized. In other literatures [5] [34], the disjunctive assumption has also been applied which has significantly downsized BRB to be with 3-8 rules while maintaining a relatively high modeling accuracy.

In this study, the disjunctive BRB is applied to model the pipeline leakage detection case.

For consistency with previous studies, *PressureDiff* and *FlowDiff* are assumed be within the bounds of [-10, 2] and [-0.04, 0.04], respectively. The utilities of the output, the *leakage size*, are assumed as follows:

$$U(D_1, D_2, D_3, D_4, D_5)=(0, 2, 4, 6, 8) \tag{33}$$

The use of multiple scales (with utilities) is to make better use of the nonlinearity modeling ability of BRB. As a good nonlinearity simulator, BRB can model any nonlinearity with the introduction of multiple scales (instead of just one scale) [10] [11].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2892348, IEEE Transactions on Fuzzy Systems

As for the optimization model, $v = 0$ since there is no need for differentiation among different scales. As for the optimization algorithms, DE is applied as the optimization engine. its parameters are set as follows: *ng*=1000, *np*=30, *cr*=0.9, *F*=0.5. The initial number of rules is set as 3 and the maximum is set at 12. A total of 30 runs are conducted.

E. *Case II: Results and discussion*

The results comparison between this study and previous studies are shown in Table VII. Table VIII shows the BRB with 5 rules. Fig. 4 compares the outputs and MSEs between this study and [5].

TABLE VII. RESULTS COMPARISON FOR TESTING DATASET IN CASE II

No.	year	description	assumption	MSE (MAE)	Size(training/testing)	NOR
1	2007[29]	local training	conjunctive	0.4049	500/2008	56
2	2009[38]	online updating	conjunctive	0.7880	800/2008	56
3	2010*[37]	sequential learning	conjunctive	0.0241	305/17	5
4	2011[11]	adaptive learning	conjunctive	0.3990	500/2008	56
5	2013**[18]	extended BRB	conjunctive	(0.2169)	/2008	900
6	2015[8]	ER rule	disjunctive	0.3709	/2008	15
7	2016[25]	dynamic rule adjustment	conjunctive	0.5040 0.4450	900/2008 900/2008	14 6
8	2017[5]	bi-level	**disjunctive**	**0.2917**	**500/2008**	**5**
9	2018**[34]	JOPS	conjunctive **disjunctive**	0.3998 **0.3974**	900/2008 **900/2008**	6 **3**
10	2018[6]	AIC	conjunctive	0.3411	500/2008	8
11	This study	generic disjunctive BRB optimization	**disjunctive**	**0.3741** **0.2848** **0.2679**	**500/2008** **500/2008** **500/2008**	**3** **5** **12**

Note: * [41] produces a very small MSE since it concentrated on a very small portion of the total 2007 sets of data. ** [24] [38] used the mean absolute error (MAE) as the optimization objective instead of the MSE, which may have resulted in a lower MAE.

TABLE VIII. BRB WITH 5 RULES FOR CASE II BY DE

No.	weight	attributes		leakage size (utility)					leakage size
		PressureDiff	*FlowDiff*	0	2	4	6	8	(integrated)
1	1.0000	-10.0000	-0.0200	0.9998	0.0000	0.0002	0.0000	0.0000	0.0007
2	0.0456	-1.2168	-0.0196	0.5114	0.3467	0.0848	0.0466	0.0106	1.3967
3	0.0089	-7.0499	-0.0015	0.0009	0.1002	0.0004	0.3144	0.5841	6.7612
4	0.0046	1.7975	0.0399	0.0127	0.6255	0.0022	0.0326	0.3269	4.0709
5	0.6852	2.0000	0.0400	0.9999	0.0000	0.0000	0.0001	0.0000	0.0005

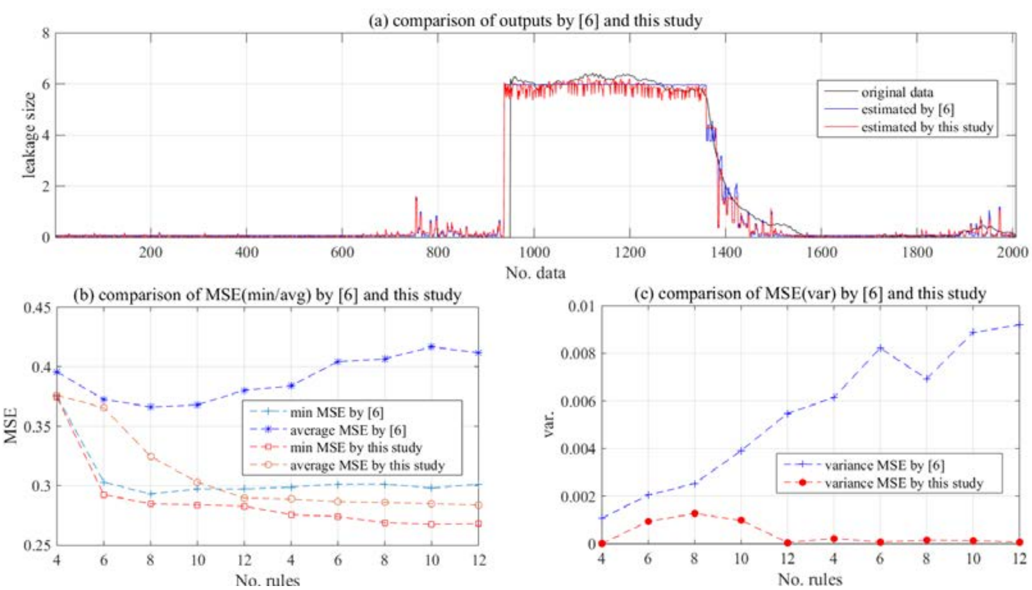

Figure 4. Comparisons on outputs and MSEs for Case II

Several discussions can be drawn:

(1) The size of BRB has been reduced. In comparison with [29] [37] [38], a minimum of 3 rules have produced as good results as 56 rules. With only 5 rules, the modeling accuracy could be within 0.3.

(2) As indicated by the results in Table VII, especially in comparison with [34] and [5] which also used the disjunctive BRB, the results produced by this study is also superior. Compared with [34], the BRB with 3 rules produced an MSE which is 5.87% smaller. Compared with [5], the BRB with 5 rules (as given in Table VIII) produced an MSE which is 2.56% smaller. Furthermore, an even smaller MSE of 0.2672 has been achieved by a BRB with 12 rules.

(3) The new optimization algorithm is also more efficient. This is because the increase in the number of the decisive variables is far fewer: only the new rule is optimized instead of the complete BRB. As indicated by Fig. 3(b), the results by [5] (increase of min/average MSE along with more rules) have shown signs of over-fitting while it is not observed in the results by this study.

(4) Furthermore, the new optimization algorithm is also more robust. Upon further exploration, it shows a very small variance of MSE of 30 runs has been derived by comparing [5] and this study. This is also partially due to the reduced number of decisive variables (computational requirement is reduced). With the computational resources being focusing on the additional new rule, a more consistent result can be derived.

Remark 7. Although only DE is applied as the optimization engine in this study, other EAs can also be effective candidates. In [5] [6] [35] [36], researchers have attempted to use different EAs, e. g., the genetic algorithm (GA), the particle swam optimization (PSO), etc. After all, this study does not intent to identify any optimization algorithm or means as the optimal solution to all problems. In fact, it is believed that the selection of an optimization engine should be case specific and the characteristics and requirements a practical case must be taken under consideration so that a feasible and even optimal solution can be obtained.

F. Case III: Background

Threat level assessment requires gathering data from multiple sensors and then integrating them into one unified result for defense analysis [21].

At time *t*, multiple suspected targets are detected with unknown intention towards a strategic facility with high value.

To assess the threat level of a suspected target, multiple types of information gathered from different sources (e.g., sensors, human) must be considered. For these suspected targets, the following seven factors are listed as follows.

Distance (km): The distance between the suspected targets and the strategic facility.

Positional angle (mil): The angle from the strategic facility to the suspected targets in the airspace.

Navigational angle (°): The navigational direction of the suspected targets in the airspace.

Horizontal velocity (m/s^{-1}): The velocity in the horizontal direction of the suspected targets.

Electromagnetic interference degree: The electromagnetic interference degree by the suspected targets on the radar and communications. This is a subjective factor mainly derived by experts.

Radar cross section (m^2): The size of the suspected targets on the radar screen.

Intension: A subjective factor that depicts the possible intentions of the suspected targets, namely attack (A), cover (C), or surveillance (S).

The assessment result is the threat level: high, medium, low, or unknown.

A total of 14 sets of data gathered from a historic database were used as the training data. A total of 12 sets of data were used as the testing data by gathering information on the new suspected targets.

G. Case III: Model construction and optimization

It is improbable to construct a conjunctive BRB for this problem because there would be $3^7/4^7/5^7$ rules if each influential factor (attribute in BRB) had 3/4/5 referenced values. Specifically for Case III, only a disjunctive BRB is applicable. Thus, a disjunctive BRB with 5 rules was constructed for this case as follows.

First, identify the lower and upper bounds of the influential factors as the attributes of the BRB model. The lower and upper bounds for the attributes are 50/100/120/5/1/1/S, and 6500/300/350/350/10/7/A, respectively.

Then, three experts were invited to each produce a disjunctive rule based on their knowledge and experience, as well as their observation of the training dataset.

All three experts agreed upon the two following rules:

$$\begin{aligned} &R_1: \text{if} \quad (x_1 \text{ is } 50) \vee (x_2 \text{ is } 100) \vee \cdots \vee (x_7 \text{ is S}), \\ &\quad \text{then } \{(L,0.6),(M,0.2),(H,0.2)\} \quad \text{with } \theta_1 = 1 \\ &R_2: \text{if} \quad (x_1 \text{ is } 6500) \vee (x_2 \text{ is } 300) \vee \cdots \vee (x_7 \text{ is A}), \\ &\quad \text{then } \{(L,0.2),(M,0.2),(H,0.6)\} \quad \text{with } \theta_2 = 1 \end{aligned} \tag{34}$$

Rule 1 states that when either of the attributes is at its lower bound, then the threat level is “low” with a belief of “0.6”. Rule 2 denotes that when either of the attributes is in its upper bound, then the threat level is “high” with a belief of “0.6”. Because these are disjunctive rules, the beliefs are not given in an absolute fashion (no belief for any scale is 100%).

Then, the experts were asked to provide one additional rule based on their knowledge and the training data.

Expert 1 provided the rule as in (35). His reasons were (1) if the positional angle is relatively low, or (2) if the radar cross section is small, or (3) if the distance is relatively close, then the threat level is medium with a high certainty.

$$\begin{aligned} R_3: \text{if} \quad & (x_1 \text{ is } 2500) \vee (x_2 \text{ is } 150) \vee (x_3 \text{ is } 200) \vee (x_4 \text{ is } 200) \vee \\ & (x_5 \text{ is } 8) \vee (x_6 \text{ is } 3) \vee (x_7 \text{ is ((S,50\%),(C,50\%))}), \\ \text{then } & \{(L,0.15),(M,0.7),(H,0.15)\} \quad \text{with } \theta_3 = 1 \end{aligned} \tag{35}$$

Expert 2 provided the rule as in (36). His reasons were (1) if the distance is far, or (2) if radar cross section is large, or (3) if the intention is classified as “cover” or “attack”, or (4) if the positional angle is small, then the threat level is high with a high certainty.

$$\begin{aligned} R_4: \text{if} \quad & (x_1 \text{ is } 5000) \vee (x_2 \text{ is } 110) \vee (x_3 \text{ is } 150) \vee (x_4 \text{ is } 150) \vee \\ & (x_5 \text{ is } 2) \vee (x_6 \text{ is } 8) \vee (x_7 \text{ is ((C,50\%),(A,50\%))}), \\ \text{then } & \{(L,0.1),(M,0.3),(H,0.6)\} \quad \text{with } \theta_3 = 1 \end{aligned} \tag{36}$$

Expert 3 provided the rule as in (37). His reasons were (1) if the distance is relatively far, or (2) if the positional angle is small, or (3) if the horizontal velocity angle is small, then the threat level is relatively low.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2892348, IEEE Transactions on Fuzzy Systems

$$\begin{aligned} R_5: \text{if}\quad & (x_1 \text{ is } 4000) \vee (x_2 \text{ is } 120) \vee (x_3 \text{ is } 180) \vee (x_4 \text{ is } 100) \vee \\ & (x_5 \text{ is } 6) \vee (x_6 \text{ is } 6) \vee (x_7 \text{ is } ((\text{C},30\%),(\text{A},70\%))), \\ & \text{then } \{(L,0.6),(M,0.2),(H,0.2)\} \quad \text{with } \theta_5 = 1 \end{aligned} \tag{37}$$

The above five disjunctive rules are used as the initial BRB with updated orders.

Next, the optimization model and algorithm proposed in Parts C and D in Section IV is used to train and update the initial BRB.

For the optimization model, the lower and upper bounds are defined as follows:

$$\begin{aligned} lb_m &= [50 \;\; 100 \;\; 120 \;\; 5 \;\; 1 \;\; 1 \;\; S], m = 1, \ldots, 7 \\ ub_m &= [6500/300/350/350/10/7/A], m = 1, \ldots, 7 \end{aligned} \tag{38}$$

Since the two rules agreed by the three experts are with the lower and upper bounds of the attributes, the referenced values for the attributes in the two rules are free from training:

$$A_m^1 = lb_m; A_m^M = ub_m \tag{39}$$

Since there is no incomplete information, thus,

$$\sum_{n=1}^{N} \beta_{n,k} = 1 \tag{40}$$

Although there is no "unknown" in this case, the differentiation ratio v is not set to be "0",

$$0 < v < 1 \tag{41}$$

With θ_k and $\beta_{n,k}$ being $(0,1]$ and $[0,1]$, respectively, the optimization model is given as follows:

$$\min \quad MSE(A_m^k, \theta_k, \beta_{n,k}, v) \tag{42}$$

s. t.

$$lb_m \le A_m^k \le ub_m; A_m^1 = lb_m; A_m^M = ub_m \tag{42a}$$

$$0 < \theta_k \le 1; 0 < v < 1 \tag{42b}$$

$$0 \le \beta_{n,k} \le 1; \sum_{n=1}^{N} \beta_{n,k} = 1 \tag{42c}$$

where $k = 1, \ldots, K; n = 1, \ldots, N; m = 1, \ldots, M$.

To solve the above optimization model, the classic DE algorithm is used again [20] with following settings: the population was set at 20; the generation was set at 500; the crossover rate was 0.9; the mutation rate was 0.5; $v = 0.1$. This experiment was conducted 30 times.

After optimization, the optimized model is given in Table IX with $v = 0.1946$, indicating that the differentiation ratio was significantly bigger than the initial estimation. It indicates that the optimized BRB model had higher differentiation ability.

Remark 8. After optimization, the referenced values remain rather close to those provided by the experts. However, the weights and beliefs of the scales of the rules have changed. Rules 2/5 are assigned relatively small weights, denoting that the two rules are of smaller confidence. Moreover, the beliefs for "low" and "high" are rather close for the optimized Rules 1/2, which is rather interesting. This is probability caused by the conflicted referenced values for different attributes. The optimized Rule 5 also shows such conflict result: the belief for "low" is higher than that for "medium".

TABLE IX. UPDATED BRB FOR CASE III OPTIMIZED BY DE

No.	weight	Factors							threat Level		
		1	2	3	4	5	6	7	low	medium	high
1	0.9613	50.0000	100.0000	120.0000	5.0000	1.0000	1.0000	(S, 100%)	0.5786	0.1118	0.3097
2	0.6862	4003.1790	120.8667	187.4815	115.5275	6.4264	6.1214	(C, 24.84%), (A, 75.16%)	0.5558	0.1059	0.3383
3	0.8954	2450.7829	166.1030	222.8361	214.5813	8.0451	3.1496	(S, 42.61%), (C, 57.39%)	0.1312	0.6938	0.1750
4	0.9798	5345.8261	105.4191	132.2365	151.8877	2.1009	6.9262	(C, 17.86%), (A, 82.14%)	0.1066	0.3231	0.5703
5	0.6164	6500.0000	300.0000	350.0000	350.0000	10.0000	7.0000	(S, 100%)	0.2525	0.1704	0.5770

H. Case III: Comparative results

The Bayesian approach, Support Vector Machine (SVM), and Neural Network (NN) were used in a comparative fashion for verification. The results are shown in Table X.

TABLE X. RESULTS COMPARISON FOR TESTING DATASET IN CASE III

No.	Min	Avg	Var
Initial BRB by experts		0.1667	
Optimized BRB by DE	**0**	**0.0806**	**1.6683E-03**
Bayesian		0.2500	
SVM		0.8690	
LSSVM		0.7500	
Neural network	**0.0833**	0.1722	1.5294E-02

For the Bayesian approach [12], the Matlab toolbox was used. The first six factors were divided into two partitions (Bayesian cannot process the qualitative index). The MSE for the testing dataset attained 0.25. We further increased the partitions to 3/4/5; however, the MSE also increased. The MSE for the training dataset remained zero when the partition number increased; this result is partially because of the fact that more partitions for Case III result in over-fitting.

For SVM, SMO in Weka [27] and LSSVM [28] were applied. For SMO, the batch size was set at 100, *c* was set to one, and the *PolyKernel* was used as the kernel function. Moreover, 10-fold cross validation was applied. The result for SMO was not acceptable; the MSE for the testing dataset was 0.8690. For LSSVM, the *RBF_kernel* was used as the kernel function with $\Gamma = 5000$ and $\sigma^2 = 80$. The result for LSSVM was also not satisfactory: the MSE for the testing dataset was 0.7500. This is partially because the training dataset is overly small, which does not support constructing a precise model.

For NN [3], the parameters were set as follows: the epoch maximum of was set to 5000 (although all of the 30 runs reached the stop optimal objective within 500 epochs), the goal was set as 3E-08, the number of middle transit neurals was set as 15, the middle nodes layer was set at three, and the middle neural transit/ output neural transit / network training functions were set as *tansig* / *logsig* / *trainscg*, respectively.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2892348, IEEE Transactions on Fuzzy Systems

Thirty runs were conducted for NN. The minimal MSE for the testing dataset was 0.0833 for 20 runs. However, in the remaining ten runs, the MSEs were as high as 0.3333/0.4167 indicating that 4/5 out of 12 testing cases could not be accurately identified. This is partially because neural network suffers from over-fitting because the MSEs for the training dataset all reach zero within only a small number of epochs.

Comparing the modeling process and results of the different approaches, the following conclusions are drawn:

(1) BRB demonstrated superior performance in comparison of MSEs with the other approaches. BRB by the experts could identify 10 out of 12 cases (error of 0.1667), similar to NN and outperform the other approaches; the optimized BRB produced the optimal result with 100% accuracy. Moreover, the optimized BRB was the most robust in comparing the variances.

(2) The result by Bayesian was acceptable. It suffered from the over-fitting problem because the MSE for the training dataset reached zero quickly. The performance of Bayesian could be improved by optimizing its parameters.

(2) NN demonstrated acceptable nonlinearity modeling ability even without considering the experts knowledge. However, its stability was less than BRB because the variance of 30 runs was relatively greater than that of BRB. Moreover, it also suffered from the over-fitting problem, similar to Bayesian.

(4) Comparatively, SVM failed in solving this case because it could not produce satisfactory results. This could be partially because there was not sufficient data to train the model.

(5) For all four approaches, the parameter configurations have significant influence on their modeling ability. By further optimizing the parameters, the results in Table X could be further improved. However, further discussion of this topic is outside of the main purpose of this study.

(6) As demonstrated in Case III, one of the main advantages of BRB is its openness and accessibility to experts and decision makers. From the initial modeling, to the settings of optimization parameters, to the final optimized BRBs as well as the assessment results, experts and decision makers can participate and understand each and every step. In one word, it is a white box which can provide direct access to humans.

I. Further Discussion

As the disjunctive BRB is designed to downsize BRB especially when faced with the combinatorial explosion problem, it is in nature an effective tool in complex systems modeling. So far, many approaches have been development for the same purposes, among which deep neural network and the fuzzy system are two powerful techniques.

For deep neural network, it has shown very outstanding performances in solving many theoretical and practical cases in recent years [23], especially by the success of AlphaGo [24]. However, as it is based on the neural network, it is still a black box approach. Of course, being a black box approach does not invalidate it but certainty makes it less applicable for problems that require decision process transparency. Back to its merits, deep neural network has the ability of identifying the hidden structure in a sense of Reinforcement Learning (RL), which could be very useful for BRB structure learning. Likewise, as one of BRB's greatest advantages is the modeling transparency, it could be used to designing more accessible networks for deep neural network so as to provide better understandability for decision makers.

Another powerful tool is the fuzzy systems [16]. A fuzzy system is comprised of multiple fuzzy rules while BRB is comprised of multiple belief rules. They both can very well represent and model multiple types of information with ambiguous and fuzzy characteristics. From this perspective, they both have their merits. As for the differences, the root lies in that their theoretical bases are different: the fuzzy system approach is based on fuzzy set theory while BRB is based on the D-S evidence theory. They can also help each other by providing their state-of -the –art merits. One of the advantages of the fuzzy system approach is that their knowledge presentation base for human linguistic terms is far richer, e. g., intuitionistic terms, hesitant linguistic terms, etc [16]. It can be of great use for BRB. Likewise, BRB has developed a knowledge presentation based on the power set which can also be helpful for the fuzzy systems [19] [31] [36].

In one word, multiple approaches can and should work together to solve a complex problem when needed. Even in developing AlphaGo, the deep neural network and the Monte-Carlo simulation are combined [24].

Another important question is the efficiency of the proposed disjunctive BRB approach. Verified by three case studies, its efficiency has been preliminary validated. But, we would like to further discuss this matter from two perspectives, against all problems in a general sense and concerning its computation performance or real time performance.

First of all, the disjunctive BRB approach is not designed as a universal approach for all complex systems modeling problems. The very motive of proposing such a disjunctive BRB approach is to solve the combinatorial explosion problem faced by the conventional conjunctive BRB modeling. The disjunctive BRB has been used in solving several problems in previous studies, e. g., the classification problems [7], the gas turbine signal inference [5], and also the cases in this study.

There are mainly two ways of using the disjunctive BRB approach. The first is to use it as a data-driven approach, e. g., Case II in this study. If so, it is applicable to many problems with tagged labels as its optimization process is essentially supervised learning (SL) which normally requires a larger amount of data for training and learning. Another way is to use it as an expert system approach, e. g., Case III in this study.. If so, experts and decision makers can be directly involved in each and every step of the modeling, inferencing, and even the optimizing process. This is especially fit for problems without many historic data but be made up by experts' involvement. Either way, the downsized disjunctive BRB can be quite effective by integrating multiple types of information under uncertainty and providing direct access to experts and decision makers.

Concerning its real time performance, the BRB modeling process requires some time to collect relative data and information from heterogeneous sources. For a BRB with many rules, this process can be quite time-consuming which goes back the very motive of this study: to downsize BRB. For the BRB inferencing process, it is an analytical process and the computation time is directly related the number of parameters (or the rules) [7]. After all, the inferencing process is normally implemented in programming (as in MATLAB in this study) which could be finished instantly. For the optimization process, it depends on the selected optimization engine. If deterministic techniques are applied, the real time ability can be ensured [38]. However, if EAs are applied as the optimization engine, then the time required would be also related to certain parameter settings [6] [35], including the number of generation and individuals, and also the selection of optimization strategies (mutation/crossover/selection operators), etc.

In fact, "real time" is a relative concept. Besides the optimization techniques, the computational power also plays an important role. For example, using 40 search threads, 48 CPUs and 8 GPUs, even complex problems such as Go with $b^d (b \approx 250, d \approx 150)$ possible sequences of moves in a search tres be solved within seconds whereas normal computers could take months if not years [24].

However, there are not such abundant resources for most problems. Thus, in practical conditions when BRB has been initialized, an online approach can be developed with deterministic optimization techniques whereas only offline approach is possible if EAs are applied.

VI. Conclusion

This study proposed a generic modeling, inferencing, and optimization approach for disjunctive BRB. The main motive of proposing the generic approach is to solve the combinatorial explosion problem for the conjunctive BRB when there are over-numbered (referenced values for) the attributes.

This study focuses on two aspects. The first is to mathematically define the BRB space as an attribute product measure space. It was proved that as long as the base of the BRB space is included, a BRB is complete regardless of whether it is conjunctively or disjunctively constructed.

The second is to propose a generic modeling, inferencing, and optimization approach for disjunctive BRB based on the mathematical definition of the BRB space. Details procedures are given, including generic BRB construction, rule-activation, and matching-degree calculation procedures, as well as a generic optimization model and optimization algorithm.

Three cases were studied for verification. It was validated that disjunctive BRB can facilitate downsizing BRBs and maintain a rather high accuracy. With the disjunctive BRB modeling, inferencing, and optimization approach, the combinatorial explosion problem can be addressed effectively. Compared with other data-driven approaches, BRB can support gathering, integrating, and inferring based on experts' knowledge, especially when there is insufficient data.

A slight difference between Case II and Case III should be made clear. In Case II, BRB is used more as a data-driven approach while it is more as an expert system approach in Case III. Both are validated by case study results with superior performances. In practical conditions, it is not easy or even unnecessary to differentiate them from each other. As long as the application requirements are met, BRB can be an effective modeling tool with direct access to human knowledge.

Further discussions on comparison with deep neural network and fuzzy systems, real time performance and efficiencies have been conducted as well.

For future research, additional theoretical and practical work addressing disjunctive BRB construction as well as its training and learning should be undertaken. Moreover, the proposed approach should also be tested on more open benchmark cases for verification.

Appendix

The proof is given in the supplementary material.

References

[1] H. Akaike, *Akaike's Information Criterion*, Springer Berlin Heidelberg, 2011.
[2] V. I. Bogachev. *Measure Theory*. Higher Education Press, 2010.
[3] G. A. Carpenter, S. Martens, and O. J. Ogas, "Self-organizing information fusion and hierarchical knowledge discovery: a new framework using ARTMAP neural networks," *Neural Netw.*, vol. 18, pp. 287-295, 2015.
[4] L. L. Chang, Y. Zhou, J. Jiang, M. J. Li, and X. H. Zhang, "Structure learning for belief rule base expert system: a comparative study," *Knowl.-Based Syst.*, vol. 39, pp. 159-172, 2013.
[5] L. L. Chang, Z. J. Zhou, Y. W. Chen, T. J. Liao, Y. Hu, and L. H. Yang, "Belief Rule Base Structure and Parameter Joint Optimization under the Disjunctive Assumption for Nonlinear Complex System Modeling," *IEEE Trans. Syst., Cybern., Syst.*, vol. PP, pp.1-13, 2017.
[6] L. L. Chang, Z. J. Zhou, Y. W. Chen, X. B. Xu, J. B. Sun, and T. J. Liao, "Akaike information criterion-based conjunctive belief rule base learning for complex system modeling", *Knowl.-Based Syst.*, accepted.
[7] L. L. Chang, Z. J. Zhou, Y. You, L. H. Yang, and Z. G. Zhou, "Belief rule based expert system for classification problems with new rule activation and weight calculation procedures," *Inform. Sci.*, vol. 336, no.1, pp. 75-91, 2016.
[8] Y. Chen, Y. W. Chen, X. B. Xu, C. C. Pan, J. B. Yang, and G. K. Yang, "A data-driven approximate casual inference model using the evidential reasoning rule," *Knowl.-Based Syst.*, vol. 88, pp. 264-272, 2015.
[9] Y. W. Chen, J. B. Yang, C. C. Pan, D. L. Xu, and Z. J. Zhou, "Identification of uncertain nonlinear systems: Constructing belief rule-based models," *Knowl.-Based Syst.*, vol. 73, pp.124-133, 2015.
[10] Y. W. Chen, J. B. Yang, D. L. Xu, and S. L. Yang, "On the inference and approximation properties of belief rule based systems," *Inform. Sci.*, vol. 234, pp.121-135, 2013.
[11] Y. W. Chen, J. B. Yang, D. L. Xu, Z. J. Zhou, and D. W. Tang, "Inference analysis and adaptive training for belief rule based system," *Expert Syst. Appl.*, vol. 38, pp. 12845-12860, 2011.
[12] S. Das, R. Grey, and P. Gonsalves, "Situation assessment via Bayesian belief networks," *Proc. Int. Conf. Inform. Fusion*, pp. 664-671, 2002.
[13] T. Denux, N. E. Zoghby, V. Cherfaoui, and A. Joulet, "Optimal Object Association in the Dempster–Shafer Framework," *IEEE Trans. Syst., Man, Cybern., B, Cybern.*, vol. 44, pp. 2521-2531, 2014.
[14] D. Molina, A. Latorre, and F. Herrera, "An Insight into Bio-inspired and Evolutionary Algorithms for Global Optimization: Review, Analysis, and Lessons Learnt over a Decade of Competitions," 2018, [Online]. Available: https://doi.org/10.1007/s12559-018-9554-0.
[15] G. L. Li, Z. J., Zhou, C. H. Hu, L. L. Chang, Z. G. Zhou, and F. J., Zhao, "A new safety assessment model for complex system based on the conditional generalized minimum variance and the belief rule base," *Safety Sci.*, vol. 93, pp. 108-120, 2017.
[16] H. Liao, X. M. Mi, Z. S. Xu, J. P. Xu, and F. Herrera, "Intuitionistic fuzzy analytical network process," *IEEE Trans. Fuzzy Syst.*,, vol. 26, pp. 2578-2590, 2018.
[17] B. D. Liu, *Uncertainty theory*, 5th ed., Springer-Verlag, Berlin, 2015.
[18] J. Liu, L. Martinez, A. Calzada, and H. Wang, "A novel belief rule base representation, generation and its inference methodology," *Knowl.-Based Syst.*, vol. 53, pp. 129-141, 2013.
[19] Z. G. Liu, Q. Pan, J. Dezert, and A. Martin, "Combination of classifiers with optimal weight based on evidential reasoning," *IEEE Trans. Fuzzy Syst.*,, vol. 26, pp. 1217-1230, 2018.
[20] K. Price, R. Kenneth, and J. Lampinen, *Differential Evolution-A Practical Approach to Global Optimization*, Springer Science & Business Media, Berlin, 2006.
[21] K. K. Rohitha, G. K. Hewawasam, K. Premaratne, and M. L. Shyu, "Rule Mining and Classification in a Situation Assessment Application: A Belief-Theoretic Approach for Handling Data Imperfections," *IEEE Trans. Syst., Man, Cybern., B, Cybern.*, vol. 37, pp. 1446-1459, 2007.
[22] E. E. Savan, J. B. Yang, D. L. Xu, and Y. W. Chen, , "A genetic algorithm search heuristic for belief rule-based model-structure validation," *Proc. 2013 IEEE Int. Conf. Syst., Man, Cybern.*, pp. 1373-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2892348, IEEE Transactions on Fuzzy Systems

1378, 2013.

[23] J. Schmidhuber, “Deep learning in neural networks: an overview,” *Neural Netw.*, vol. 61, pp. 85-117, 2015.

[24] D. Silver, A. Huang, C. J. Maddison, et al, “Mastering the game of Go with deep neural networks and three search,” *Nature*, vol. 529, pp. 484-489, 2016.

[25] Y. M. Wang, L. H. Yang, Y. G. Fu, L. L. Chang, and K. S. Chin, “Dynamic rule adjustment approach for optimizing belief rule-base expert system,” *Knowl.-Based Syst.*, vol. 96, pp. 40-60, 2016.

[26] Y. M. Wang, J. B. Yang, and D. L. Xu, “Environmental impact assessment using the evidential reasoning approach,” *Eur. J. Oper. Res.*, vol. 174, no. 3, pp. 1885-1913, 2006.

[27] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*, 4th ed., Morgan Kaufmann, 2016.

[28] J. Xu, X. Zhang, and Y. Li, “Kernel MSE algorithm: a unified framework for KFD, LS-SVM and KRR,” *Proc. IJCNN'01.*, pp. 1486-1491, 2001.

[29] D. L. Xu, J. Liu, J. B. Yang, G. P. Liu, J. Wang, I. Jenkinson, and J. Ren, “Inference and learning methodology of belief-rule-based expert system for pipeline leak detection,” *Expert Syst. Appl.*, vol. 32, pp. 103-113, 2007.

[30] X. B. Xu, J. Zheng, J. B. Yang, D. L. Xu and Y. W. Chen, “Data classification using evidential reasoning rule,” *Knowl.-Based Syst.*, vol. 116, pp.144-151, 2017.

[31] J. B. Yang, J. Liu, J. Wang, H. S. Sii, and H. W. Wang, “Belief rule-base inference methodology using the evidential reasoning approach-RIMER,” *IEEE Trans. Syst., Man, Cybern., A, Syst., Humans*, vol. 36, pp. 266-285, 2006.

[32] J. B. Yang and D. L. Xu, “Nonlinear information aggregation via evidential reasoning in multiattribute decision analysis under uncertainty,” *IEEE Trans. Syst., Man, Cybern., A, Syst., Humans*, vol. 32, pp. 376-393, 2002.

[33] L. H. Yang, Y. M. Wang, L. L. Chang, and Y. G. Fu, “A disjunctive belief rule-based expert system for bridge risk assessment with dynamic parameter optimization model,” *Knowl.-Based Syst.*, vol. 113, pp. 459-474, 2017.

[34] L. H. Yang, Y. M. Wang, J. Liu, and L. Martines, “A joint optimization method on parameter and structure for belief-rule-based systems,” *Knowl.-Based Syst.*. vol. 142, pp. 220-240, 2018.

[35] Z. J. Zhou, L. L. Chang, C. H. Hu, X. X. Han, and Z. G. Zhou, “A New BRB-ER-Based Model for Assessing the Lives of Products Using Both Failure Data and Expert Knowledge,” *IEEE Trans. Syst., Cybern., Syst.*, vol. 46, pp. 1529-1543, 2016.

[36] Z. J. Zhou, C. H. Hu, G. Y. Hu, X. X. Han, B. C. Zhang, and Y. W. Chen, “Hidden behavior prediction of complex systems under testing influence based on semi-quantitative information and belief rule base,” *IEEE Trans. Fuzzy Syst.*, vol. 23, pp.2371-2386, 2015.

[37] Z. J. Zhou, C. H. Hu, J. B. Yang, D. L. Xu, M. Y., Chen, and D. H. Zhou, “A sequential learning algorithm for online constructing belief-rule-based systems,” *Expert Syst. Appl.*, vol. 37, pp. 1790-1799, 2010.

[38] Z. J. Zhou, C. H. Hu, J. B. Yang, D. L. Xu, and D. H. Zhou, “Online updating belief rule based system for pipeline leak detection under expert intervention,” *Expert Syst. Appl.*, vol. 36, pp. 7700-7709, 2009.

Lei-lei Chang received a B.Eng degree from Central South University, Changsha, China, in 2008 and M. Eng and Ph.D. degrees from the National University of Defense Technology, Changsha, China, in 2010 and 2014, respectively.

He is currently a lecturer at the High-Tech Institute of Xi’an. He has authored three books and published approximately 20 articles. His research interests include BRB structure and parameter learning, especially in disjunctive BRB optimization for complex system modeling.

Zhi-Jie Zhou received B.Eng and M. Eng degrees from the High-Tech Institute of Xi’an, Xi’an, China, in 2001 and 2004, respectively, and a Ph.D. degree from Tsinghua University, Beijing, China, in 2010.

He is a professor at the High-Tech Institute of Xi’an. He has published approximately 70 articles. His research interests include belief rule base, dynamic system modeling, and hybrid quantitative and qualitative decision modeling.

Huchang Liao (M’13–SM’17) received his Ph.D. degree in Management Science and Engineering from the Shanghai Jiao Tong University, Shanghai, China, in 2015.

Dr. Liao is currently a Research Fellow in the Business School at the Sichuan University, Chengdu, China. His current research interests include decision analysis under uncertainty, fuzzy set and fuzzy systems, computing with words, intelligent decision support systems, healthcare management, etc.

Yu-Wang Chen received a Ph.D. degree in control theory and control engineering from Shanghai Jiao Tong University, Shanghai, China, in 2008.

He is currently a Senior Lecturer in decision sciences in the Alliance Manchester Business School, University of Manchester, Manchester, U.K. His research interests include decision and risk analysis under uncertainties, modeling and optimization of complex systems, operational research, and data analytics.

Xu Tan received the Ph.D degree in management science and engineering from National University of Science and Technology in 2009.

He is now a professor at the Shenzhen Institute of Information Technology, P. R. China. He has received several grants from both the National Science Foundation and the Ministry of Education in China. He has also joined the Pengcheng Scholar Funded Scheme which is a special grants program for decorated scholars. His research interests include intelligent decision, machine learning

Francisco Herrera (SM’15) received his M.Sc. in Mathematics in 1988 and Ph.D. in Mathematics in 1991, both from the University of Granada, Spain. He is currently a Professor in the Department of Computer Science and Artificial Intelligence at the University of Granada. He currently acts as Editor in Chief of the international journals "Information Fusion" (Elsevier) and “Progress in Artificial Intelligence (Springer). His current research interests include among others, soft computing (including fuzzy modeling and evolutionary algorithms), information fusion, decision making, biometric, data preprocessing, data science and big data.