A Comparison: Fuzzy Signatures and Choquet Integral

B.S.U. Mendis and T.D. Gedeon

***Abstract*— Fuzzy Signatures are hierarchical multi aggregative descriptors of objects. They have reduced computational complexity compared to formal fuzzy rule based systems. Weighted Relevance Aggregation enhances the performance of hierarchical Fuzzy Signatures. Thus, they are very robust and flexible under perturbed input data. On the other hand the Choquet Integral, which is based on fuzzy measures, is a powerful aggregation tool in multi-criteria decision making. We compared Fuzzy Signatures and the Choquet Integral as practical applications for hierarchical and non-hierarchical data aggregation/organization methods.**

***Index Terms*— Fuzzy Signatures, Aggregation Operators, generalised Weighted Relevance Aggregation Operator, Fuzzy Integrals, Choquet Integral, Fuzzy Measure, Multi-criteria Decision Making.**

I. Introduction

FUZZY Signature concept is a generalization of the vector valued fuzzy sets concept in [16]. Fuzzy signatures can model hierarchically structured problems with the help of hierarchically structured vector valued fuzzy sets and a set of not-necessarily homogenous and hierarchically organised aggregation functions. The set of aggregation functions map, from lower branches to the higher branches, the different universes of discourse of the hierarchical fuzzy signature structure. We argue that these properties help fuzzy signatures to model problems similar to the nature of human hierarchical approaches to problem solving. An important advantage of the fuzzy signature concept is that it can be used to compare degrees of similarity or dissimilarity of two slightly different objects, which have the same fuzzy signature skeleton. Additionally, fuzzy signatures are capable of dealing with missing input data. Thus, medical/economic diagnoses, web and document information retrieval, data mining are the obvious applications of fuzzy signatures.

We investigate the different ways of aggregating hierarchical Fuzzy Signatures in [19]. Further, in [24] we enhanced the inference in fuzzy signatures, by introducing the Weighted Relevance Aggregation (WRA) method. The concept behind the Weighted Relevance aggregation method is that the weights in each branch of the fuzzy signature are the observations of the relevance of that branch to its higher level branches in the hierarchical fuzzy signature structure. Thus, this method introduces extra knowledge into the fuzzy signature structure to classify vague data. In addition, it enhances the adaptability of hierarchical fuzzy signatures to different problem domains. In [23] we showed the methodology of learning Weighted Relevancies from real world data. The results of experiments in [23] showed that weight learning was successful and effective.

Later, we further generalised these Weighted Relevancies and aggregation functions in hierarchical fuzzy signatures, into one operator called Weighted Relevance Aggregation Operator (WRAO) [21]. WRAO allows a user to learn both aggregation function and weighted relevance at the same time for each node in the hierarchical fuzzy signatures structure. Thus, WRAO simplifies the learning of hierarchical fuzzy signature models from data. In [22] we have shown a successful way to extract WRAO for hierarchical Fuzzy Signatures based on the Levenberg-Marquardt (LM) optimization method [18], [17]. Experiments in [22] showed that LM method can learn both aggregations and weighted relevancies for hierarchical fuzzy signatures.

In some situations it is necessary to reduce or aggregate information to become compatible with information obtained from another source, where some detail variables are missing or have simply been removed from the input data. In such situations, human experts can still make decisions based on existing knowledge. For example, medical practitioners can make decisions based on the available data for a patient in urgent situations. In [20], we investigated the ability of fuzzy signature structures to cope with substantially missing information, as simulated by an experiment in which some major branches were removed from the structure. We also investigated the removal of significant number of individual data items from inputs. The results demonstrated that fuzzy signatures are highly robust under both of these conditions.

Sugeno [29], [25] proposed the concept of Fuzzy Measures and Integral in which it finds a kind of disordered mean in the desecrate case. In the Multi-Criteria Decision Making (MCDM) literature [9], it has been shown that weighted average, medians, a large part of generalised means [7], weighted maximum and minimum [6], and Ordered Weighted Averaging (OWA) operators [33] are special cases of Fuzzy Integrals. Therefore Fuzzy Integrals are powerful tools that cover wide ranges of aggregation criteria and distinguished from others in that it is able to represent the interactions between criteria ranging from redundancy to synergy. Thus, it is worth comparing a Fuzzy Integral with our novel concept of hierarchical Fuzzy signatures, which organise and aggregate data hierarchically to describe real world data better. In this, paper we use the Choquet Integral [26] as the Fuzzy Integral to be compared with Fuzzy Signatures.

The rest of the paper is organised as follows. In section II, we discuss the concept of hierarchical Fuzzy Signatures and Weighted Relevance Aggregation (WRA). Section III reviews the Fuzzy Measures and Choquet Integral and it also discusses the learning of Fuzzy Measure Parameters. Finally,

B.S.U. Mendis and T.D. Gedeon are with the Department of Computer Science, The Collage of Engineering and Computer Science, The Australian National University, Australia (email: {sumudu, tom}@cs.anu.edu.au).

978-1-4244-1819-0/08/$25.00©2008 IEEE

Authorized licensed use limited to: Australian National University. Downloaded on January 25,2021 at 12:04:26 UTC from IEEE Xplore. Restrictions apply.

in section IV, we compare the Fuzzy Signatures and Choquet Integral against two real world problems.

II. Fuzzy Signature Concept

Fuzzy signatures can describe, compare and classify objects with complex structures and interdependent features. The hierarchical organisation of fuzzy signatures expresses the structural complexity of a problem. The local preference relations among the hierarchies and sub-branches of a fuzzy signature can be used to approximate the global preference relation of a decision problem.

1) Vector Valued Fuzzy Sets: The fuzzy signature concept is a generalization of the Vector Valued Fuzzy Sets (VVFS) concept. The early work of Kóczy [15] introduced the Vector Valued Fuzzy Sets concept. The VVFS is a special form of an *L*-fuzzy set, and can be denoted in the following form:

$$A : X \to [0,1]^n . \tag{1}$$

It is obvious that $L = [0,1]^n$ is in (1) and thus VVFS is *L*-fuzzy. The qualitative meaning of an object is represented by the quantities of the VVFS. Applications of VVFS can be found in [13], [28].

The notation of the vector valued fuzzy set A is written as $\overline{A} = (x, q_{\overline{A}})$ and the membership function $q_{\overline{A}}$ can be defined as, $q_{\overline{A}} : X \to [0,1]^n$, where $x \in X$.

2) Hierarchical Fuzzy Signatures: Fuzzy signatures are fuzzy descriptors of real world objects. They represent objects with the help of a sets of quantities that are arranged in a hierarchical structure expressing interconnectedness and set of non-homogeneous qualitative measures, which are the interdependencies among the quantities of each set, to aggregate these hierarchies. Thus, fuzzy signatures are capable of handling problems that are complex and inherently hierarchical.

Additionally, the fuzzy signature concept is a good solution to the rule explosion problem in fuzzy logic, as fuzzy signatures are hierarchically structured and inherently sparse. In this section, we discuss the concept of fuzzy signatures as a practical approach that organises and aggregates data hierarchically. Now, we recall the fuzzy signature concept introduced in [16].

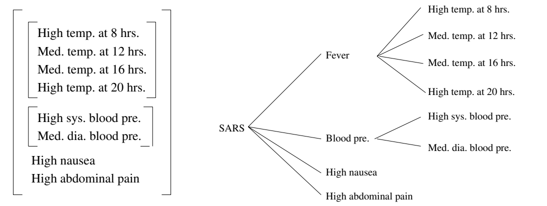
Definition 1: Fuzzy Signature is a VVFS, where each vector component is another VVFS (branch) or a atomic value (leaf), and denoted by,

$$A : X \to [a_i]_{i=1}^{k} \left(\equiv \prod_{i=1}^{k} a_i \right) . \tag{2}$$

$$where\ a_i = \begin{cases} [a_{ij}]_{j=1}^{k_i} & ; if\ branch \\ [0,1] & ; if\ leaf \end{cases}$$

and Π describes the Cartesian product.

The figure 1(a) shows an example fuzzy signature [32]. This fuzzy signature describes an individual SARS patient, which is a data point among many SARS data collected in the year 2003 [31], [27], [14]. The figure 1(b) shows the hierarchical view of the fuzzy signature shown in figure 1(a).



(a) As a Vector (b) As a Tree

Fig. 1: Example Fuzzy Signature

Figure 2 shows an example for a simple aggregation of a fuzzy signature using *max* and *min* as the aggregation functions.

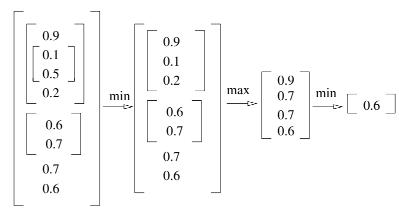


Fig. 2: Aggregation of fuzzy signatures

3) Weighted Relevance Aggregation (WRA): The Weighted Relevance Aggregation provides an additional meaning to the fuzzy signature structure by introducing the weighted relevance of each branch to its higher branches of the fuzzy signature structure. That is, the weighted relevance reflects the idea that some branches provide higher values to the next level of the fuzzy signature structure. Some other branches in the same parent branch provide relatively lower values to the next level of the fuzzy signature structure. In this way WRA enhances the accuracy of the final results of the Fuzzy Signature. In [24], we discussed a method of learning weights in WRA automatically. In [23], we have shown the successfulness of the weights extraction method in [24].

We further generalise the weights and the aggregation into one operator called Weighted Relevance Aggregation Operator (WRAO) [21]. This subsection briefly describes the generalised Weighted Relevance Aggregation (WRAO) operator [21] for fuzzy signatures. In [22], we showed that WRAO enhances the accuracy of the results of fuzzy signatures, by allowing better adaptation to the meaning of the decision making process. Further, WRAO helps to reduce the number of individual fuzzy signatures needed for the decision making process, by absorbing more patterns into these recognized by one Fuzzy Signature.

Now, we recall the definition of WRAO in [21]. All the notation in definition 2 refer to the arbitrary fuzzy signature

Authorized licensed use limited to: Australian National University. Downloaded on January 25,2021 at 12:04:26 UTC from IEEE Xplore. Restrictions apply.

in figure 3.

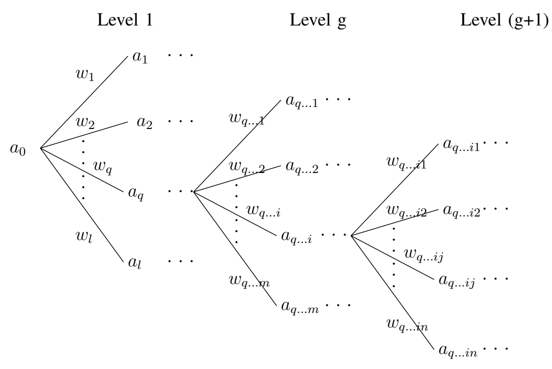


Fig. 3: An Arbitrary Fuzzy Signature

Definition 2: The generalised Weighted Relevance Aggregation Operator (WRAO) of an arbitrary branch $a_{q...i}$ with n sub-branches, $a_{q...i1}, a_{q...i2}, \ldots, a_{q...in} \in [0,1]$, and weighted relevancies, $w_{q...i1}, w_{q...i2}, \ldots, w_{q...in} \in [0,1]$, for an arbitrary fuzzy signature, figure 3, is a function g : $[0,1]^{2n} \to [0,1]$ such that,

$$a_{q...i} = \left[\frac{1}{n} \sum_{j=1}^{n} (a_{q...ij} \bullet w_{q...ij})^{p_{q...i}} \right]^{\frac{1}{p_{q...i}}} \tag{3}$$

The WRAO must satisfy the following three properties,

(i) $w_{q...ij} \in [0,1]$

(ii) $\bigvee_{j=1}^{n} w_{q...ij} \leq 1$

(iii) $p_{q...i} \neq 0$

In [21], we prove the following two properties for WRAO.

Theorem 1: Let $a_{q...i}$ be an arbitrary branch with n sub-branches, $s_{q...i1}, s_{q...i2}, \ldots, s_{q...in}$, and weighted relevancies, $w_{q...i1}, w_{q...i2}, \ldots, w_{q...in}$, for an arbitrary fuzzy signature (figure 3). Then WRAO in definition (2) holds the following properties.

(i) Partially Idempotent w.r.t $s_{q...ij}$, when all $w_{q...ij}$ are fixed and vice versa,

(ii) Commutative, and

(iii) Partially Monotonic w.r.t $s_{q...ij}$ when all $w_{q...ij}$ are fixed and vice versa.

The partial idempotency and monotonicity is adequate to satisfy the requirement to be an aggregation function [5] as weights, $w_{q...i1}, w_{q...i2}, \ldots, w_{q...in}$, in WRAO are fixed for any instance of a fuzzy signature in the decision making phase, and both weights and aggregation operators vary simultaneously only in the learning phase.

Theorem 2: The WRAO in definition (2) holds the following characteristics.

(a) $p_{q...i} \to 0$ then WRAO $\to$ geometric mean

(b) $$\lim_{p_{q...i} \to +\infty} g(s_{q...i1}, \ldots s_{q...in}; w_{q...i1}, \ldots, w_{q...in}) = max(s_{q...i1}w_{q...i1}, \ldots s_{q...in}w_{q...in})$$

(c) $$\lim_{p_{q...i} \to -\infty} g(s_{q...i1}, \ldots s_{q...in}; w_{q...i1}, \ldots, w_{q...in}) = min(s_{q...i1}w_{q...i1}, \ldots s_{q...in}w_{q...in})$$

(d) $p = 1$ then WRAO$\to$ arithmetic mean

(e) $p = -1$ then WRAO$\to$ harmonic mean

4) Levenberg-Marquardt Learning of WRAO for Fuzzy Signatures: In this subsection we explain the method of learning WRAO from real world data briefly, with more detailed explanations to be found in [21]. First, to avoid the first 2 constraints on the weighted relevance factor $w_{q...ij}$ in definition 2, we replace it by the following sigmoid function:

$$w_{q...ij} = \frac{1}{1 + e^{-\lambda_{q...ij}}} \tag{4}$$

where $\lambda_{q...ij} \in \Re$. Now, the equation (3) can be modified as follows,

$$a_{q...i} = \left[\frac{1}{n} \sum_{j=1}^{n} \left(s_{q...ij} \bullet \left[\frac{1}{1 + e^{-\lambda_{q...ij}}} \right] \right)^{p_{q...i}} \right]^{\frac{1}{p_{q...i}}} \tag{5}$$

Now, the constrained optimisation problem become an unconstrained optimisation problem. The $p_{q...i}$ and $\lambda_{q...ij}$ are called the aggregation factor of $q \ldots i$th branch and weighted relevance factor of $q \ldots ij$th sub-branch of the fuzzy signature in figure 3, respectively. This form of WRAO (5) can be readily used for gradient based learning.

The parameters we need to learn are the aggregation factor $p_{q...i}$ and weighted relevance factors $\lambda_{q...ij}$ for each WRAO at each node of the fuzzy signature structure in figure 3. First we can obtain the partial derivatives of the equation (5) w.r.t. $p_{q...i}$,

$$\frac{\partial a_{q...i}}{\partial p_{q...i}} = \left[\frac{a_{q...i}^{1-p_{q...i}}}{n p_{q...i}^2} \right] \left\{ \sum_{j=1}^{n} t \ln(t) - n t' \ln\left(t'\right) \right\} \tag{6}$$

where $t = (a_{q...ij} w_{q...ij})^{p_{q...i}}$ and $t' = a_{q...i}^{p_{q...i}}$. Similarly, we obtain the patrial derivatives of the equation (5) w.r.t. $\lambda_{q...ik}$

$$\frac{\partial a_{q...i}}{\partial \lambda_{q...ik}} = \left[\frac{1}{n} \sum_{j=1}^{n} (s_{q...ij} \bullet w_{q...ij})^{p_{q...i}} \right]^{\frac{1}{p_{q...i}} - 1} \bullet T \tag{7}$$

where $w_{q...ij} = \frac{1}{1+e^{-\lambda_{q...ij}}}$ and $T = \left\{ \frac{d([s_{q...ik} \bullet w_{q...ik}]^{p_{q...i}})}{d\lambda_{q...ik}} \right\}$.

We used, the Levenberg-Marquardt (LM) method [17], [18] for learning of WRAO parameters. The LM algorithm is a widely used advanced optimization algorithm that outperforms simple gradient descent and other gradient methods when applied in a wide variety of problems. The LM algorithm is a pseudo-second order, Sum of Square Error (SSE) based optimization method, in which the Hessian matrix is estimated using the gradients [17], [18]. The two

Authorized licensed use limited to: Australian National University. Downloaded on January 25,2021 at 12:04:26 UTC from IEEE Xplore. Restrictions apply.

equations, (6) and (7) above, together with the chain rule for partial derivatives have been used to calculate the Jacobian, which is then used to approximate the Hessian matrix for LM learning. The detailed discussion of the method of using LM for learning WRAO can be found in [22].

III. Fuzzy Measures and Choquet Integral

In this section we discuss the concepts fuzzy measure and Choquet Integral for Multi-Criteria Decision Making (MCDM). MCDM is a well-known paradigm for intelligent decision making and is more widely used compared to the paradigm Decision Making Under Uncertainty.

A. Fuzzy Measure

The main characteristic of classical measures theory is additivity. Many engineering applications were successfully designed with this property, but when it comes to soft computing applications the additivity property can be too rigid. The fuzzy measure is a generalisation of the additive measures as it replaces additivity by the weaker condition of monotonicity [29].

Definition 3: A fuzzy measure on a discrete set $N = \{1, \ldots, n\}$ is a set function $v : 2^N \rightarrow [0, 1]$ that satisfies the following conditions:

(i) Boundary: $v(\emptyset) = 0$, $v(N) = 1$
(ii) Monotonicity: $A, B \subseteq N$ and $A \subseteq B$ then $v(A) \leq v(B)$

B. Choquet Integral

The Choquet integral has been introduced to the fuzzy community by Murofushi and Sugeno [26]. During the decade since its introduction, the Choquet Integral has gained considerable attention and success [10], [11].

Definition 4: Let v be a fuzzy measure on X, whose elements are denoted $x_1, \ldots, x_n$. The discrete Choquet Integral of a function $f : X \rightarrow \mathbb{R}^+$ can be written as

$$C_v(f) = \sum_{i=1}^{n} (f(x_{(i)}) - f(x_{(i-1)}))v(A_i) \tag{8}$$

,where $\cdot_{(i)}$ is a permutation on X such that $f(x_{(1)}) \leq f(x_{(2)}) \leq \ldots \leq f(x_{(n)})$, $A_{(i)} = x_{(1)}, \ldots, x_{(n)}$, and $f(x_{(0)}) = 0$

The Choquet Integral is a powerful tool in MCDM as it expresses a certain kind of interaction between different criteria. Some successful applications of Choquet Integral in various fields, such as subjective evaluation, designing of speakers, and time series modeling, can be found in [11]. Unfortunately, all Fuzzy Integrals suffer from the problem of exponential growth of fuzzy measure parameters with respect to the number of criteria.

C. Learning Fuzzy Measure Parameters for Choquet Integral

Beliakov and his team in [4], [3], have shown a method of learning fuzzy measure parameters for Choquet Integral from data. Further, they provide a software package, called *Fuzzy Measure Tool* for the learning of fuzzy measure parameters. We used the Fuzzy Measure Tool for the learning of fuzzy measure parameters for the Choquet Integral. This subsection will briefly discuss the method, with more detailed discussion to be found in [4].

Definition 5: The möbius transformation of a fuzzy measure v is a set function defined for every $A \subseteq N$ as

$$M(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} v(B) \tag{9}$$

The möbius transformation is invertible, and one can recover v by using its inverse, called the Zeta transform,

$$a(A) = \sum_{B \subseteq A} M(B), \quad where\ \forall A \subseteq N \tag{10}$$

The Choquet Integral in equation (8) can be represented in an alternative way using the Möbius transformation, in equation (9), [2], [12].

$$C_v(f) = \sum_{A \subseteq N} M(A)\ h_A(x) \tag{11}$$

where $h_A(x) = min_{i \in A} x_i$.

As we explained earlier, the computational complexity of fuzzy measures are exponential and it becomes too large when the number of inputs goes above 7. Grabisch in [12] discussed that *additive* fuzzy measures can reduce the 2^n number of coefficients to n, but this may reduce the expressive power of the Fuzzy Integral. He suggested a method to reduce the computational complexity based on *k-order additivity* and keep the expressive power of the Choquet Integral to a certain extent. The following interpretation of the definition of *k-order additivity* can be found in [4].

Definition 6: A fuzzy measure v is called k-additive $(1 \leq k \leq n)$ if its Möbius transformation verifies $M(A) = 0$ for any subset A with more than k elements, $\mid A \mid \geq k$, and there exists a subset B with k elements such that $M(B) \neq 0$.

Now, the goal is to find a fuzzy measure v, such that the function $f = C_v$ approximates the desired result y_i, such that $f(X_i) \approx y_i$, where $i \in \{1, \ldots, m\}$ and $m \in \mathbb{N}$. The satisfaction of the approximate equalities $f(X_i) \approx y_i$ is translated into the minimization problem given by,

$$minimize\ \parallel f(X_i) - y_i \parallel \tag{12}$$

Now, in the case when f is the Choquet Integral with respect to a fuzzy measure v, C_v, the above expression can be written as follows and it is subject to satisfying the basic properties of the Choquet Integral [26],

$$minimize\ \parallel C_v(x_{1i}, \ldots, x_{ni}) - y_i \parallel \tag{13}$$

Now according to Beliakov [4], using definitions 5 and 6, the equation (13) can be translated into the following constrained optimization problem,

$$minimize\ \parallel \sum_{A \mid\ |A| \leq k} h_A(x_i) m_A - y_i \parallel \tag{14}$$

such that the following constraints are satisfied,

Authorized licensed use limited to: Australian National University. Downloaded on January 25,2021 at 12:04:26 UTC from IEEE Xplore. Restrictions apply.

$$\sum_{B\subseteq A|\ j\in B,|B|\leq k} m_B \geq 0$$

,

$$\forall A \subseteq N, |\ A\ | \geq 1,\ and\ all\ j \in A$$

$$m_j \geq 0,\ j = 1, \ldots, n,$$

$$\sum_{B\subseteq N|midB|\leq k} m_B = 1$$

It appears that in the Fuzzy Measure Tool, [1], [4], the minimization problem in the expression (14) is further converted to a linear programming problem and the l_1 norm is used to calculate the error.

IV. The Comparison: Fuzzy Signatures and Choquet Integral

In this section, two real world problems, namely High Salary Selection of employees and SARS patient classification problem are used for the comparison of Integral and Fuzzy Signature. The scope of the experiments is to demonstrate if possible the success and the effectiveness of hierarchically organised decision making. First we discuss the Fuzzy Classification Error (FYCLE) to overcome the practical issues which arise when we need to compare the results of two different systems.

A. Fuzzy Classification Error

We formulate the Fuzzy Classification Error (FYCLE) in the following way. We call it *Fuzzy* as it consider smooth transition between good and bad classifications using several categories of error. First, we specify that both *desired* output and *predicted* output of an experiment are in the range $[0, 1]$. Next, we define a set of rules for the classification and these rules are visualised in the following figure 4.

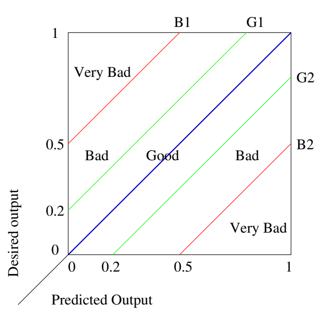


Fig. 4: Fuzzy Classification Error Rules

According to figure 4, there are 3 categories of classifications that can occur, they are *Good*, *Bad*, and *Very Bad (VB)*. Now we assume the pair of *predicted* and *desired* values, of the ith input, respectively taken as *X* and *Y* coordinates of the point P_i on the 2 dimensional fuzzy classification error rule space, figure 4. The fuzzy classification error of an arbitrary point P_i can be written as,

$$FYCLE(P_i) = \begin{cases} 0 \ \ if \ \ P_i \in Good \\ 0.5 \ \ if \ \ P_i \in Bad \\ 1 \ \ if \ \ P_i \in Very\ Bad \end{cases}$$

Let us consider the 4 straight lines, $B1$, $B2$, $G1$, and $G2$, in figure 4. In this experiment, they are equivalent to,

$B1 \equiv y - x - 0.5$
$G1 \equiv y - x - 0.2$
$G2 \equiv y - x + 0.2$
$B2 \equiv y - x + 0.5$

Now, The fuzzy classification error of an arbitrary point P_i can be calculated as,

$$FYCLE(P_i) = \begin{cases} 0 \ \ if \ \ G1(P_i) \leq 0 \ \mathit{AND}\ G2(P_i) \geq 0 \\ 0.5 \ \ if \ \ (B1(P_i) \leq 0 \ \mathit{AND}\ G1(P_i) > 0) \ \mathit{OR} \\ \qquad\qquad (G2(P_i) < 0 \ AND \ B1(P_i) \geq 0) \\ 1 \ \ if \ \ B1(P_i) > 0 \ \mathit{OR}\ B2(P_i) < 0 \end{cases}$$

Next, the Sum of Fuzzy Classification Error (SYCLE) for a set of data with m records can be calculated as,

$$SYCLE(P) = \sum_{i=1}^{m} FYCLE(P_i) \ \ where\ m \in \mathbb{N} \qquad (15)$$

B. High Salary Selection Problem

We select the High Salary Selection problem, as discussed in [8], as the first experiment. The problem is to find the degree of relevance for having a high salary based on the *contacts*, *age*, and *work experience* of an employee. Figure 5 shows a High Salary Selection Fuzzy Signature, which is obtained using domain expert knowledge, for the high salary selection problem. Note that $@_i$ and w_i in figure 5 represent the aggregation function and weighted relevance of the node i, respectively.

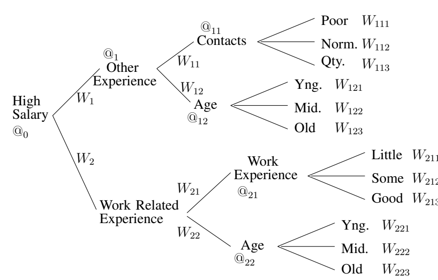


Fig. 5: High Salary Selection Fuzzy Signature

In general, Choquet Integral takes normalised inputs and produce the output in the same range, which is a rank between [0,1] of the likelihood of getting a high salary. On the other hand, Fuzzy Signatures take fuzzyfied inputs and produce the fuzzy values as the output. Fuzzy Signatures, for the High Salary Selection problem, give the degree of membership in the high salary fuzzy set of an employee's salary as the output result. We faced a practical issue at the very beginning of the comparison of Fuzzy Integrals and

Authorized licensed use limited to: Australian National University. Downloaded on January 25,2021 at 12:04:26 UTC from IEEE Xplore. Restrictions apply.

Fuzzy Signatures concepts, that is to decide which output target among the two output distributions, that is normalised output and fuzzyfied output, is best for learning and testing of Fuzzy Integral based systems and Fuzzy Signature based system. One option may be to consider using the Fuzzy Integrals to aggregate the fuzzyfied input data but it is obvious that this can increase the computational complexity of the Fuzzy Integral based system, as this method increases the number of input dimensions by at least 3 times more than the number of normalised input dimensions.

Also, it is obvious that the fuzzyfied output target is the best for the experiments with the Fuzzy Signatures as Fuzzy Signatures take fuzzy values for the inputs. Moreover, for Fuzzy Integrals the fuzzyfied output target values can be hard to learn as they are fuzzified, according to the high salary fuzzy set, from the normalised distribution of output data. For the comprehensibility of this experiment, we decided to experiment with Choquet Integral and Fuzzy Signatures both with normal and fuzzyfied output targets to discover which output target is good for each method.

We used the two learning methods explained in the previous section to extract Fuzzy measures and WRAO for the Choquet Integral and Fuzzy Signature respectively for the High Salary Selection problem. Now, we can use the fuzzy classification error to visualise and classify the results of the experiments. Table I shows the MSE and SYCLE for both training and testing phases of the experiment with all four possibilities that is Choquet Integral and Fuzzy Signatures with both normal and fuzzyfied output targets. Figures 6 and 7 illustrate the fuzzy classification error of the best test results for Choquet Integral and High Salary Selection Fuzzy Signature respectively.

TABLE I: High Salary Selection Experiment

	Test MSE	Test SYCLE	Train MSE	Train SYCLE
Integral (norm. output)	0.02661	54.5	0.02620	55.5
Fuzzy Sig.(norm. output)	0.0216	47.5	0.0216	37.5
Integral (fuzzy output)	0.0837	185	0.1052	210.5
Fuzzy Sig. (fuzzy output)	**0.0149**	**23.5**	**0.0152**	**20**

The results of this experiment show that Choquet Integral can learn and classify the data in High Salary Selection problem. But, High Salary Selection Fuzzy Signature has a much lower sum of fuzzy classification error (SYCLE) and lower mean squared error (MSE) as shown in table I, compared to that of the Choquet Integral. Further, Fuzzy Signature model has reduced SYCLE in the testing phase compared to that of the training phase, which importantly shows that Fuzzy Signature training is generalising and avoiding over fitting problems as well.

C. SARS Patient Classification Problem

In this experiment we use SARS patient classification problem to compare the performance of the two methods Choquet Integral and Fuzzy Signature.

Medical practitioners know that for a certain disease, such as SARS, they need to check the patient for possible fever,

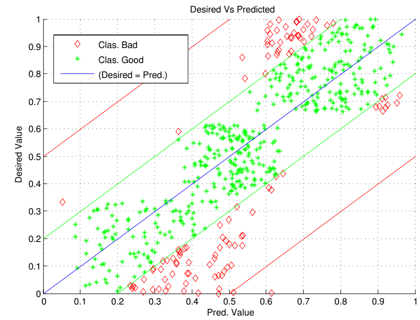


Fig. 6: Test Fuzzy Classification Error Salary: Choquet Integral

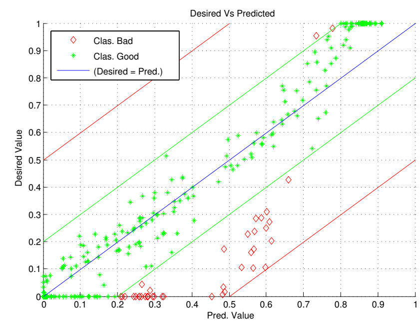


Fig. 7: Test Fuzzy Classification Error Salary Fuzzy Signature

hypertension, conditions of nausea, and abdominal pain [14], [27], [31]. In addition, it is fairly important to monitor the fever regularly during the day, as well as the blood pressure. Figure 8 shows a SARS polymorphic fuzzy signature, which is constructed based on domain expert knowledge. Each symptom check has been divided into a number of doctors' diagnosis levels, such as *slight*, *moderate*, and *high* for body temperature (fever), *low*, *normal*, and *high* for both measurements of blood pressure, *slight*, *medium*, and *high* for nausea, and *slight*, and *high* for abdominal pain. Also, the SARS fuzzy signature contains three levels of hierarchies, which can be aggregated using different aggregations.

In figure 8, the notations a_{ij}, $@_{ij}$, and w_{ij} represent the input value, aggregation function, and weight for the branch ij of the SARS Fuzzy Signature. Test and train data sets, which were used for the experiments, are a combination of SARS, blood pressure, pneumonia, and normal patients data. The desired output of this experiment is a classification that needs to give full degree of confidence, ie. 1, for the SARS patient data and zero degree of confidence for the other condition and normal people data.

Authorized licensed use limited to: Australian National University. Downloaded on January 25,2021 at 12:04:26 UTC from IEEE Xplore. Restrictions apply.

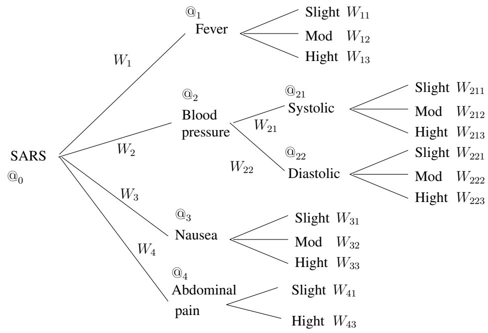


Fig. 8: SARS Patient Classification Fuzzy Signature

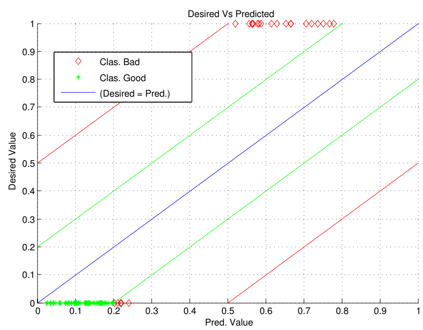


Fig. 9: Training Predicted Vs Desired SARS: Integral

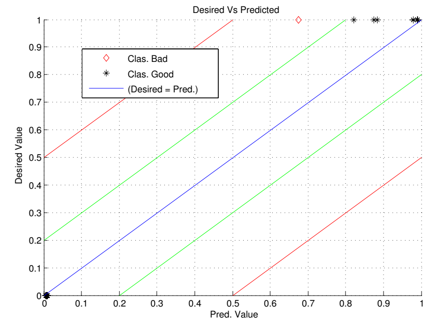


Fig. 10: Training Predicted Vs Desired SARS Signature

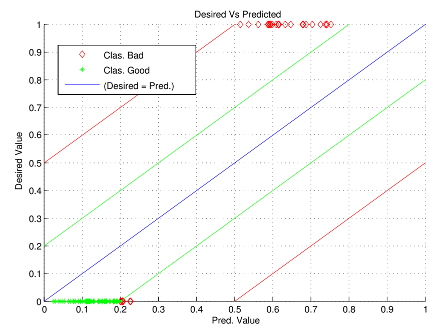


Fig. 11: Test Predicted Vs Desired SARS: Integral

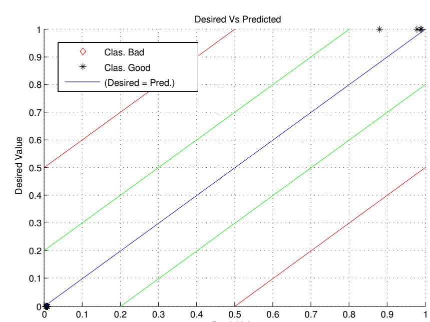


Fig. 12: Test Predicted Vs Desired SARS Signature

Figures 9 and 10 show the training results of Choquet Integral and Fuzzy Signature. Figures 11 and 12 show the test results of the two methods. Table II shows the MSE and SYCLE for training and testing phases of the two methods.

TABLE II: SARS Patient Classification Experiment

	Train MSE	Train SYCLE	Test MSE	Test SYCLE
Integral	0.04253	13.5	0.04948	16.5
Fuzzy Signature	**0.0017**	**0.5**	**0.0001**	**0**

The results of the 2nd experiment show that Choquet Integral can not learn or classify the SARS patients data properly. The SARS patients classification Fuzzy Signature has very much less SYCLE and MSE as shown in table II, compared to that of the Choquet Integral. Further, in this experiment also, we observed that the Fuzzy Signature model has reduced SYCLE in the testing phase compared to that of the training phase.

According to the results of the two real world experiments, it is clear that our Fuzzy Signature concept out-performs the Choquet Integral in both experiments. An interesting future research direction would be to compare Fuzzy Signatures with hierarchically organised Choquet Integral [30].

Authorized licensed use limited to: Australian National University. Downloaded on January 25,2021 at 12:04:26 UTC from IEEE Xplore. Restrictions apply.

V. Conclusions

Fuzzy Signatures are less computationally complex compare to the exponential growth of Fuzzy Measure parameter against the number of input dimensions. Fuzzy Signatures and Choquet Integrals are compared for two real world problems, importantly, Fuzzy signatures work better for both real world problems. From the results of the two experiments, we conclude the importance of hierarchical organization and aggregation of data and also the importance of use of non-homogeneous aggregation functions for sub-sets of the problem.

References

[1] G. Beliakov, Aggregation operator approximation tool (aotool), Tech. rep., http://www.it.deakin.edu.eu/gleb.

[2] G. Beliakov, Programming library fmtools for handling and fitting fuzzy measures, Tech. rep., Deakin University (2007). URL `http://www3.it.deakin.edu.au/~gleb/`

[3] G. Beliakov, R. Mesiar, L. Valaskova, Fitting generated aggregation operators to empirical data, INTERNATIONAL JOURNAL OF UNCERTAINTY FUZZINESS AND KNOWLEDGE BASED SYSTEMS 12 (2) (2004) 219–236.

[4] G. Beliakov, A. Pradera, T. Calvo, Aggregation Functions: A Guide for Practitioners, Springer, Heidelberg, Berlin, New York, 2007.

[5] G. Beliakov, J. Warren, Appropriate choice of aggregation operators in fuzzy decisionsupport systems, Fuzzy Systems, IEEE Transactions on 9 (6) (2001) 773–784.

[6] D. Dubois, H. Prade, Weighted minimum and maximum operations in fuzzy set theory., INFO. SCI. 39 (2) (1986) 205–210.

[7] H. Dyckhoff, W. Pedrycz, Generalized means as model of compensative connectives, Fuzzy Sets and Systems 14 (2) (1984) 143–154.

[8] T. Gedeon, K. Wong, D. Tikk, Constructing hierarchical fuzzy rule bases for classification, Fuzzy Systems, 2001. The 10th IEEE International Conference on 3.

[9] M. Grabisch, Characterization of fuzzy integrals viewed as aggregation operators, Fuzzy Systems, 1994. IEEE World Congress on Computational Intelligence., Proceedings of the Third IEEE Conference on (1994) 1927–1932.

[10] M. Grabisch, Fuzzy integral in multicriteria decision making, Fuzzy Sets and Systems 69 (3) (1995) 279–298.

[11] M. Grabisch, The application of fuzzy integrals in multicriteria decision making, European Journal of Operational Research 89 (3) (1996) 445–456.

[12] M. Grabisch, k-order additive discrete fuzzy measures and their representation, Fuzzy Sets and Systems 92 (2) (1997) 167–189.

[13] M. Hajnal, L. Kóczy, Classification of textures by vectorial fuzzy sets,", approximate reasoning in decision analysis, edited mm gupta and e. sanchez (1982).

[14] L. Hsu, C. Lee, J. Green, B. Ang, N. Paton, L. Lee, J. Villacian, P. Lim, A. Earnest, Y. Leo, Severe acute respiratory syndrome (sars) in singapore: clinical features of index patient and initial contacts, Emerg Infect Dis 9 (6) (2003) 713–7.

[15] L. Kóczy, Vector valued fuzzy sets, BUSEFAL- BULL STUD EXCH FUZZIN APPL (1980) 41–57.

[16] L. Kóczy, T. Vámos, G. Biró, Fuzzy signatures, in: EUROFUSE-SIC '99, 1999.

[17] K. Levenberg, A method for the solution of certain nonlinear problems in least squares, Q. Appl. Math 2 (1944) 164–168.

[18] D. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, Journal of the Society for Industrial and Applied Mathematics 11 (2) (1963) 431–441.

[19] B. S. U. Mendis, T. D. Gedeon, L. T. Kóczy, Investigation of aggregation in fuzzy signatures, in: 3rd International Conference on Computational Intelligence, Robotics and Autonomous Systems, Singapore, vol. CD ROM, 2005.

[20] B. S. U. Mendis, T. D. Gedeon, L. T. Kóczy, Flexibility and robustness of hierarchical fuzzy signature structures with perturbed input data, in: International Conference of Information Processing and Management of Uncertainty in Knowledge Based Systems (IPMU), IPMU proceedings, Paris, 2006.

[21] B. S. U. Mendis, T. D. Gedeon, L. T. Kóczy, Generalised weighted relevance aggregation operators for hierarchical fuzzy signatures, in: International Conference on Computational Inteligence for Modelling Control and Automation and International Conference on Intelligent Agents Web Technologies and International Commerce (CIMCA'06), Sydney, Australia, 2006.

[22] B. S. U. Mendis, T. D. Gedeon, L. T. Kóczy, Learning generalized weighted relevance aggregation operators using levenberg-marquardt method, in: Sixth International Conference on Hybrid Intelligent Systems (HIS'06) and 4th Conference on Neuro-Computing and Evolving Intelligence (NCEI 06'), New Zealand, 2006.

[23] B. S. U. Mendis, T. D. Gedeon, L. T. Kóczy, Learning weights from observations for hierarchical fuzzy signatures, in: 3rd International Conference Soft Computing and Intelligent Systems and 7th International Symposium Advanced Intelligent Systems, Tokyo, Japan, 2006.

[24] B. S. U. Mendis, T. D. Gedeon, L. T. Kóczy, On the issue of learning weights from observations for fuzzy signatures, in: World Automation Congress (WAC), Budapest, Hungary, 2006.

[25] T. Murofushi, M. Sugeno, Fuzzy measures and fuzzy integrals, Fuzzy Measures and Integrals: Theory and Applications, Studies in fuzziness and soft computing 3.

[26] T. Murofushi, M. Sugeno, An interpretation of fuzzy measures and the choquet integral as an integral with respect to a fuzzy measure, Fuzzy Sets and Systems 29 (2) (1989) 201–227.

[27] S. Poutanen, D. Low, B. Henry, S. Finkelstein, D. Rose, K. Green, R. Tellier, R. Draker, D. Adachi, M. Ayers, et al., Identification of severe acute respiratory syndrome in canada, N Engl J Med 348 (2003) 1995–2005.

[28] W. Slany, J. Vascak, A consistency checker for a fuzzy diagnosis system applied to warmrolling-mills in steelmaking plants, Fuzzy Systems, 1996., Proceedings of the Fifth IEEE International Conference on 1.

[29] M. Sugeno, Theory of fuzzy integrals and its applications, Ph.D. thesis, Tokyo Institute of Technology (1974).

[30] M. Sugeno, K. Fujimoto, T. Murofushi, Hierarchical decomposition theorems for choquet integral models, Fuzzy Systems, 1995. International Joint Conference of the Fourth IEEE International Conference on Fuzzy Systems and The Second International Fuzzy Engineering Symposium., Proceedings of 1995 IEEE International Conference on 4.

[31] K. Tsang, P. Ho, G. Ooi, W. Yee, T. Wang, M. Chan-Yeung, W. Lam, W. Seto, L. Yam, T. Cheung, et al., A cluster of cases of severe acute respiratory syndrome in hong kong (2003).

[32] K. Wong, T. Gedeon, L. Kóczy, Construction of fuzzy signature from data: an example of sars pre-clinical diagnosis system, Fuzzy Systems, 2004. Proceedings. 2004 IEEE International Conference on 3.

[33] R. Yager, On ordered weighted averaging aggregation operators in multicriteria decisionmaking, IEEE Transactions on Systems, Man and Cybernetics 18 (1) (1988) 183–190.

Authorized licensed use limited to: Australian National University. Downloaded on January 25,2021 at 12:04:26 UTC from IEEE Xplore. Restrictions apply.