IMPROVING INTERNAL-VALUED INFERENCING WITH LIKELIHOD RATIO

Lim Chee Kau

Department of Artificial Intelligence, Faculty of Computer Science and Information Technology,
University of Malaya, Kuala Lumpur, Malaysia

E-mail: limck@um.edu.my

DOI: https://doi.org/10.22452/mjcs.vol32no3.3

ABSTRACT

This paper point out the limitations of Interval-Valued Inferencing as a defuzzification method for inference engines based on the Bandler-Kohout subproduct. As an improvement, a measurement on the likelihood of an inference result in an acceptance/rejection band suggested. With this improvement, more meaningful results are generated from a Bandler-Kohout subproduct based inference system, especially if it is implemented as a medical decision support system. To demonstrate the capability of this improvement, an experiment with a popular dataset is carried out.

Keywords: BK Subproduct, Interval-Valued Inferencing, Defuzzification

1.0 INTRODUCTION

Over the past few decades, fuzzy logic and related technologies have been widely used to deal with problems such as software development [1,2], computer vision [3,4], network management [5,6], data analysis [7,8] and medical [9,10]. To cater the problems of different fields, various technologies have been proposed, for example fuzzy logic controller [1,5,6], fuzzy clustering [4], fuzzy regression [7] and some even combine fuzzy logic with other technologies such as artificial neural networks [2,9] and genetic programming [8,10].

Among all, fuzzy logic systems based on the Bandler-Kohout (BK) subproduct [11,12] received high reputation with their remarkable success. As a study of fuzzy relations, the BK subproduct finds its advantages in developing inference engines for numerous applications, such as medical expert systems [13–16], path finding of autonomous underwater vehicles [17,18], land evaluation [19], scene understanding [20], human motion analysis [21] and etc. Recently, a study of Stepnicka and Jayaram [22] has also proved that the BK subproduct is as outstanding as the popular Compositional Rule of Inference (CRI)[23] in term of both efficiency and effectiveness.

Although the mechanism of inference engines developed by the BK subproduct work differently compare to the CRI-based inference engines, but the flow of information processing in both inference models are somewhat of similar (Fig. 1). For both models, an inference process starts with the fuzzifying of input data. Next, the inference engine performs computations on the fuzzified data based on the predefined rules or information stored in knowledge base. Once all input signals are processed, an aggregator is used to aggregate the results. Finally, a defuzzification module is used to convert the output, which is in the form of fuzzy membership, to useful information. Defuzzification modules are required by fuzzy systems so that the fuzzy results produced by the inference engines can be interpreted by subsequent applications that only deal with binary logic. The discussion of this paper focuses on the defuzzification module of BK subproduct based fuzzy logic systems.

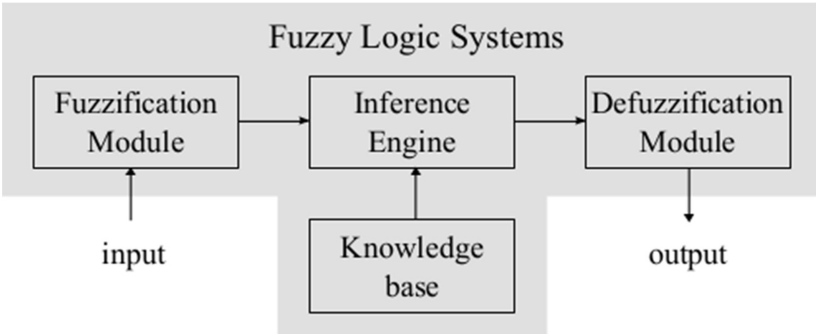

Fig. 1. Typical structure of a fuzzy logic system.

One of the interesting properties of BK subproduct is it performs inferences using fuzzy implication operators, which can be defined based on the needs of applications. With this special property, Interval-Valued Inferencing (IVI) was proposed in [24] to obtain more reliable results in reasoning. With IVI, an interval, instead of a point-value, is generated as the output of an inference. The advantage of having intervals as outputs makes BK subproduct-based inference engines more reliable, especially when it is acquired as an inference engine in a medical decision support system [13], [14].

However, despite of the great idea, two drawbacks are found in this IVI technique, especially when it is implemented in medical decision support systems. Firstly, effectiveness is low in many cases because of redundant computations. For an inference, this algorithm requires two computations with different fuzzy implication operators on a set of input. This is beyond reproach if both computations provide two information which are both helpful in decision making. However, in many cases the algorithm only make use on one of the two outputs. Secondly, this algorithm leads to binary classifications (i.e. an inference can only be 'accepted' or 'not accepted'), which is not a desire feature in medical decision support systems. In the implementation of medical decision support systems, medical experts see technologies as 'mediators', rather than a replacement [25]. Therefore, the challenge of this paper is to improve the classical IVI so that with no additional computational cost to the inference engine, the inference results can be interpreted effectively. Consequently, information-rich recommendations, which are preferred [26] can be provided rather than a simple classification results.

The objective of this paper is twofold. First, we want to improve the effectiveness of IVI with a newly proposed likelihood ratio, so that the results of computations with both fuzzy implication operators are utilized. Second, to provide a defuzzification algorithm for the medical decision support systems so that information-rich recommendations, but not classifications are delivered. On the other hand, this paper has no intention to investigate the proper acceptance and rejection threshold values for inference engines. The determination of these thresholds is another open research problem that independent from the algorithm discussed in this paper.

Improving the IVI with the likelihood ratio is significant. It is because with a better defuzzification algorithm, the results from inference engines can be interpreted accurately and effectively. In the case of medical decision support systems, applying this likelihood ratio in classification of patients not only return crisp classifying results, but also the likelihood of each patient suffering with the disease. In addition, users of this defuzzification algorithm have the flexibility of fixing a convenience likelihood ratio (for example, likelihood ratio ≥ 0.5 as demonstrate in the experiment in Section 5) to return strong classification results.

The rest of this paper is organized as follow: BK subproduct is explained in Section 2, along with the IVI defuzzification algorithm in some of its applications. The shortcomings of IVI in BK subproduct based systems are explained in Section 3. Section 4 explains the improved defuzzification algorithm, which involves the calculation of likelihood ratio of inferences. Experiment on this defuzzification algorithm using common dataset is discussed in Section 5. Lastly, the conclusion of this paper is drawn in Section 6

2.0 BK SUBPRODUCT REVISIT

2.1 BK Subproduct

We start the discussion with a short revision on the BK subproduct in crisp sets. Assume that there are set $A = \{a_i \mid i = 1, \ldots, m\}$ and set $B = \{b_j \mid j = 1, \ldots, n\}$. R is defined as a relation from A to B such that $R \subseteq A \times B$. If there is a set $C = \{c_k \mid k = 1, \ldots, k\}$ such that A has no direct relation with C, but B is in relation S with C ($S \subseteq B \times C$), we can trace the indirect relation from A to C in term of relations R and S using the BK subproduct, denoted $R \triangleleft S$:

$$R \triangleleft S = \{(a,c) \mid (a,c) \in A \times C \text{ and } aR \subseteq Sc\} \quad (1)$$

In (1), aR is the image of a in set B under relation R, whereas Sc is the image of c in B under converse relation of S, i.e. S^T . One can see that BK subproduct gives all ordered pairs of (a,c) such that the image of a under relation R in B is among the subset of c under S^T in B (Fig. 2). With this composition of relations, one can find the relation between an object, $a \in A$ and target, $c \in C$ if a set with common features, $B' \subseteq B$ appears in the middle.

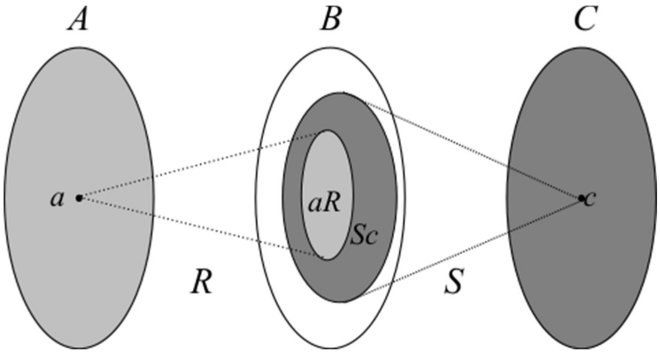

Fig. 2. Using BK subproduct, the relations between 2 sets which are not related directly can be retrieved if a set with common features exist.

Therefore, when a particular $a' \in A$ is concerned, using the BK subproduct is helpful in retrieving a set of $c \in C$ where $a'R$ is a subset of Sc for all the c. This makes BK subproduct becomes a tool in retrieving indirect relations:

$$a'(R \triangleleft S) = \{c \mid c \in C \text{ and } a'R \subseteq Sc\} \quad (2)$$

For example, if A is a set of patients and C is a set of diseases, indirect relations between these two sets can be defined through a set of signs and symptoms, B. Therefore, R is the relation that identifying the signs and symptoms developed by a patient, whereas S shows signs and symptoms that used to characterize a disease. If a particular patient a' developed a set of signs and symptoms, $B' = a'R$, and this $a'R$ is a subset of Sc', then we can conclude that all the signs and symptoms that developed by patient a' are part of the signs and symptoms of disease(s) c'. In another word, disease(s) c' could be the disease(s) that patient a' suffer from.

As one can notice, the central thought of BK subproduct is a subsethood measure. Therefore, this crisp BK subproduct can be extended to fuzzy version easily by introducing a fuzzy subsethood measure - i.e. the possibility of a set is the subset of another given set. As proposed in [11], in the same universe $X = \{x_i \mid i = 1, \ldots, n\}$, the degree of subsethood of a fuzzy set A in B, $\pi(A \subseteq B)$ can be expressed as :

$$\pi(A \subseteq B) = \bigwedge_{x \in U} \big(A(x) \rightarrow B(x)\big) \quad (3)$$

where $\bigwedge$ is defined as infimum operator in harsh criterion, or arithmetic mean in mean criterion, $A(x)$ and $B(x)$ are membership functions of x in A and B respectively, and $\rightarrow$ denotes the fuzzy implication operator. There is a number of definitions for the fuzzy implication operator in the literature (Table 1).

Table 1. Fuzzy Implication Operators

Name	Symbol	Definition
S# - Standard Sharp	$r \rightarrow_{S\#} s$	$\begin{cases} 0, & \text{iff } r \neq 1 \text{ or } s = 1 \\ 1, & \text{otherwise} \end{cases}$
S - Standard Strict	$r \rightarrow_{S} s$	$\begin{cases} 0, & \text{iff } r \leq 1 \\ 1, & \text{otherwise} \end{cases}$
S* - Standard Star	$r \rightarrow_{S*} s$	$\begin{cases} 0, & \text{iff } r\text{f } r \leq s \\ 1, & \text{otherwise} \end{cases}$
G43 - Gaines 43	$r \rightarrow_{G43} s$	$\min(1, \frac{r}{s})$
G43' - Modified Gaines 43	$r \rightarrow_{G43'} s$	$\min(1, \frac{r}{s}, \frac{1-r}{1-s})$
KD - Kleene-Dienes	$r \rightarrow_{KD} s$	$\max(s, 1-r)$
R - Reichenbach	$r \rightarrow_{R} s$	$1-r+rs$
L - Łukasiewicz	$r \rightarrow_{Ł} s$	$\min(1, 1-r+rs)$

With this fuzzy subsethood measure definition, it is easy to rewrite (1), the expression of BK subproduct in fuzzy environment. For the case of 2 elements $a \in A$ and $c \in C$, the fuzzy relation from a to c is given as:

$$R \lhd S(a,c) = \bigwedge_{b \in B} \big(R(a,b) \rightarrow S(b,c)\big) \tag{4}$$

The expression returns a value in the interval [0, 1]. A return value close to 1 indicate strong relation between a and c, and vise-versa. When working as an inference engine, (4) evaluates the strength of relations between all (a, c) pairs and let the defuzzification module interprets the results according to the requirements of the application.

2.2 Defuzzification of BK Subproduct Based Fuzzy Logic Systems

Defuzzification is a process that converts fuzzy outputs to crisp values. Some popular defuzzification algorithm for CRI-based inference engines include Center of Area (COA), Mean of Maximum (MOM) and etc [27].

In some other systems, a predefined option with highest membership grade is selected as the defuzzified value. Some BK subproduct based fuzzy logic systems produce outputs with methods that are quite similar to this [17–19].

In these systems, for a particular a', a set of possible options, c_i were evaluated with inference engine derived from BK subproduct. For each option, a membership grade that represents the composite relation $a'(R \lhd S)c_i$ is computed. Finally, an option with highest membership grade is chosen as the ultimate output of the system.

Some other inference systems prefer indiscriminative reasoning where non mutually exclusive results are accepted [14, 24, 28]. In these systems, BK subproduct has exhibited its advantage as inference engine. With different fuzzy implication operators (Table 1), different outputs may be produced even with the same set of inputs. To increase the credibility of inferences, these systems adopted checklist paradigm [29] to identify a pair of implication operators - Kleene-Dienes and Łukasiewicz implication operators as the two end points of implications. Hence, an interval is obtained for each inference, instead of a single point value. An inference is considered as accepted if and only if the whole result interval is sitting in the acceptance band [β,1], where β is the predefined acceptance threshold value. On the other hand, an inference is rejected if the whole result interval is sitting in the rejection band [0,α], where α is the predefined rejection threshold value. This model of defuzzification is called IVI [24].

For example, in [13], BK subproduct was used to identify the relation between patients and diseases. Kleene-Dienes and Łukasiewicz fuzzy implication operators are used to generate an interval for each inference. A patient is suspected as suffering for a particular disease if the result interval is sitting in the acceptance band, which is [0.8, 1.0]. If the interval is sitting in the interval [0, 0.3], it is considered as rejected. The remaining interval, (0.3,0.8) is considered as gray area where the inference results are not accepted and rejected.

3.0 LIMITATION OF IVI

The adoption of IVI is to increase the accuracy [14] of inferences. An interval has to be considered, instead of a point value to confirm that an inference is accepted or rejected. However, inattention to a property of fuzzy implication operators in the design lead to the low effectiveness in the systems that implementing this defuzzification algorithm. We explain the details in the following.

From the definition of Kleene-Dienes and Łukasiewicz fuzzy implication operators in Table 1, one can easily prove that:

$$\forall r,s \in [0,1] \qquad (r \rightarrow_{KD} s) \leq (r \rightarrow_{Ł} s) \qquad (5)$$

Readers can also refer to Table 2 for example of calculations of $\rightarrow_{KD}$ and $\rightarrow_{Ł}$. With (5), it is clear that for an inference to obtain an interval in the acceptance band, computation using $\rightarrow_{KD}$ is sufficient. This is because whenever a computation with $\rightarrow_{KD}$ is in the acceptance band, the computation with $\rightarrow_{Ł}$ which yields a result greater or equal to $\rightarrow_{KD}$ is always in the acceptance band too.

$$\forall r,s \in [0,1] \qquad r \rightarrow_{KD} s \geq \beta \;\Rightarrow\; r \rightarrow_{Ł} s \geq \beta \qquad (6)$$

On the other hand, if the computation using $\rightarrow_{KD}$ is not in the acceptance band, the result of computation using $\rightarrow_{Ł}$ is not relevant anymore because the interval is not entirely sitting in the acceptance band and the inference is not going to be accepted. To check whether an inference is rejected, a similar argument hold, $\rightarrow_{Ł}$ is the only needed fuzzy implication for this purpose.

Furthermore, this defuzzification algorithm may run into the trap that facing by crisp systems - details are oversimplified by applying threshold values. For example, if β is 0.80, then the result interval [0.80, 0.95] will be accepted but [0.79, 0.95] will be rejected, although the difference between these intervals are very small.

With the problems that we find in this section, we can conclude that [14, 24, 28] which perform IVI do not really benefit from the intervals that obtained from the computations, although [14] claimed that intervals have better accuracy over point values. Hence, a solution that based on the ratio of result intervals in the acceptance or rejection bands is proposed in the following section.

Table 2. Values generated by the Kleene-Dienes ($r \rightarrow_{KD} s$) and Łukasiewicz implication operators ($r \rightarrow_{Ł} s$)

(a) $r \rightarrow_{KD} s$

s \ r	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.1	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	1.0
0.2	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	1.0
0.3	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.9	1.0
0.4	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.8	0.9	1.0
0.5	0.5	0.5	0.5	0.5	0.5	0.6	0.7	0.8	0.9	1.0
0.6	0.4	0.4	0.4	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.7	0.3	0.3	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.8	0.2	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.9	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
1.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0

(b) $r \rightarrow_{Ł} s$

s \ r	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
0.2	0.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
0.3	0.8	0.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
0.4	0.7	0.8	0.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0
0.5	0.6	0.7	0.8	0.9	1.0	1.0	1.0	1.0	1.0	1.0
0.6	0.5	0.6	0.7	0.8	0.9	1.0	1.0	1.0	1.0	1.0
0.7	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.0	1.0	1.0
0.8	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.0	1.0
0.9	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.0
1.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0

4.0 IMPROVED IVI WITH LIKELIHOOD RATIO

Bandler and Kohout [11] revised the theory of fuzzy subsets proposed by Zadeh [30], which proposed that a fuzzy set A is a subset of another fuzzy set B if and only if:

$$\forall x \in U, \qquad A(x) \leq B(x) \tag{7}$$

With this theory of fuzzy subsets, a fuzzy set is either utterly a subset of the other fuzzy set or the other way round. Apparently, this theory of fuzzy subset is "an unconscious step backward to the realm of dichotomy" [11]. To rectify this problem, a subsethood theory that based on the implication operators was proposed (discussed in Section 2.1.).

Based on the same justification, we propose a new IVI scheme that can rectify the limitations that addressed in Section 3. This new scheme not only provides information about whether an inference result is in the predefined band (either acceptant or rejection), but also measures the likelihood of an inference based on a threshold value.

Basically, this improved method measures the likelihood ratio of an inference result. It evaluates the ratio of an output interval in the acceptant and rejection bands. If the whole output interval is in the acceptance (or rejection) band, it is reasonable to assume that the inference is completely accepted (or rejected) and the likelihood is 1.0. Otherwise, the ratio of the interval in the corresponding evaluation band gives the likelihood of this inference.

Assume that the result of an inference is an interval $x = [k,l]$, where k and l are corresponding to the results of Kleene-Dienes and Łukasiewicz fuzzy implication operators respectively. If $l > \beta$, it implies that at least a portion of $[k, l]$ is in the acceptance band (Fig. 3). Therefore, we can measure the likelihood of this inference with result x being accepted at threshold value β with a piecewise function, Θ_x^β:

$$\Theta_x^\beta = \begin{cases} 1, & \text{if } k \geq \beta \\ \dfrac{l-\beta}{l-k}, & \text{if } k < \beta < l \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

and $\Theta \in [0,1]$. Similarly, the measurement of the likelihood of an inference with interval x in rejection band at threshold value α is given by Ψ:

$$\Psi_x^\alpha = \begin{cases} 1, & \text{if } l \leq \alpha \\ \frac{\alpha - k}{l - k}, & if\ k < \alpha < l \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

and $\Psi \in [0,1]$.

The magnitudes of Θ and Ψ show the likelihoods of inferences in acceptance and rejection bands respectively. Instead of utterly accepting or rejecting an inference, this new scheme provides better tolerant towards uncertainty in inputs and the choice of acceptance and rejection threshold values. In term of effectiveness, the computations based on both implication operators $\rightarrow_{KD}$ and $\rightarrow_{Ł}$ are used as compliment to each other in producing information-rich results. Compare to the classical IVI which require two computations but only one is used in generating results, obviously the proposed likelihood ratio is more effective with no additional computations required.

5.0 EXPERIMENT AND RESULT

5.1 Dataset

To verify the performance of the proposed defuzzification algorithm, an experiment involving a publicly available Wisconsin Breast Cancer (Diagnostic) dataset [31] is carried out. The Wisconsin Breast Cancer (Diagnostic) dataset has a number of 569 samples, which are divided into 2 classes (357 benign and 212 malignant). Besides the patient IDs and classes, another 30 features are provided as hints for reasoning.

As discussed in [16], BK subproduct based inference engines demonstrate a unique advantage in the training process, i.e. a small amount of data is sufficient to train the inference systems. Therefore, only 10% of the samples in Wisconsin Breast Cancer (Diagnostic) dataset (36 benign and 21 malignant) are used to train the system and the remaining 90% (321 benign and 191 malignant) are tested in this experiment.

5.2 Inference System

BK subproduct is implemented as a soft classifier in this experiment. To be specific, equation (10) is used as inference structure. The Kleene-Dienes and Łukasiewicz fuzzy implication operators (refer Table (1)) are used to generate an interval for each set of input. For the defuzzification, these outputs, or the intervals are sent to (8) for the computation of likelihood.

$$R \lhd S(a,c) = \frac{1}{N}\sum_{b=1}^{N}(R_{ab} \rightarrow S_{bc}) \tag{10}$$

It is worth to be reminded that the objective of this experiment is not to test the performance of BK subproduct, but to demonstrate the usefulness of the proposed defuzzification algorithm in interpreting the inference results compared to conventional IVI. Since our focus is on inference system with non-discriminative reasoning where non-mutual exclusive results are accepted, classifying a sample into 2 species or more are allowed. With the proposed likelihood ratio, we expect to see the inference systems to present the confidence level of relating a sample to the typical attributes of a species in term of likelihood. In case of strict-cut type classification is needed, the adoption of likelihood ratio also shows better performance compared to the classical IVI.

Method proposed in [32] is used to generate membership functions corresponding to the features shown by each attribute in both datasets. In short, the method derives clusters in a class using fuzzy equivalence relations and measure the membership degrees in these clusters. First, for a feature, the data in training set is sorted in ascending order.With the data arranged in sequence, fuzzy compatibility relation between pairs of training data is retrieved. Next, max-min transitive closure of the fuzzy compatibility relation is retrieved to obtain the fuzzy equivalence relation. This fuzzy equivalence relation can be considered as a measure of the similarity between a pair of data. With appropriate α-cut values, partition the whole range of data into multiple partitions (The choice of the value of α is data dependent). Now, each partition constructed is a fuzzy subset for the feature. With the maximum, minimum and mean of the data in a subset, a triangular membership function can be constructed. Then, this membership function is used to retrieve membership degree of elements. The above steps are repeated for all the features to construct all the membership functions required.

To test the system, all the remaining samples in the dataset are used. The measurements of the attributes are converted to membership grades with membership functions generated above. This forms a matrix that represents $R(a,b)$.

5.3 Results

Before the results of the improved defuzzification algorithm is discussed, the results of classical IVI with $\beta = 0.90$ is presented in Table 3 as reference. Out of the 357 benign samples, 286 are inferred as benign. Whereas the ratio of malignant samples being diagnosed correctly as malignant is 113 over 212.

Table 3. Number of samples accepted with classical IVI with $\beta = 0.9$

Class / Inferred as	Benign	Malignant
Benign	286	2
Malignant	54	113

For the result with the proposed likelihood ratio, the classical evaluation tools such as measuring of "true positive" and "true negative" is not applicable here. This is because the propose method is not a strong classifier by nature. Therefore, the arithmetic means of the likelihood to accept with threshold value 0.9 ($\Theta^{0.9}$) is measured. The result is presented in Table 4 which we can see on average the likelihood of a patient being inferred correctly is high. On average, the likelihood when a benign case being inferred as benign is 0.98, whereas the likelihood of malignant inferred as malignant is 0.84. For some benign cases, the system leaves room for medical experts to further explore the cases with an average of likelihood 0.61. However, the average likelihood of a malignant case inferred as benign is low as expected (0.15).

Table 4. Means of the likelihood ratios of all the inferences with $\beta = 0.90$

Class / Inferred as	Benign	Malignant
Benign	0.98	0.15
Malignant	0.61	0.84

In some cases, if a strong classification is required, one can always convert the soft classifier to a strong classifier if a level of likelihood ratio can be determined. For example, if all the inferences with likelihood ratio $\Theta \geq 0.5$ are considered as accepted, the classification is strong and we obtain Table 5. A likelihood ratio $\Theta = 1.0$ will bring the result in Table 3. Furthermore, if the highest likelihood ratios among classes represent the most acceptable inference results, then Table 6 presents the total correct inferences by the inference engine with $\beta = 0.9$.

Table 5. Number of inferences accepted with $\beta = 0.9$ and $\Theta \geq 0.5$

Class / Inferred as	Benign	Malignant
Benign	319	45
Malignant	130	172

Table 6. Number of correct inference based on highest likelihood ratio. The figures in brackets are the number of samples that showing same likelihood ratio for both benign and malignant inferences.

Class	Correct inferences
Benign	318 (2)
Malignant	147 (36)

The acceptance threshold, β controls the degree of an inference to be accepted or not for both the classical IVI and the proposed likelihood ratio. Therefore, it is expected that for both the methods, consequences of increasing or decreasing the value of β are tighten or loosen of an inference to be considered as accepted. For the purpose of comparison, experiment with different values of β, which are not optimal is also conducted. The Table 7 shows the results of classical IVI, with $\beta = 0.85$ and $\beta = 0.95$. Comparing Table 7 with Table 3 shows that the results are greatly affected by the value of the acceptance threshold, β. Lower the acceptance threshold allows more inferences being accepted, but increasing the acceptance threshold causes almost all of the inferences not being accepted.

Table 7. Number of samples accepted with classical IVI with $\beta = 0.85$ or $\beta = 0.95$.

(a) $\beta = 0.85$

Class / Inferred as	Benign	Malignant
Benign	319	166
Malignant	176	184

(b) $\beta = 0.95$

Class / Inferred as	Benign	Malignant
Benign	0	0
Malignant	0	4

On the other hand, inference results of the proposed likelihood ratio method is presented in Table 8. The numbers indicate the accepted inferences with corresponding β and the likelihood ratio $\Theta \geq 0.5$. As expected, the likelihood ratio method also does not work well with non-optimal acceptance threshold values. However, if Table 8 is being compared with Table 7, we find the relative advantage of newly proposed likelihood ratio method with high threshold values. At a high threshold values, not almost all inferences are rejected (IVI in Table 7(b)), but still a high ratio of benign/malignant cases are inferred as benign/malignant (Table 8(b)).

Table 8. Number of inferences accepted with $\beta = 0.85$ and $\beta = 0.95$, and $\Theta \geq 0.5$

(a) $\beta = 0.85$

Class / Inferred as	Benign	Malignant
Benign	321	240
Malignant	190	189

(b) $\beta = 0.95$

Class / Inferred as	Benign	Malignant
Benign	23	0
Malignant	7	52

From the results, it is obvious that the likelihood ratio method has advantages compare to the classical IVI method, for both optimal and non-optimal but higher acceptance threshold values cases. An optimized acceptance threshold value is good in achieving better inference results for both methods, however, optimizing the acceptance threshold values is not within the scope of this research.

6.0 CONCLUSION

BK subproduct has its advantages in developing fuzzy inference engines. However, IVI, the classical defuzzification algorithm is low in terms of effectiveness and the results produced for medical decision support systems are not in the preferred form. As a result, the likelihood ratio is introduced in this paper to improve the defuzzification process. With this improvement, an inference is not accepted (rejected) utterly, but a degree of likelihood of acceptance (rejection) is provided. This likelihood is a measure of the ratio of an output interval in an acceptance band.

Defuzzification is an important process in fuzzy logic systems. With a good defuzzification algorithm, information generated from inference engine can be interpreted correctly and more meaningful results can be obtained. With this likelihood ratio, performance of the BK subproduct based inference systems are expected to improve. This conclusion is verified with the experiments using breast cancer dataset in this paper.

To extend the capability of BK subproduct based inference systems, one of the topics that worth to investigate in the future is optimization of the acceptance threshold. A minor change of acceptance threshold affects the results of inference systems, regardless of classical IVI or proposed likelihood ratio is used as the defuzzification algorithm. An optimized acceptance threshold can refine the classification resolution of the inference system. To address this problem, machine learning techniques such as regression analysis can be adopted.

ACKNOWLEDGEMENT

This research was supported by the BKP Research Grant of University Malaya (Project No. BK071-2016).

REFERENCES

[1] S. M. Ting, T. Ling, K. Phang, L. C. Sun, and S. Gan, "Enhanced software development effort and cost estimation using fuzzy logic model," *Malaysian Journal of Computer Science*, vol. 20, no. 2, 2007, pp. 199–207.

[2] W. L. Du, D. Ho, and L. F. Capretz, "Improving software effort estimation using neuro-fuzzy model with seer-sem," *arXiv preprint arXiv*:1507.06917, 2015.

[3] H. Liu, Z. Ju, X. Ji, C. S. Chan, and M. Khoury, "Fuzzy qualitative human motion analysis," in *Human Motion Sensing and Recognition*. Springer, 2017, pp. 67–93.

[4] V. John, S. Mita, Z. Liu, and B. Qi, "Pedestrian detection in thermal images using adaptive fuzzy c-means clustering and convolutional neural networks," in *2015 14th IAPR International Conference on Machine Vision Applications (MVA)*, May 2015, pp. 246–249.

[5] S.-T. Sheu, C.-L. Lin, and J. Chen, "Reducing signaling cost of location management in cellular network using fuzzy logic controller," in *Wireless Communications, Networking and Mobile Computing (WiCOM 2015), 11th International Conference on*. IET, 2015, pp. 1–8..

[6] M. Collotta, G. Scat, S. Tirrito, R. Ferrero, and M. Rebaudengo, "A parallel fuzzy scheme to improve power consumption management in wireless sensor networks," in *Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA)*, Sept 2014, pp. 1–4.

[7] Y.-H. O. Chang and B. M. Ayyub, "Fuzzy regression methods a comparative assessment," *Fuzzy Sets and Systems*, vol. 119, no. 2, 2001, pp. 187 – 203.

[8] A. Lahsasna and W. C. Seng, "An improved genetic-fuzzy system for classification and data analysis," *Expert Systems with Applications*, vol. 83, 2017, pp. 49 – 62.

[9] A. T. Azar and A. E. Hassanien, "Dimensionality reduction of medical big data using neural-fuzzy classifier," *Soft Computing*, vol. 19, no. 4, 2015, pp. 1115–1127.

[10] B. Dennis and S. Muthukrishnan, "AGFS: Adaptive genetic fuzzy system for medical data classification," *Applied Soft Computing*, vol. 25, 2014, pp. 242–252..

[11] L. J. Kohout and W. Bandler, "Fuzzy Power Sets And Fuzzy Implication Operators," *Fuzzy Sets and Systems*, vol. 4, 1980, pp. 13–30.

[12] L. J. Kohout and W. Bandler, "Relational-Product Architectures for Information Processing," *Information Sciences*, vol. 37, 1985, pp. 25–37.

[13] K. M. Yew and L. J. Kohout, "Interval-based reasoning in medical diagnosis," in *Proceedings Intelligent Information Systems*. IIS'97. IEEE Comput. Soc, 1997, pp. 32–36.

[14] K. M. Yew and L. Kohout, "Classes of Iinterval-valued Fuzzy Relational Inference Templates," *Malaysian Journal of Computer Sciences*, vol. 9, no. 1, 1996, pp. 42–48.

[15] C. K. Lim and C. S. Chan, "Logical Connectives and Operativeness of BK Sub-triangle Product in Fuzzy Inferencing," *International Journal of Fuzzy Systems*, vol. 13, no. 4, 2011, pp. 237–245.

[16] C. K. Lim and C. S. Chan, "A weighted inference engine based on interval-valued fuzzy relational theory," *Expert Systems with Applications*, vol. 42, no. 7, 2015, pp. 3410–3419.

[17] Y.-i. Lee and Y.-G. Kim, "Comparison of fuzzy implication operators by means of fuzzy relational products used for intelligent local path-planning of AUVs," *Soft Computing*, vol. 13, no. 6, May 2008, pp. 535–549.

[18] L.-d. Bui and Y.-g. Kim, "An Obstacle-avoidance Technique for Autonomous Underwater Vehicles Based On BK-products of Fuzzy Relation," *Fuzzy Sets and Systems*, vol. 157, 2006, pp. 560 – 577.

[19] R. Groenemans, V. E. Ranst, and E. Kerre, "Fuzzy relational calculus in land evaluation," *Geoderma*, vol. 77, 1997, pp. 283–298.

[20] E. Vats, C. K. Lim, and C. S. Chan, "An improved BK sub-triangle product approach for scene classification," *Journal of Intelligent and Fuzzy Systems*, vol. 29, no. 5, 2015, pp. 1923–1931.

[21] E. Vats, C. K. Lim, and C. S. Chan, "Early Human Actions Detection using BK Sub-triangle Product," in *IEEE Conference on Fuzzy Systems*, 2015, pp. 1–8.

[22] M. Stepnicka and B. Jayaram, "On the Suitability of the Bandler Kohout Subproduct as an Inference Mechanism," *Fuzzy Systems, IEEE Transactions on*, vol. 18, no. 2, 2010, pp. 285–298.

[23] L. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," *Systems, Man and Cybernetics, IEEE Transactions on*, no. 1, 1973, pp. 28–44.

[24] L. J. Kohout and W. Bandler, "How the checklist paradigm elucidates the semantics of fuzzy inference," *IEEE International Conference on Fuzzy Systems*, 1992, pp. 571–578.

[25] V.L.Patel, D.R.Kaufman, andJ.F.Arocha, "Emerging paradigms of cognition in medical decision-making," *Journal of biomedical informatics*, vol. 35, no. 1, 2002, pp. 52–75.

[26] M. A. Musen, B. Middleton, and R. A. Greenes, "Clinical decision-support systems," in *Biomedical informatics*. Springer, 2014, pp. 643–674.

[27] G. J. Klir and B. Yuan, *Fuzzy Sets And Fuzzy Logic: Theory And Applications*. New Jersey: Prentice Hall, 1995.

[28] L. J. Kohout, I. Stabile, H. Kalantar, and F. San-Andres, "Parallel interval-based reasoning in medical knowledge-based system Clinaid," *Reliable Computing*, vol. 1, no. 2, 1995, pp. 109–140.

[29] L. J. Kohout and W. Bandler, "Semantics of implication operators and fuzzy relational products," *International Journal of Man-Machine Studies*, vol. 12, no. 1, Jan. 1980, pp. 89–116.

[30] L. A. Zadeh, "Fuzzy Sets," *Information and Control*, vol. 8, no. 1, Jan. 1965, p. 338.

[31] M. Lichman, "UCI machine learning repository," 2013. [Online]. Available: http://archive.ics.uci.edu/ml

[32] T. P. Wu and S. M. Chen, "A new method for constructing membership functions and fuzzy rules from training examples." *IEEE transactions on systems, man, and cybernetics. Part B*, vol. 29, no. 1, Jan. 1999, pp. 25–40.