Information Sciences 194 (2012) 85–106

Contents lists available at SciVerse ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

# Complex Structured Decision Making Model: A hierarchical frame work for complex structured data

B.S.U. Mendis *, T.D. Gedeon

*Research School of Computer Science, The Australian National University, Canberra, ACT 0200, Australia*

ARTICLE INFO

*Article history:*
Available online 30 August 2011

*Keywords:*
Complex Structured Decision Making
Fuzzy signatures
Polymorphic fuzzy signatures
Weighted Relevance Aggregation Operator
Fuzzy integrals
Choquet integral

ABSTRACT

We introduce a hierarchical framework we call Complex Structured Decision Making model for complexly structured knowledge representation in intelligent decision making. We show that our model extends non-hierarchical (flat) decision making models to hierarchical decision making models that are similar to comprehensible human decision making processes. Further, we make an argument that hierarchial representation of knowledge in a Complex Structured Decision Making Model simplifies the approximation of aggregation functions to easily adapt to the underline relation of the system. Additionally, using a real world complex structured data set, we show that hierarchically organized Fuzzy Integrals, e.g. Choquet Integral, and Sugeno Integral and Fuzzy Signatures outperform these non-hierarchical Fuzzy Integrals.



## 1. Introduction

Human decision making is a comprehensible hierarchical process, in which cognition processes lead to the selection of a set of actions among many variations. We are inspired by the hierarchical organization of the human decision making process as it reduces the space and time complexity, as it may require less information compared to a flat organization, and it simplifies the approximation of the aggregation of information. It is obvious that the hierarchical organization of data reduces the space and time complexity of the decision making model [12,22]. The hierarchical structure of a decision making system makes it possible to neglect some input information during the decision making [30,36] whereas non-hierarchical ('flat') decision making systems always need all the input information for their decision making process [38,50]. This is an additional advantage of hierarchical decision making systems in situations when all the information may not be available. As the hierarchical systems organize data into small groups of hierarchies, it is easy to approximate the aggregations of these small sets of information to achieve the final result. That is, a hierarchical decision making system uses a set of hierarchically organized local aggregation functions to approximate the desired global preference relation of the system. On the other hand, flat decision making systems try to approximate the global preference relation using a single aggregation function. This may reduce the accuracy of a flat decision making system, especially when large numbers of input variables are available.

In this paper we introduce our Complex Structured Decision Making (CSDM) model, which can be seen as a hierarchically extended application of the Multi-criteria Decision Making (MCDM) paradigm [10]. A CSDM model represents knowledge in the form of hierarchically organized information. We investigate Fuzzy Integrals [13] and Fuzzy Signatures [27,29] using the Weighted Relevance Aggregation Operator (WRAO) [28] for the selection of best aggregation modules for the CSDM method. However, Fuzzy Integral models are computationally very impractical for large data sets, as they need $2^n$ parameters, where $n$ is the number of input dimensions. Thus, the computational cost increases significantly with any increase of the number of

* Corresponding author.
*E-mail addresses:* sumudu@cs.anu.edu.au (B.S.U. Mendis), tom@cs.anu.edu.au (T.D. Gedeon).

0020-0255/$ - see front matter 
doi:10.1016/j.ins.2011.08.010

input dimensions. Our Complex Structured Decision Making approach uses hierarchical structured information representation and advanced aggregation functions to handle complexly structured data more efficiently and effectively.

In Section 2 we discuss the Multi-criteria Decision Making (MCDM) paradigm and fundamentals of Fuzzy Integrals as aggregation methods for MCDM. In Section 3, we introduce the Complex Structured Decision Making (CSDM) model and further propose fuzzy signatures with Weighted Relevance Aggregation Operator as a practical model for CSDM. In Section 4 we show two methods of extracting fuzzy measures and WRAO for Fuzzy Integrals and Fuzzy Signatures, respectively. Then in Section 5 we show the various results of two experiments with real world complex problems, where one experiment is based on personnel management and other experiment is based on medical diagnosis. In this section we also introduce the concept of fuzzy classification error for better visualization of test results of an experiment compared to sum of squares error.

## 2. Multi-criteria decision making and fuzzy integrals

In this section we discuss the fuzzy measure concept and the Multi-criteria Decision Making (MCDM) paradigm. The MCDM is a well-known paradigm for intelligent decision making and is widely used compared to the Decision Making Under Uncertainty paradigm. Next, we discuss fuzzy measures and Fuzzy Integrals, in order to use them for aggregation in MCDM.

### 2.1. 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 [39,41,43] the additivity property can be too rigid. The fuzzy measure concept is a generalization of additive measure concept as it replaces the additivity by the weaker condition of monotonicity [37,18,17].

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

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

### 2.2. Multi-criteria decision making paradigm

We recall the Multi-criteria Decision Making (MCDM) paradigm [10,45,19].

**Definition 2.** A Multi-criteria Decision Making (MCDM) method is a triple $(N, X, \succeq)$, where

(i) $N = \{1,\ldots,n\}$ is the set of criteria to satisfy
(ii) X is the Cartesian product, $X = X_1 \times X_2 \times \cdots \times X_i$ that corresponds to the set of alternatives $X_i$, being the evaluation scale related to criterion $i$ $(i \in N)$
(iii) $\succeq$ is a preference relation on $X$

An aggregation function $M_v: L^n \to L$ is defined from a fuzzy measure $v$ and local utility function $u_i: X_i \to L$, where $(i \in N)$, such that

$U_\upsilon(x) := M_\upsilon[u_1(x_1),\ldots,u_n(x_n)]$.

where $x \in X$ and $U_v : X \to \mathbb{R}$ is a global preference function, which satisfies

$x \succeq y \iff U_v(x) \geqslant U_v(y)(x, y \in X)$

Here $x = (x_1,\ldots,x_n)$ is the result of alternatives after applying their criteria.

The following two examples describe an applications of the MCDM to a real world scenario.

**Example 1.** A personnel manager has decided to increase the salary of 3 employees, called Smith, McCreath, and Graham. The problem is to find the appropriate salaries for the 3 employees (or technically "alternatives") using the criteria *age*, *contacts*, and *experience*. We can write the rating of *Smith's* salary to be a *High Salary (HS)*, using the MCDM method as,

$HS(Smith) = M_\upsilon[u_1(age(Smith)), u_2(contacts(Smith)), u_3(experience(Smith))]$, where $M_v$ is an appropriate aggregation function and $u_1,\ldots,u_3$ are local utility functions, which are defined as in Definition 2, being mappings from the range of $(criteria)_i : \to L$.

**Example 2.** We consider the problem of finding a Qualitative Grading (QG) scheme for a high school with respect to following 3 subjects: *mathematics* (M), *physics* (P), and *literature* (Lit.) [15]. Let us assume school has 3 students namely Euler, Einstein, and Hercules. For an example, we can write the Qualitative Grade of *Euler's* total marks, using the MCDM as,

$QG(Euler) = M_v[u_1(Math(Euler)), u_2(Physics(Euler)), u_3(Literature(Euler))]$, where $M_v$ is an appropriate aggregation function and $u_1, \ldots, u_3$ are local utility functions, which are defined as in Definition 2, being mappings from the range of $(criteria)_i : \rightarrow L$.

### 2.3. Sugeno integral

Fuzzy Integrals were introduced by Sugeno (they also called Sugeno Integrals) [38,37].

**Definition 3.** Let $v$ be a normalized fuzzy measure on $X$, whose elements are denoted $x_1, \ldots, x_n$. The discrete Sugeno Integral of a function $f: X \rightarrow [0,1]$ can be written as

$$S_v(f) = \bigvee_{i=1}^{n} (f(x_{(i)}) \wedge v(A_i)) \tag{1}$$

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

### 2.4. Choquet integral

The Choquet integral was introduced to the fuzzy community [42,40] by Murofushi and Sugeno [34]. During the decade since its introduction, the Choquet Integral has gained a considerable attention and success [14,15,49].

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

$$C_v(f) = \sum_{i=1}^{n} (f(x_{(i)}) - f(x_{(i-1)}))\, v(A_i) \tag{2}$$

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

The Fuzzy Integral is a powerful tool in MCDM as it expresses a certain kind of interaction between different criteria. Some successful applications of Fuzzy Integrals in various fields, such as subjective evaluation, design of speakers, risk assesment and time series modeling, can be found in [15,6]. Unfortunately, all Fuzzy Integrals suffer from the problem of exponential growth of fuzzy measure parameters (weights) with respect to the number of criteria. Example 3 shows the expressive power of Choquet Integral as compared to the conventional weighted mean method.

**Example 3.** We continue the students qualitative examination marks classification problem in Example 2. The teacher of these students has to evaluate them according to their level in 3 subjects: Mathematics (M), Physics (P), and Literature (Lit.). She has to decide the evaluation of students according to the following criteria,

(i) Scientific subjects (mathematics and physics) are more important.
(ii) The 2 scientific subjects are similar, that is if a is student good in one of these subjects, that student is also good at the other subject.
(iii) Students good at scientific subjects and literature are rather uncommon and must be favored.

As in Table 1, the teacher decides the weights of 3, 3, and 2 for the subjects mathematics, physics, and literature respectively for a weighted mean type evaluation. Table 2 shows the fuzzy measure parameters for the same students' evaluation. Finally, Table 3 shows the corresponding students' data and the results (last 2 columns) of weighted mean and Choquet Integral based evaluations.

According to the results of the 2 evaluations, it is clear that the Choquet Integral has better ranked the students according to teacher's preference function.

## 3. Complex Structured Decision Making Method

In this section we first introduce the Complex Structured Decision Making (CSDM) model as a hierarchically extended application of the Multi-criteria Decision Making (MCDM) paradigm [10]. CSDM models represent knowledge in the form of hierarchically organized information for better modeling of the decision making process. Next we discuss the theory of

**Table 1**
Weights for weighted mean.

| M | P | L |
|---|---|---|
| 3 | 3 | 2 |

**Table 2**
Fuzzy measure parameters (weights).

| M | P | L | MP | ML | PL | MPL |
|---|---|---|---|---|---|---|
| 0.45 | 0.45 | 0.3 | 0.5 | 0.9 | 0.9 | 1 |

**Table 3**
Qualitative grading of students using WM & CI.

| Student | Mathematics | Physics | Literature | Weighted mean | Choquet integral |
|---|---|---|---|---|---|
| Euler | 18 | 16 | 10 | 15.25 | 13.9 |
| Einstein | 10 | 12 | 18 | 12.75 | 13.6 |
| Hercules | 14 | 15 | 15 | 14.62 | 14.9 |

fuzzy signatures [27,29] and the Weighted Relevance Aggregation Operator (WRAO) [28] as an aggregation selection for CSDM model.

### 3.1. Complex Structured Decision Making

In general, decision making systems need to handle problems which are structurally complex [12,47] and have a global preference [9] between the outcomes. The structural complexity of a problem expresses the correlation of different input dimensions, in a hierarchical structure. Thus, the structural complexity expresses the "interconnectedness" of the system based on the correlations of the inputs of a system. The ideal hierarchical organization of a decision problem enhances the accuracy of the results and reduces computational space and time complexity. The global preferences of the outcomes of the system against the set of input dimensions is a relation, which can be approximated using a hierarchically organized set of non-homogeneous aggregation functions. We call this the "interdependent feature" of the system as it expresses the inter-dependance of the input dimensions in order to preserve the monotonicity of the system. We argue that hierarchical organization of input data will simplify the approximation of the global preference relation of a system, by using sets of local preferences for a lower number of input dimensions, which are arranged hierarchically. Most existing methods focus only on structural complexity [36,8,7] or on the interdependent feature of the problem with non-hierarchical approaches [50,38].

The MCDM model is mainly used for generation of decision making systems based on the underlying preference relation of the decision making problem. As we discussed in the introduction, we are inspired by the human comprehensibility of decision making, where information is organized hierarchically. In this paper, we present a method in which we consider both structural complexity and the interdependent features of the problem simultaneously to construct the decision making system.

**Definition 5.** A Complex Structured Decision Making (CSDM) method is a triple $(N, X', \succeq)$, where

(i) $N = \{1,\ldots,n\}$ is the set of criteria to satisfy,
(ii) $X'$ is the Cartesian product, $X' = X'_1 \times X'_2 \times \cdots \times X'_m$, where $m \leqslant 2^n - 1$ and $X'_i \in \rho^*(X)$ such that $\rho^*(X) = \{\prod_{i\in S} X_i, S \subseteq N \text{ with } S \neq \emptyset \text{ and } S \neq N\}$ corresponds to the set of alternatives $X_i$ being the evaluation scale related to criterion i,
(iii) $\succeq$ is a preference relation on $X'$.

Now let $x' = (x'_1, \ldots, x'_k) \in X'$, where $k \leqslant 2^n - 1$ and $x = (x_1,\ldots,x_n) \in X$.

(a) $U_0^v$ is a preference function s.t. $U_0^v : X' \to L$.
(b) $M_0^v$ is an aggregation functions s.t. $M_0^v : L^k \to L$.
(c) $\{u_i\}$ is the set of local utility functions s.t. $u_i : X_i \to L$.

Where

$$\overline{U}_0^v(x) := M_0^v[u'_1(x'_1), \ldots, u'_k(x'_k)]$$
$$\text{and} u'_j(x'_j) = \begin{cases} u_i(x_i); & \text{if } x'_j \in x \\ U_j^v(x'_j); & \text{otherwise} \end{cases}$$

Note that, $U_j^v(x'_j)$ recursively follows this definition to define the next hierarchy.

In this definition, $i \in N, j \in [1\ldots k]$, and $v$ is a fuzzy measure.

Further, the preference relation $U_0^v$ is monotonic, similar to Definition 2.

According to this definition, it is clear that the original form of the MCDM can be obtained when $X'$ reduces to $X$ and $k = n$. In practical applications, a more restricted case of CSDM i.e. $k \leqslant n$, needs to be applied to reduce the computational complexity. In the following two examples, we illustrate CSDM in practice. Example 4 expresses a simple situation and Example 5 expresses a more complex situation.

**Example 4.** The Fig. 1(a) shows a complex structure for a decision problem with criteria $\{C_1, C_2, C_3\}$.

Let us assume an alternative, $a = (a_1, a_2, a_3)$ and in Fig. 1, each $x_i = C_i(a_i)$. Then, we get $x = (x_1, x_2, x_3)$ and $\rho^*(x) = \{x_1, x_2, x_3, x_1x_2, x_1x_3, x_2x_3\}$. Further, $x'_0 = (x_1x_2, x_3)$.

Now, we can formulate the first level of the complex structured decision problem as:

$U_0^v(x_1x_2x_3) = M_0^v(u'_1(x'_1), u_3(C_3(a_3)))$

and now $u'_1(x'_1)$ recursively follows the Definition 5 to create a further hierarchy i.e. now we get $x = (x_1, x_2)$, $\rho^*(x) = \{x_1, x_2\}$, and $x'_1 = (x_1, x_2)$. According to Fig. 1(a), we write the new branch as,

$U_1^v(x'_1) = M_1^v(u_1(C_1(a_1)), u_2(C_2(a_2)))$

**Example 5.** The Fig. 1(b) shows a complex structure for a decision problem with criteria $(C_1, C_2, C_3, C_4)$ and in Fig. 1, $x_i = C_i(a_i)$.

Let us assume an alternative, $a = (a_1, a_2, a_3, a_4)$. Then, we get $x = (x_1, x_2, x_3, x_4)$ and $\rho^*(x) = \{x_1, x_2, x_3, x_4, x_1x_2, x_1x_3, x_1x_4, x_2x_3, x_2x_4, x_3x_4, x_1x_2x_3, \ldots\}$ Further we have $x'_0 = (x_1x_2x_3, x_4)$.

Now we can formulate the first level of the complex structure as:

$U_0^v(x_1x_2x_3x_4) = M_0^v(u'_1(x'_1)), u_4(C_4(a_4)))$

Now $u'_1(x'_1)$ recursively follows the Definition 5 to create the next level i.e. now we get $x = (x_1, x_2, x_3)$, $\rho^*(x) = \{x_1, x_2, x_3, x_1x_2, x_1x_3, x_2x_3\}$, and $x'_1 = (x_1, x_2x_3)$. According to Fig. 1(b), we write the new branch as, $U_1^v(x'_1) = U_1^v(x_1x_2x_3) = M_1^v(u_1(C_1(a_1)), u'_{12}(x'_{12}))$.

Again $u'_{12}(x'_{12})$ recursively follows the Definition 5

i.e. now we get $x = (x_2, x_3)$, $\rho^*(x) = \{x_2, x_3\}$, and $x'_{12} = (x_2, x_3)$. According to Fig. 1(b), we write the new branch as,$U_{12}^v(x'_{12}) = M_2^v(u_2(C_2(a_2)), u_3(C_3(a_3)))$

It is clear that according to Definition 5, to represent Complex Structured Decision Making Models, we need a hierarchical aggregation function that can facilitate the formulation of both structure and multi-aggregations. It is already well known in fuzzy theory that we can use fuzzy sets to represent local utility functions [4,5] in Multi-Criteria Decision Making problems. Therefore, we use the same approach in the Complex Structured Decision Making method. Next, we discuss the concept of Fuzzy Signatures as a practical approach to the Complex Structured Decision Making (CSDM) model.

### *3.2. Hierarchical fuzzy signatures*

Fuzzy signatures can describe, compare and classify objects with complex structures and interdependent features [26,30]. The hierarchical organization of fuzzy signatures expresses the structural complexity of a problem effectively. The hierarchically organized set of aggregation functions of fuzzy signature can be used to approximate the underline preference relation of the problem easily compare to conventional flat approaches.

#### *3.2.1. 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 potentially 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 in the Complex Structured Decision Making context.

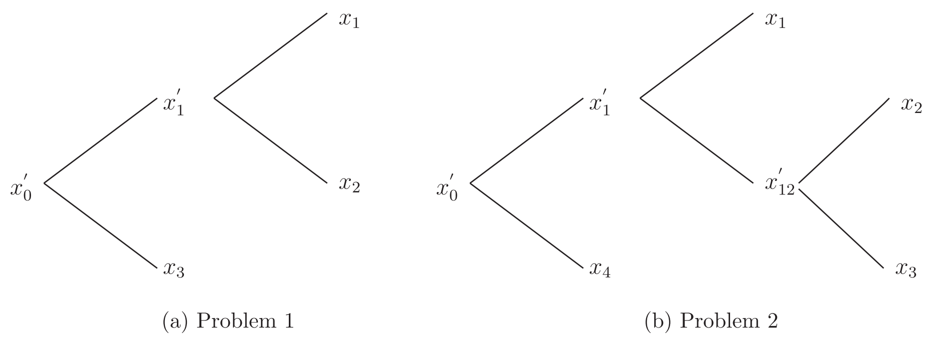


**Fig. 1.** Complex structures of two decision problems.

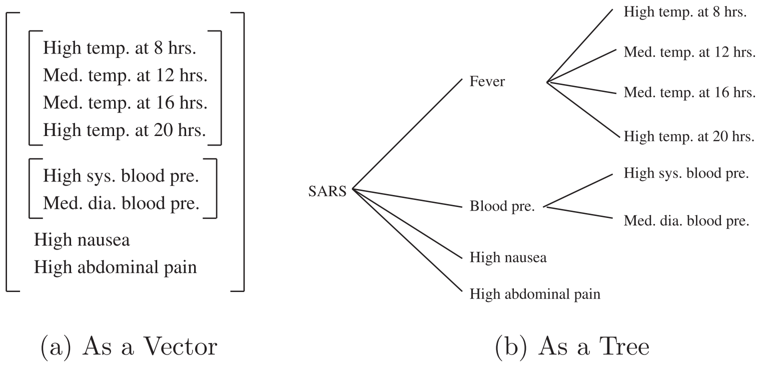


**Fig. 2.** Example fuzzy signature.

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 to the Complex Structured Decision Making (CSDM) model. Now, we recall the fuzzy signature concept introduced in [21].

**Definition 6.** Fuzzy Signature is a VVFS, where each vector component is another VVFS (branch) or a atomic value (leaf), and denoted by,

$$A : X \to [a_i]_{i=1}^{k} \left( \equiv \prod_{i=1}^{k} a_i \right), \tag{3}$$

where

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

and $\Pi$ describes the Cartesian product.

The Fig. 2(a) shows an example fuzzy signature [48]. This fuzzy signature describes an individual SARS patient, which is a data point[1] among many SARS data collected in the year 2003 [20,35,44]. The Fig. 2(b) shows the hierarchical view of the fuzzy signature shown in Fig. 2(a).

Fig. 3 shows an example aggregation of a fuzzy signature using *max* and *min* as the aggregation functions. In this example the final aggregated value of the fuzzy signature $S$ of data point $d$ is 0.6. Now, we introduce our notation in Lemma 2 to represent this value.

**Lemma 1.** *Let $S_{d_k}$ be a fuzzy signature that represent the data point $d_k$. The notation $S_{d_k}(d_k)(\in [0,1])$ represents the final aggregated value of the fuzzy signature $S_{d_k}$.*

Now, using Lemma 2, we can write the final atomic value of the fuzzy signature $S$ in the Fig. 3 as $S(d) = 0.6$.

*3.2.2. Polymorphic fuzzy signatures*

A fuzzy signature is a real world descriptor of an individual data point. Thus, in the initial concept of fuzzy signature, as shown in Fig. 2, one fuzzy signature is needed to model each individual data point of that problem. But in most situations, in real world decision making applications, people may not be interested or able to invest a large amount of time and funding in achieving the best possible solution.

In some situations, we observed [29,32] that we may be able to find a single fuzzy signature for a set of individual data points and for reducing the number of fuzzy signatures required to implement a decision making model. We call such a fuzzy signature a Polymorphic Fuzzy Signature (PFS) for the set of data points it represents.

**Definition 7.** Let $X$ be the universe of the parent node and $X_1 \times X_2 \times \cdots \times X_n$ be the domain of $n$ children. The mapping between parent and children of the polymorphic fuzzy signature $\widehat{S}_i$ is as follows:

$$\widehat{S}_i : X \to \left[a^i\right]_{i=1}^{n} \left( \equiv \prod_{i=1}^{n} a_i \right). \tag{4}$$

[1] In the fuzzy signature concept, a data point means a collection of data which represents an event, e.g. in medical applications, a patient's data record of a whole day can be considered a single data point [48].

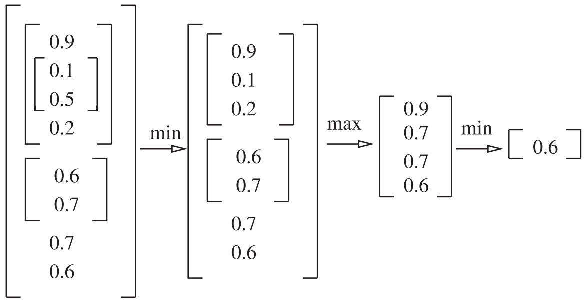


**Fig. 3.** Aggregation of fuzzy signatures.

$$a_i = \begin{cases} [a_{ij}]_{j=1}^{n_i}; & \text{if branch } (n_i > 1) \\ F_i(x_i); & \text{if leaf} \end{cases} \tag{5}$$

Where $F_i$ is a fuzzy subset of universe $X_i$ such that $F_i(x_i) \in [0,1]$ and $x_i \in X_i$.

As a summery of the Definition 7, unlike in ordinary fuzzy signatures, in the polymorphic fuzzy signatures leaf nodes are fuzzy sets. Fig. 4 shows an example polymorphic fuzzy signature structure with two arbitrary levels $g$ and $(g + 1)$. In this figure, the node $q,\ldots,i$ decomposes to another set of branches: $[a_{q,\ldots,ij}]_{l=1}^{n}$. The branch $q,\ldots,i1$ is a leaf that defines inputs from a fuzzy subset $F_{q,\ldots,i1}$.

**Lemma 2.** *Let $\widehat{S_i(d_k)}(\in [0,1])$ represents the degree of match between polymorphic fuzzy signature $\hat{S}_i$ and the data point $d_k$.*

Below, we formulate an objective function to measure the optimality of a polymorphic fuzzy signature [29,32]. This could be especially used for learning of polymorphic fuzzy signatures from data.

**Definition 8.** Let $\{d_1, d_2, \ldots, d_n\}$ be a collection of data points for a certain problem and let $A = \{S_{d_1}, S_{d_2}, \ldots, S_{d_n}\}$ be the collection of fuzzy signatures they represent respectively. Further let $B = \{\widehat{S}_1, \widehat{S}_2, \ldots, \widehat{S}_m\}$ be a set of possible polymorphic fuzzy signatures for the same problem. Then $\widehat{S}_l(\in B)$ is the optimal polymorphic fuzzy signature of the set $A$ if:

$$\sum_{i=1}^{n} |\widehat{S}_l(d_i) - S_i(d_i)| \leqslant \sum_{i=1}^{n} |\widehat{S}_k(d_i) - S_i(d_i)| \quad \forall \widehat{S}_k \in B \tag{6}$$

The following example illustrates the concept of the optimal PFS.

**Example 6.** The set of fuzzy signatures, $\{S_a, S_b, S_c, S_d\}$ in Fig. 5 describe four different SARS patients' data $\{d_a, d_b, d_c, d_d\}$.

The fuzzy signature in Fig. 6 is a possible Polymorphic Fuzzy Signature, which may drop $\delta$ in Eq. (6) to a low value, for the set of fuzzy signatures shown in Fig. 5.

In [29], we have investigated the aggregation of PFSs. From the results of that experiment we concluded that the weighted aggregation of PFSs are more accurate compared to non-weighted aggregation. In the next sub-section, we discuss a weighted aggregation method for PFS.

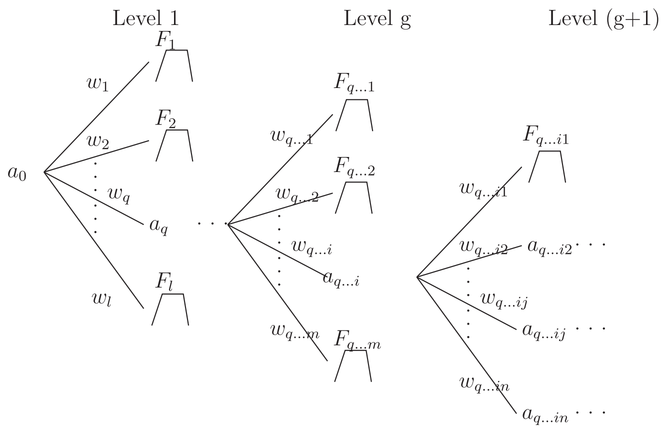


**Fig. 4.** An arbitrary polymorphic fuzzy signature (PFS).

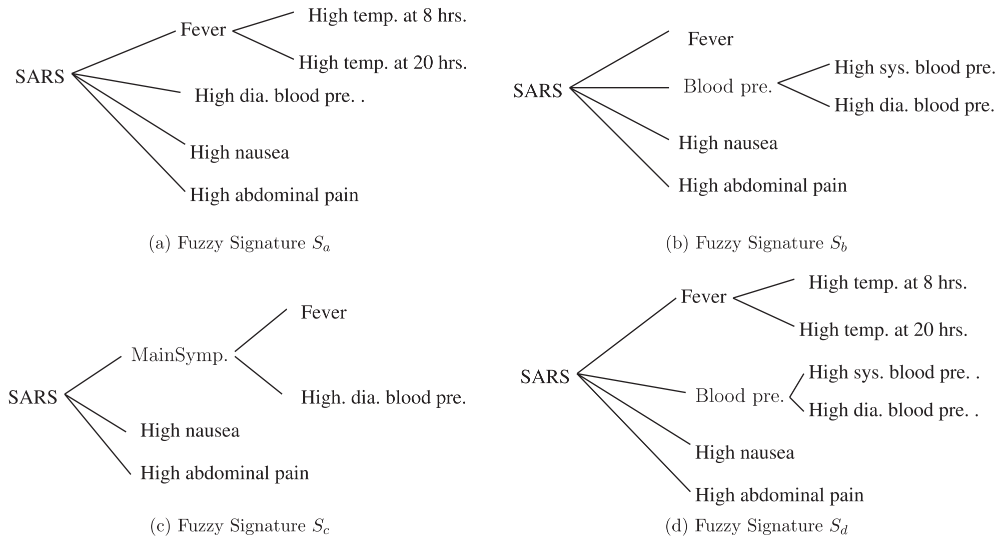


**Fig. 5.** A Set of SARS fuzzy signatures.

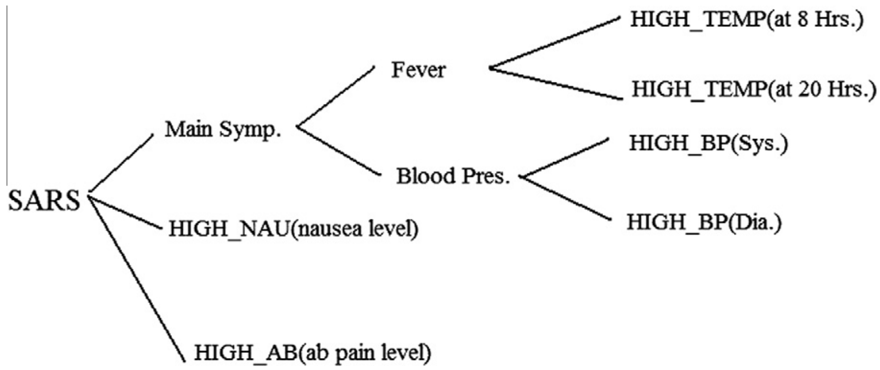


**Fig. 6.** A SARS polymorphic fuzzy signature.

### 3.2.3. Weighted Relevance Aggregation (WRA)

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. And some other branches in the same parent branch provide relatively lower values to the next level (or to the parent branch) of the fuzzy signature structure. In this way WRA enhances the accuracy of the final results of the PFS. In [33], we discussed a method of learning weights in WRA automatically. In [32], we have shown the successfulness of the weight extraction method in [33].

We further generalized the weights and the aggregation into one operator called the Weighted Relevance Aggregation Operator (WRAO) [28]. This subsection briefly describes the WRAO [28] for fuzzy signatures. In [31], 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 adding the ability to include more data points into one Polymorphic Fuzzy Signature.

Now, we recall the definition of WRAO in [28]. All the notation in the Definition 9 refer to the arbitrary fuzzy signature in Fig. 7.

**Definition 9.** The WRAO of an arbitrary branch $a_{q\ldots i}$ with n sub-branches $a_{q\ldots i1}, a_{q\ldots i2}, \ldots, a_{q\ldots in} \in [0,1]$, and weighted relevancies, $w_{q\ldots i1}, w_{q\ldots i2}, \ldots, w_{q\ldots in} \in [0,1]$, for a fuzzy signature is a function $g : [0,1]^{2n} \to [0,1]$ such that,

$$a_{q\ldots i} = \left[\frac{1}{n}\sum_{j=1}^{n}\left(a_{q\ldots ij}\cdot w_{q\ldots ij}\right)^{p_{q\ldots i}}\right]^{\frac{1}{p_{q\ldots i}}} \tag{7}$$

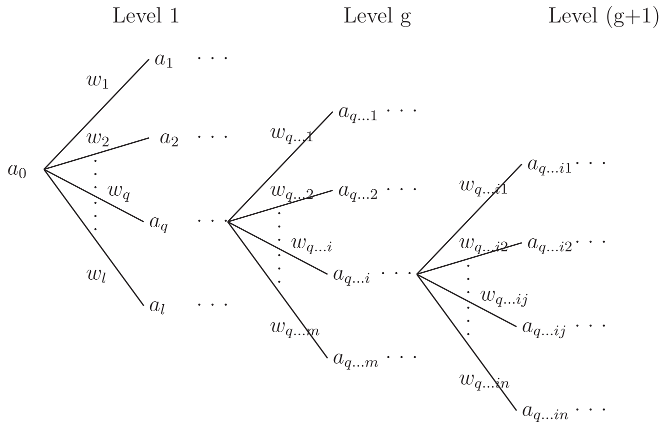


**Fig. 7.** An arbitrary fuzzy signature.

The WRAO must satisfy the following three properties,

(i) $w_{q\ldots ij} \in [0,1]$
(ii) $\bigvee_{j=1}^{n} w_{q\ldots ij} \leqslant 1$
(iii) $p_{q\ldots i} \in \Re - \{0\}$

In [28], we prove the following two properties for WRAO.

**Theorem 1.** *Let $a_{q\ldots i}$ be an arbitrary branch with n sub branches, $a_{q\ldots i1}, a_{q\ldots i2}, \ldots, a_{q\ldots in}$, and weighted relevancies, $w_{q\ldots i1}, w_{q\ldots i2}, \ldots, w_{q\ldots in}$, for an arbitrary fuzzy signature (Fig. 7). Then WRAO in Definition 9 holds the following properties.*

(i) *Partially Idempotent w.r.t $a_{q\ldots ij}$, when all $w_{q\ldots ij}$ are fixed and vice versa,*
(ii) *Commutative, and*
(iii) *Partially Monotonic w.r.t $a_{q\ldots ij}$ when all $w_{q\ldots ij}$ are fixed and vice versa.*

The partial idempotency and monotonicity is adequate to satisfy the requirement to be an aggregation function [4] as weights, $w_{q\ldots i1}, w_{q\ldots i2}, \ldots, w_{q\ldots in}$, 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 9 holds the following characteristics.*

(a) *$p_{q\ldots i} \to 0$ then WRAO → geometric mean*
(b) $\lim_{p_{q\ldots i}\to+\infty} g(a_{q\ldots i1}, \ldots a_{q\ldots in}; w_{q\ldots i1}, \ldots, w_{q\ldots in})$
$= max(a_{q\ldots i1}w_{q\ldots i1}, \ldots a_{q\ldots in}w_{q\ldots in})$
(c) $\lim_{p_{q\ldots i}\to-\infty} g(a_{q\ldots i1}, \ldots a_{q\ldots in}; w_{q\ldots i1}, \ldots, w_{q\ldots in})$
$= min(a_{q\ldots i1}w_{q\ldots i1}, \ldots a_{q\ldots in}w_{q\ldots in})$
(d) *p = 1 then WRAO → arithmetic mean*
(e) *p = – 1 then WRAO → harmonic mean*

**Example 7.** Now, we return to the student students' qualitative examination marks classification problem, from Example 3. Fig. 8 shows a PFS for the set of student data shown in Table 3. Further it shows the weighted relevancies $w_{ij}$ and aggregation factors $P_i$ of each branch. Additionally, each criterion of *Math*, *Physics*, and *Literature* are represented by fuzzy sets *High Math*, *High Physics*, and *High Literature* respectively.

Table 4, shows the normalized results of the PFS for the student evaluation. Also it shows the normalized results of Choquet Integral for the same data.

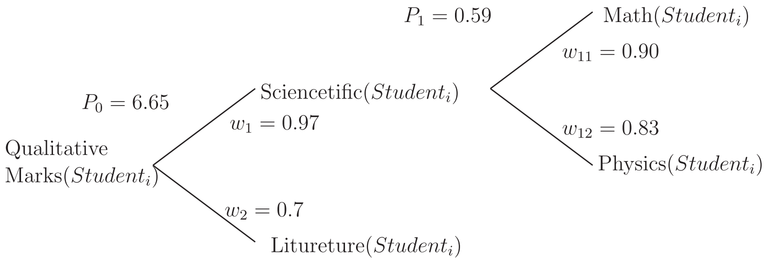


**Fig. 8.** Student PFS for qualitative marks.

**Table 4**
Qualitative grading of students using CI & FS.

| Student | Mathematics | Physics | Literature | Choquet Integral | Fuzzy signature |
|---|---|---|---|---|---|
| Euler | 18 | 16 | 10 | 0.69 | 0.69 |
| Einstein | 10 | 12 | 18 | 0.68 | 0.68 |
| Hercules | 14 | 15 | 15 | 0.74 | 0.74 |

According to the results shown in Table 4, it is clear that PFS can predict the same output as the Choquet Integral. We further extended the example and tested the both models with 200 records of students data. The results of the experiment are shown in the Fig. 9. The two plots, namely *Integral results* and *Signature results* in Fig. 9 show the results of the two methods Choquet Integral and Fuzzy Signature respectively.

Based on the results shown in Fig. 9, it can be seen that the students evaluation Fuzzy Signature can consistently approximate the results given by the Choquet Integral for the qualitative students evaluation problem. Therefore, one may come to a conclusion that there is no need for a CSDM model accompanied by fuzzy signatures, as the CSDM model together with Choquet Integral can solve the problem. In the next section, we show that the fuzzy signatures outperform the Choquet Integral in two real world problems. Further, as we discussed earlier, all Fuzzy Integrals suffer from the problem of exponential growth of fuzzy measure parameters (weights) with respect to the number of criteria. Our CSDM models have reduced space and time complexity, may need less information for decision making, and simplify the approximation of the aggregation of information compared to that of MCDM systems.

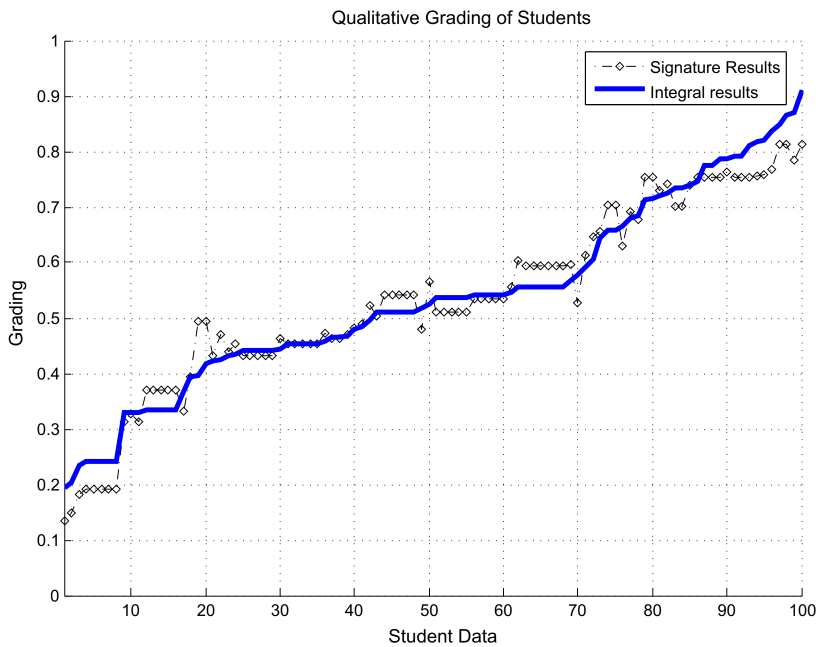


**Fig. 9.** Results of extended test with student problem.

## 4. Methods of learning WRAO and fuzzy measures

In this section we briefly explain two methods to learn WRAO and Fuzzy Measure parameters from empirical data. These two methods have been used for the experiments in the next section.

### *4.1. Levenberg–Marquardt learning of WRAO for fuzzy signatures*

In this section we discuss the method of learning WRAO from real world data briefly, with more detailed explanations to be found in [28]. First, to avoid the first two constraints on the weighted relevance factor $w_{q\dots ij}$ in Definition 9. We represents the weighted relevance factor $w_{q\dots ij}$ using the following sigmoid function:

$$w_{q\dots ij} = \frac{1}{1+e^{-\lambda_{q\dots ij}}} \tag{8}$$

where $\lambda_{q\dots ij} \in \Re$. Now, the Eq. (7) can be modified as follows,

$$a_{q\dots i} = \left[\frac{1}{n}\sum_{j=1}^{n}\left(a_{q\dots ij}\cdot\left[\frac{1}{1+e^{-\lambda_{q\dots ij}}}\right]\right)^{p_{q\dots i}}\right]^{\frac{1}{p_{q\dots i}}} \tag{9}$$

The $p_{q\dots i}$ and $\lambda_{q\dots ij}$ are called the aggregation factor of branch $q\dots i$ and the weighted relevance factor of sub-branch $q\dots ij$ of the fuzzy signature in Fig. 7, respectively. This form of WRAO (9) can be readily used for gradient based learning.

The parameters we need to learn are the aggregation factor $p_{q\dots i}$ and weighted relevance factors $\lambda_{q\dots ij}$ for each WRAO at each node of the fuzzy signature structure in Fig. 7. First we can obtain the partial derivatives of the Eq. (9) w.r.t. $p_{q\dots i}$

$$\frac{\partial a_{q\dots i}}{\partial p_{q\dots i}} = \left[\frac{a_{q\dots i}^{1-p_{q\dots i}}}{np_{q\dots i}^2}\right]\left\{\sum_{j=1}^{n} t\ln(t) - nt'\ln(t')\right\} \tag{10}$$

where $t = (a_{q\dots ij}w_{q\dots ij})^{p_{q\dots i}}$ and $t' = a_{q\dots i}^{p_{q\dots i}}$. Similarly, we obtain the partial derivatives of the Eq. (9) w.r.t. $\lambda_{q\dots ik}$

$$\frac{\partial a_{q\dots i}}{\partial \lambda_{q\dots ik}} = \left[\frac{1}{n}\sum_{j=1}^{n}(a_{q\dots ij}\cdot w_{q\dots ij})^{p_{q\dots i}}\right]^{\frac{1}{p_{q\dots i}}-1}\cdot T \tag{11}$$

where

$$w_{q\dots ij} = \frac{1}{1+e^{-\lambda_{q\dots ij}}} \quad \text{and} \quad T = \left\{\frac{d([a_{q\dots ik}\cdot w_{q\dots ik}]^{p_{q\dots i}})}{d\lambda_{q\dots ik}}\right\}$$

. We have used the Levenberg–Marquardt (LM) method [23,25] for learning 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 [23,25]. The two equations, (10) and (11) 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. A detailed discussion of the method of using LM for learning WRAO can be found in [31].

### *4.2. Automatic extraction of fuzzy measure parameters for fuzzy integrals*

Learning of fuzzy measure parameters from data has been considered by many researchers [2,24,46]. Beliakov and his team in [2,3], have shown a method of learning fuzzy measure parameters for Fuzzy Integrals from data. Further, they provided a software package, called *Aggregation Operator Approximation Tool* (AOTool [1]) and the Fuzzy Measure Tool [2] to learn Fuzzy Measure parameters for both Choquet and Sugeno Integrals. We used the Fuzzy Measure tool to learn Fuzzy Measure parameters for the Fuzzy Integrals. This sub-section will briefly discuss the method, with more detailed discussion to be found in [2,3].

**Definition 10.** The Möbius transformation of a fuzzy measure $v$ is a set function defined for every $A \subseteq N$

$$M(A) = \sum_{B\subseteq A}(-1)^{|A\setminus B|}v(B) \tag{12}$$

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

$$a(A) = \sum_{B\subseteq A} M(B), \quad \text{where } \forall A \subseteq N \tag{13}$$

The Choquet Integral in Eq. (2) can be represented in an alternative way using the Möbius transformation in Eq. (12) [16,2]

$$C_v(f) = \sum_{A \subseteq N} M(A) h_A(x) \tag{14}$$

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

As we mentioned earlier, the computational complexity of fuzzy measures are exponential and it becomes too large when the number of inputs goes above 7. Grabisch in [16] 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 in [16] 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 [2].

**Definition 11.** A fuzzy measure $v$ is called $k$-additive $(1 \leqslant k \leqslant n)$ if its Möbius transformation verifies $M(A) = 0$ for any subset $A$ with more than $k$ elements, $\mid A \mid \geqslant k$, 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

$$\text{minimize } \|f(X_i) - y_i\| \tag{15}$$

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 [34]

$$\text{minimize } \|C_v(x_{1i}, \ldots, x_{ni}) - y_i\| \tag{16}$$

Now according to Beliakov [2], using Definitions 10 and 11, Eq. (16) can be translated into the following constrained optimization problem

$$\text{minimize } \| \sum_{A\ :|A| \leqslant k} h_A(x_i) m_A - y_i \| \tag{17}$$

such that the following constraints are satisfied

$$\sum_{B \subseteq A : j \in B, |B| \leqslant k} m_B \geqslant 0$$
$$\forall A \subseteq N, |A| \geqslant 1, \quad \text{and all } j \in A$$
$$m_j \geqslant 0, \quad j = 1, \ldots, n,$$
$$\sum_{B \subseteq N || |B| \leqslant k} m_B = 1$$

It appears that in AOTool, [1,2], the minimization problem in the expression (17) is further converted to a linear programming problem and the $l_1$ norm is used to calculate the error.

## 5. Two experiments: two fuzzy integrals & fuzzy signatures

In this section, two real world problems, namely High Salary Selection of employees and SARS patient classification problem are used, applying Choquet and Sugeno Integrals and Polymorphic Fuzzy Signature (PFS). The aim of the experiments are to investigate and demonstrate if possible the success of the CSDM method and the effectiveness of hierarchical fuzzy signatures in decision making.

### *5.1. High salary selection problem*

We select the High Salary Selection problem, which is discussed in [12], as the first experiment. The problem is to find the degree of relevance of a high salary based on the *contacts*, *age*, and *work experience* of an employee. Fig. 10 shows a High Salary Selection Polymorphic Fuzzy Signature, which is obtained using domain expert knowledge, for the high salary selection problem. Note that $@_i$ and $w_i$ in Fig. 10 represent the aggregation function and weighted relevance of the node $i$, respectively.

In general, Choquet and Sugeno Integrals take normalized inputs and produce the output in the same range, which is a rank between [0, 1] of likelihood of getting a high salary. On the other hand, Fuzzy Signatures take fuzzified 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 experiment. That is to decide which output target among the two output distributions, that is normalized output and fuzzyfied output, is best for learning and testing of Fuzzy Integral based systems and Fuzzy Signature based system. Another 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 three times more than the number of normalized input dimensions.

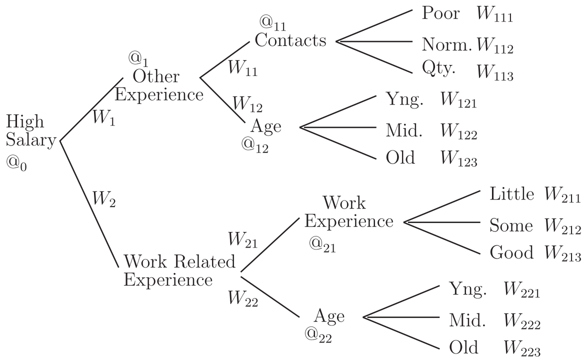


**Fig. 10.** High salary selection PFS.

Also, it is clear 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 high salary fuzzy set, from the normalized distribution of output data. For the comprehensibility of this experiment, we decided to use Choquet Integral and Fuzzy Signatures both with normal and fuzzyfied output targets to discover which output target is suitable for each method.

We used the two learning methods explained in the previous section to extract fuzzy measures and WRAO for the Fuzzy Integrals and Polymorphic Fuzzy Signature respectively for the High Salary Selection problem. The Table 5 shows the results of the experiment for all four possibilities, that is Choquet Integral and Fuzzy Signatures with both normal and fuzzyfied output targets. According to the table the Choquet Integral shows best results with the normalized outputs and the High Salary Selection PFS shows best results with the fuzzyfied outputs as expected. Therefore, the normalized output data has been selected for the experiment with Fuzzy Integrals and Fuzzified output has been selected for experiments with Polymorphic Fuzzy Signatures in this section.

The next practical issue arises when we need to compare the results. Figs. 11 and 12, and the mean squared error (MSE) shown in Table 5 experiment may illustrate the comparison of two methods. But this also does not express the situation clearly, as the values of numerical error such as MSE of the results of two methods depend on the desired output distributions given to each method. Therefore we measure classification error (see Section 5.1.1) to compare and visualize the difference of the error of the results of the two methods clearly. The next subsection briefly explains the form of the calculation of classification error we used.

### 5.1.1. Classification error

We formulate the Classification Error (CLE) in the following way. 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 visualized in the following Fig. 13.

According to Fig. 13, there are 3 categories of classifications that can occur, they are *Good*, *Bad*, and *Very Bad*. Now we assume the pair of *predicted* and *desired* values, of the *i*th input, respectively taken as *X* and *Y* coordinates of the point $P_i$ on the two dimensional classification error rule space Fig. 13. The classification error of an arbitrary point $P_i$ can be written as

$$CLE(P_i) = \begin{cases} 0 & \text{if} \quad P_i \in \text{Good} \\ 0.5 & \text{if} \quad P_i \in \text{Bad} \\ 1 & \text{if} \quad P_i \in \text{Very Bad} \end{cases}$$

Let us consider the 4 straight lines, $B1$, $B2$, $G1$, and $G2$, in Fig. 13. In this experiment, they are equivalent to,

**Table 5**
High salary selection experiment.

| | Fuzzy output | | Crisp output | |
|---|---|---|---|---|
| | MSE train | MSE test | MSE train | MSE test |
| Choquet | 0.0837 | 0.1052 | **0.0266** | **0.0262** |
| PFS | **0.0149** | **0.0152** | 0.0216 | 0.0195 |

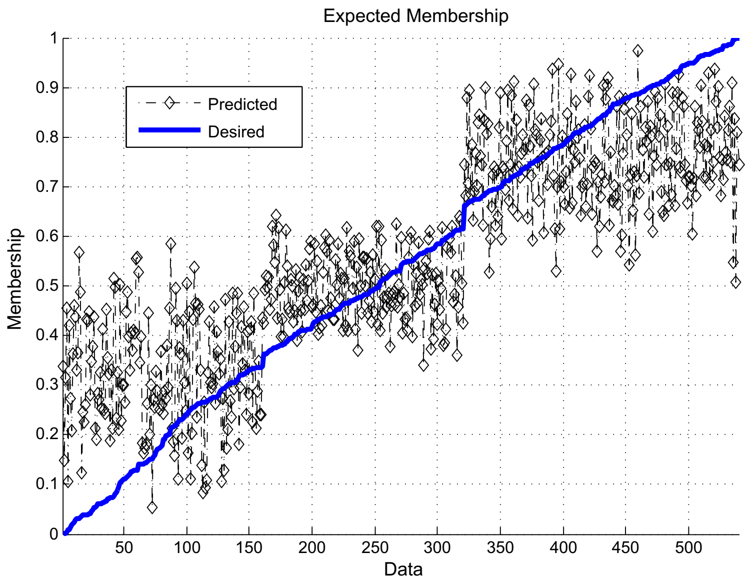


**Fig. 11.** Training results of salary experiment: Choquet integral.

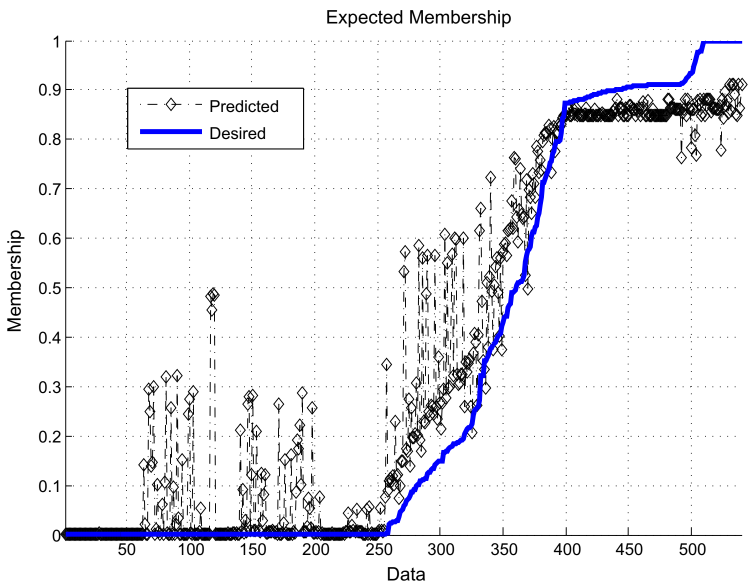


**Fig. 12.** Training results of salary experiment: PFS.

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

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

$$CLE(P_i) = \begin{cases} 0 & \text{if} \quad G1(P_i) \leqslant 0 \quad \text{and} \quad G2(P_i) \geqslant 0 \\ 0.5 & \text{if} \quad (B1(P_i) \leqslant 0 \quad \text{and} \quad G1(P_i) > 0) \text{ or} \\ & \quad\;\;\; (G2(P_i) < 0 \quad \text{and} \quad B1(P_i) \geqslant 0) \\ 1 & \text{if} \quad B1(P_i) > 0 \quad \text{or} \quad B2(P_i) < 0 \end{cases}$$

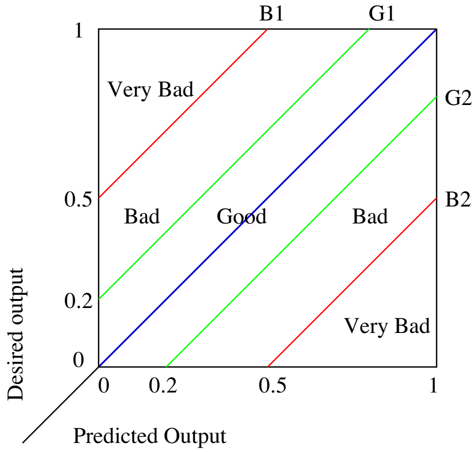


**Fig. 13.** Classification error rules.

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

$$SCLE(P) = \sum_{i=1}^{m} CLE(P_i) \text{ where } m \in \mathbb{N} \tag{18}$$

Now, we can use the classification error to visualize and classify the results of the experiments. Let us continue experiment 1, Table 6 shows the MSE and SCLE for both training and testing phases of the experiment with all four possibilities we discussed earlier. Figs. 14 and 15 illustrate the classification error of the best test results for Choquet Integral and High Salary

**Table 6**
High salary selection experiment.

| | MSE train | SCLE train | MSE test | SCLE test |
|---|---|---|---|---|
| Choquet | 0.0266 | 54.5 | 0.0262 | 55.5 |
| Sugeno | 0.0283 | 61.5 | 0.0274 | 55.5 |
| PFS | **0.0149** | **23.5** | **0.0152** | **20** |

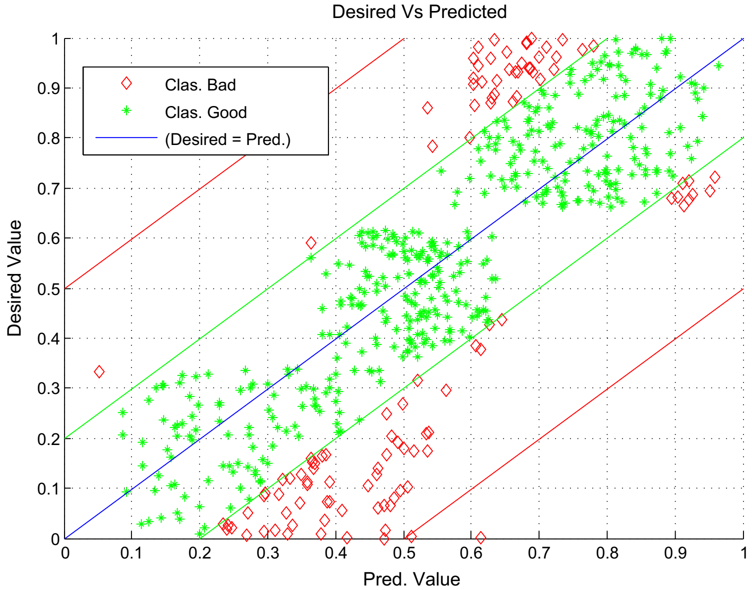


**Fig. 14.** Test classification error salary: Choquet integral.

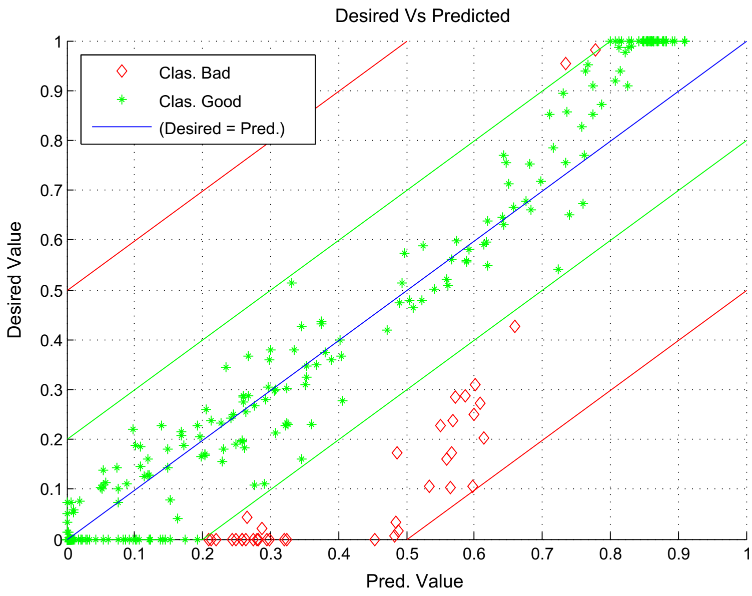


**Fig. 15.** Test classification error salary: PFS.

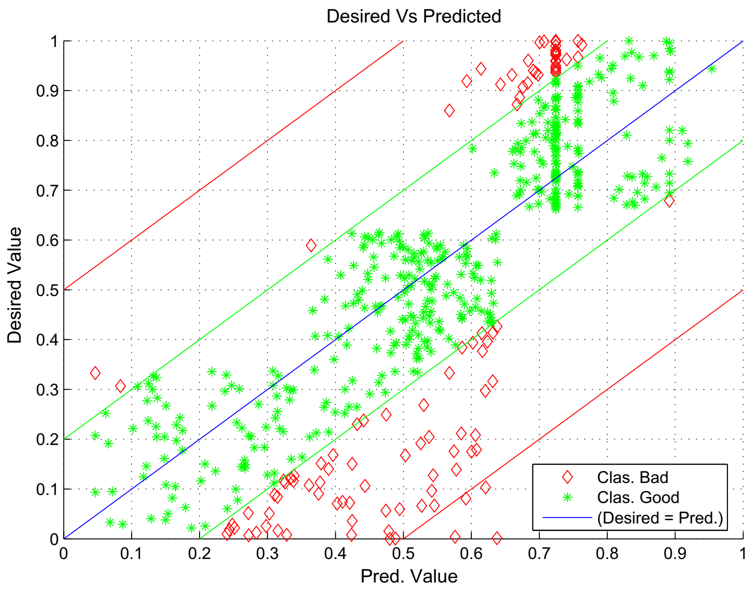


**Fig. 16.** Test classification error salary: Sugeno integral.

Selection PFS respectively. Additionally, Fig. 16 illustrates the testing results of the Sugeno Integral with normalized data, which is best for Fuzzy Integrals.

The results of this experiment show that Fuzzy Integrals can learn and classify the data in High Salary Selection problem. But, High Salary Selection PFS has a much lower sum of classification error (SCLE) and lower mean squared error (MSE) as shown in Table 6, compared to that of the Fuzzy Integrals. Further, the PFS model has reduced SCLE in the testing phase compared to that of the training phase, which importantly shows that PFS training is generalizing and avoiding over fitting problems as well.

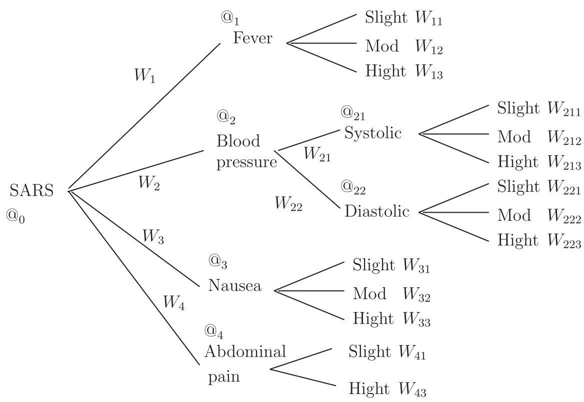


**Fig. 17.** SARS Patient classification PFS.

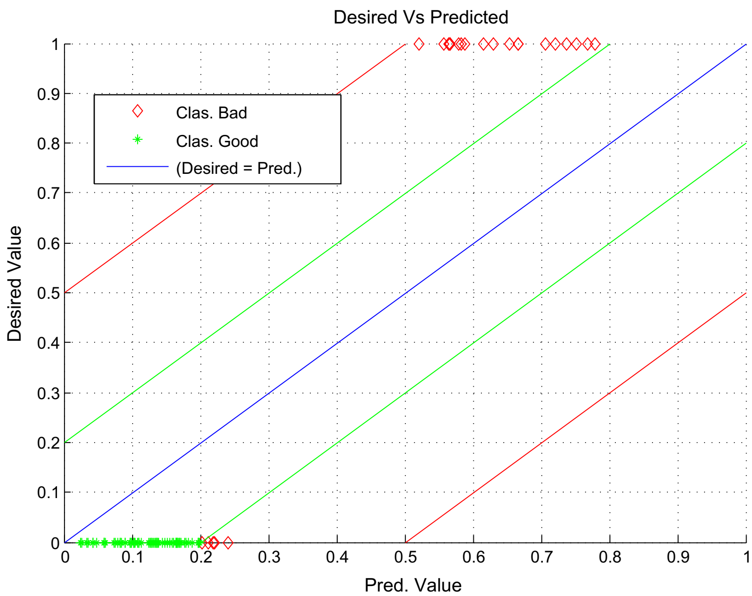


**Fig. 18.** Training predicted vs desired SARS: Choquet integral.

## 5.2. SARS patient classification problem

In this experiment we use SARS patient classification problem to compare the performance of the two methods, Fuzzy Integrals and PFS.

Medical practitioners know that for certain diseases, such as SARS, they need to check the patient for possible fever, hypertension, conditions of nausea, and abdominal pain [20,35,44]. In addition, it is fairly important to monitor the fever regularly during the day, as well as the blood pressure. Fig. 17 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 the two measurements of blood pressure, *slight*, *medium*, and *high* for nausea, and *slight*, and *high* for abdominal pain. The SARS fuzzy signature contains three levels of hierarchies, which can be aggregated using different aggregations.

In Fig. 17, the notations $a_{ij}$, $@_{ij}$, and $w_{ij}$ represent the input value, aggregation function, and weight for the branch *ij* of the SARS PFS. 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 a full degree of confidence, i.e. 1, for the SARS patient data and zero degree of confidence for the other condition, and normal data.

Figs. 18–20 show the training results of Choquet and Sugeno Integrals and PFS. Figs. 21–23 show the test results of the two methods. Table 7 shows the MSE and SCLE for training and testing phases of these methods.

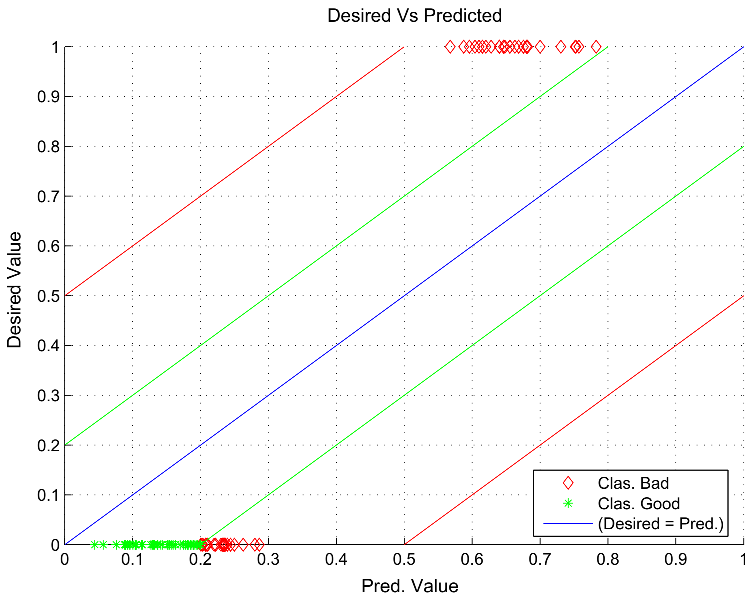


**Fig. 19.** Training predicted vs desired SARS: Sugeno integral.

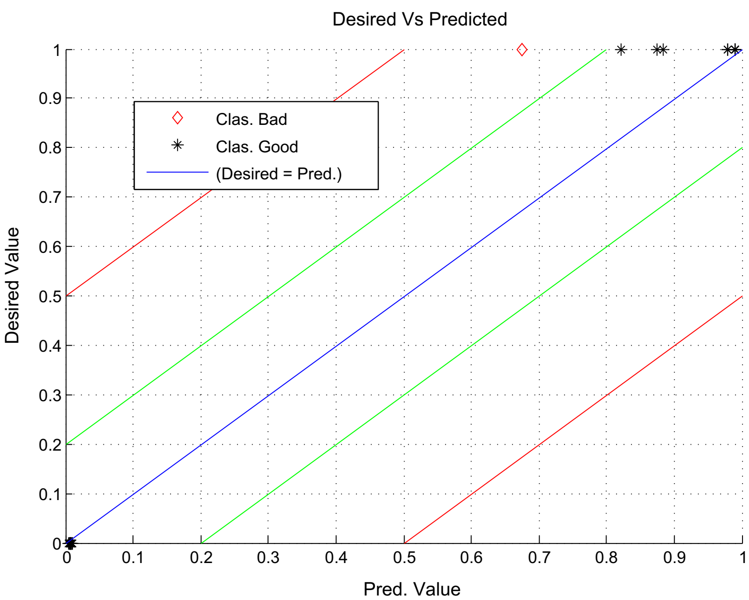


**Fig. 20.** Training predicted vs desired SARS: PFS.

The results of the second experiment show that Fuzzy Integrals can not learn or classify the SARS patients data properly. The SARS patients classification PFS has very much lower SCLE and MSE as shown in Table 7, compared to that of the Fuzzy Integrals. Further, in this experiment, we also observed that the PFS model has reduced SCLE in the testing phase compared to that of the training phase.

From the results of the two real world experiments, it is clear that our PFS outperforms the Fuzzy Integrals in both experiments. In the next sub-section, we discuss the reasons for the improved performance of the Fuzzy Signatures against the Fuzzy Integrals.

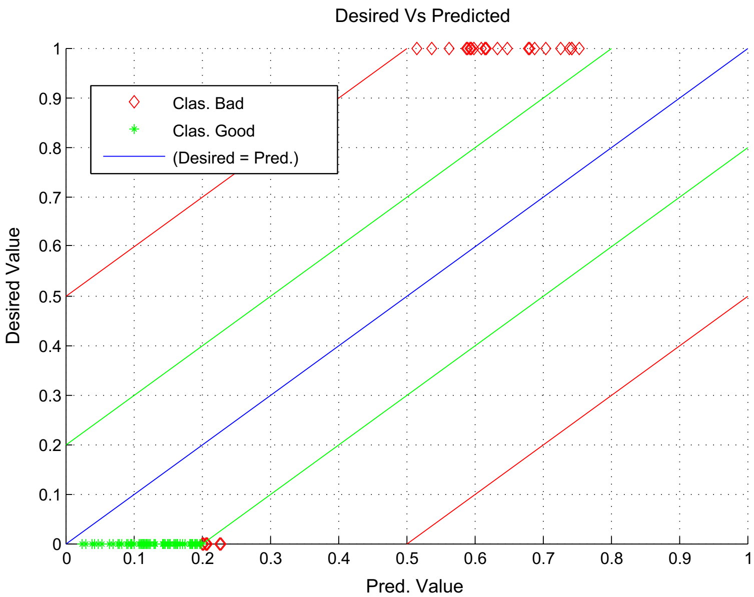


**Fig. 21.** Test predicted vs desired SARS: Choquet integral.

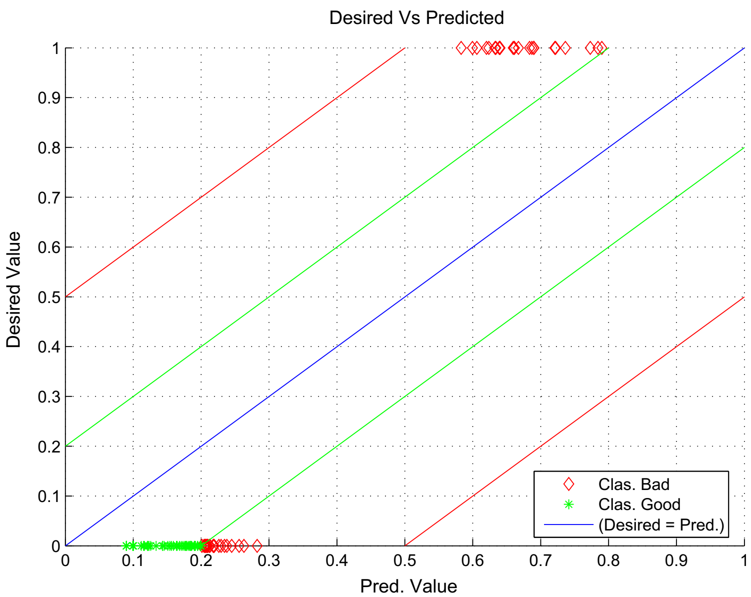


**Fig. 22.** Testing predicted vs desired SARS: Sugeno integral.

### *5.3. Discussion of experiments*

As we pointed out earlier, a CSDM system, such as Fuzzy Signatures, uses a set of hierarchically organized local aggregation functions to approximate the desired global preference function of the system. On the other hand, flat decision making systems such as Fuzzy Integrals try to approximate the global preference relation using one aggregation function. In this paper, we argued that the approximation of a global preference relation using a set of local aggregation functions is necessarily less complex. Further, the approximation of the same global preference relation using a single aggregation function is much harder. Thus this is an advantage for CSDM systems over MCDM systems. This argument is supported by the success of the Fuzzy Signature (CSDM method) based system over the Fuzzy Integral (MCDM method) based systems in the two real world experiments. Also, we have shown that for the students evaluation problem, Fuzzy Signatures can predict the same results as Choquet Integral.

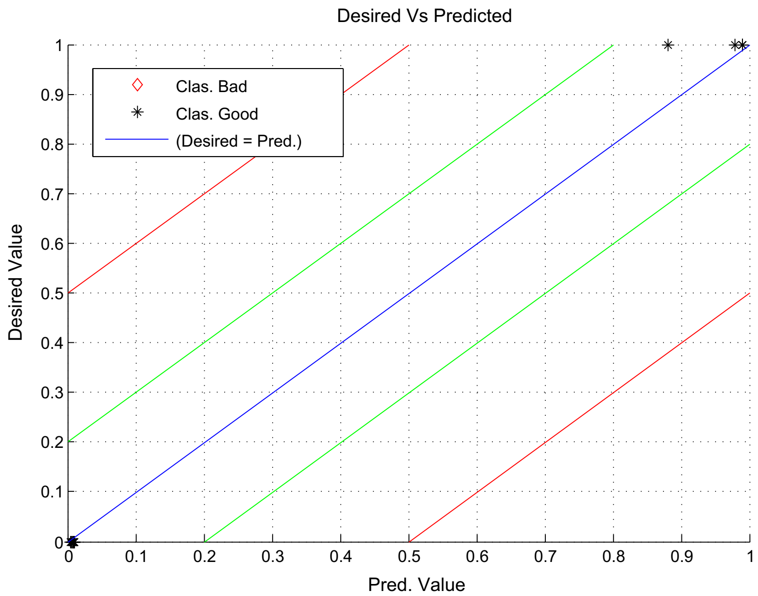


**Fig. 23.** Test predicted vs desired SARS: PFS.

**Table 7**
SARS Patient classification experiment.

| | MSE train | SCLE train | MSE test | SCLE test |
|---|---|---|---|---|
| Integral | 0.04253 | 13.5 | 0.04948 | 16.5 |
| PFS | **0.0017** | **0.5** | **0.0001** | **0** |

The WRAO has been derived similarly to the form of the generalized weighted means function discussed in [11] in order to satisfy the Weighted Relevance Aggregation concept discussed in [28]. The WRAO is a more generalized version of weighted mean as it has weaker constraints on weights, that is $\bigvee_{j=1}^{n} w_{q\ldots ij} = 1$, Definition 9, compared to needing to sum to 1 in the generalized weighted mean [11]. Grabisch in [14] has shown that the generalized weighted mean is a special class of Choquet Integral when the Fuzzy Measure used by Choquet Integral is additive and the aggregation factor $p > 0$. Further, Grabisch suggested that geometric mean, $p \to 0$ and the harmonic mean, $p \to -1$ are not Fuzzy Integrals. An implication is that WRAO also has a subset, which is a special case of Choquet Integral. The authors are now working in this direction to show the relation between WRAO and Choquet Integral mathematically.

Further, an interesting question is to investigate why WRAO has performed better than the Choquet Integral (Fuzzy Integrals). The first reason may be because WRAO may represent only a special class of Choquet Integral where this subset can represent the quasi-optimization problem better. That is, it gives a smoother optimization surface towards the best local minimum for the learning algorithm. On the other hand, a second reason may be that WRAO is not a proper subset of Choquet Integral and uses a class of aggregation functions that is outside the Choquet Integral to outperform it, again by providing a smoother optimization surface.

Here we discussed two major reasons to explain the performance of Fuzzy Signatures accompanied by the WRAO against the Choquet Integral. Both reasons point to new research directions, towards the development and understanding of the role of aggregation functions in intelligent decision making systems.

## 6. Conclusion

We discussed the MCDM paradigm and the role of Fuzzy Integrals in developing intelligent decision making systems. Then, we introduced the CSDM model in order to construct 'human like' hierarchical decision making systems. The qualitative student evaluation example, which is common in the MCDM literature, has been used to illustrates the difference between MCDM and CSDM models. Finally, using two real world experiments, namely High Salary Selection of employees and SARS patients classification problems, the performance of the Choquet Integral and Fuzzy Signatures has been compared. The results of the two experiments show that Fuzzy Signatures together with WRAO outperformed the Fuzzy Integrals. Our conclusion is that CSDM systems simplify the approximation of the aggregation functions to achieve the global preference relation of the problem, and WRAO represents a class of aggregation functions which simplify the learning and provide more adaptation to the problem.

## References


[1] G. Beliakov, Aggregation Operator Approximation Tool (aotool), Technical report <http://www.deakin.edu.au/gleb/>.
[2] G. Beliakov, Programming Library fmtools for Handling and Fitting Fuzzy Measures, Technical report, 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, J. Warren, Appropriate choice of aggregation operators in fuzzy decision support systems, IEEE Transactions on Fuzzy Systems 9 (6) (2001) 773–784.
[5] R. Bellman, L. Zadeh, Decision-making in a fuzzy environment, Management Science 17 (4) (1970) 141–164.
[6] G. Bykzkan, D. Ruan, Choquet integral based aggregation approach to software development risk assessment, Information Sciences 180 (3) (2010) 441–451.
[7] A.R. de Soto, A hierarchical model of a linguistic variable, Information Sciences 181 (20) (2011) 4394–4408. special Issue on Interpretable Fuzzy Systems.
[8] M.R. Delgado, F.V. Zuben, F. Gomide, Hierarchical genetic fuzzy systems, Information Sciences 136 (1-4) (2001) 29–52.
[9] D. Dubois, H. Fargier, P. Perny, Qualitative decision theory with preference relations and comparative uncertainty: an axiomatic approach, Artificial Intelligence 148 (1–2) (2003) 219–260.
[10] D. Dubois, J. Marichal, H. Prade, M. Roubens, R. Sabbadin, The use of the discrete Sugeno integral in decision-making: a survey, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 9 (5) (2001) 539–561.
[11] H. Dyckhoff, W. Pedrycz, Generalized means as model of compensative connectives, Fuzzy Sets and Systems 14 (2) (1984) 143–154.
[12] T. Gedeon, K. Wong, D. Tikk, Constructing hierarchical fuzzy rule bases for classification, The 10th IEEE International Conference on Fuzzy Systems, 3, 2001.
[13] M. Grabisch, Fuzzy measures and integrals: a survey of applications and recent issues, Dubois et al., 3, 1996, pp. 507–529.
[14] M. Grabisch, Fuzzy integral in multicriteria decision making, Fuzzy Sets and Systems 69 (3) (1995) 279–298.
[15] M. Grabisch, The application of fuzzy integrals in multicriteria decision making, European Journal of Operational Research 89 (3) (1996) 445–456.
[16] M. Grabisch, k-order additive discrete fuzzy measures and their representation, Fuzzy Sets and Systems 92 (2) (1997) 167–189.
[17] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap, Aggregation functions: means, Information Sciences 181 (1) (2011) 1–22.
[18] C. Guo, D. Zhang, On set-valued fuzzy measures, Information Sciences 160 (1-4) (2004) 13–25.
[19] W. Ho, X. Xu, P. Dey, Multi-criteria decision making approaches for supplier evaluation and selection: a literature review, European Journal of Operational Research 202 (1) (2010) 16–24.
[20] 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) 713–717.
[21] L. Kóczy, T. Vámos, G. Biró, Fuzzy signatures, in: EUROFUSE-SIC '99, 1999.
[22] L.T. Kóczy, K. Hirota, Approximative inference in hierarchical structured rule bases, in: 5th IFSA World Congress, pp. 1262–1265.
[23] K. Levenberg, A method for the solution of certain nonlinear problems in least squares, Quarterly of Applied Mathamatics 2 (1944) 164–168.
[24] X. Liu, L. Ma, J. Mathew, Machinery fault diagnosis based on fuzzy measure and fuzzy integral data fusion techniques, Mechanical Systems and Signal Processing 23 (3) (2009) 690–700.
[25] 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.
[26] B.S.U. Mendis, T.D. Gedeon, Aggregation selection for hierarchical fuzzy signatures: a comparison of hierarchical OWA and WRAO, in: International Conference of Information Processing and Management of Uncertainty in Knowledge Based Systems (IPMU), 2008, pp. 1–8.
[27] B.S.U. Mendis, T.D. Gedeon, Polymorphic fuzzy signatures, in: FUZZ-IEEE 2010 International Conference on Fuzzy Systems, pp. 1–7, 2010.
[28] B.S.U. Mendis, T.D. Gedeon, J. Botzheim, L.T. Kóczy, Generalised weighted relevance aggregation operators for hierarchical fuzzy signatures, in: International Conference on Computational Intelligence for Modelling Control and Automation and International Conference on Intelligent Agents Web Technologies and International Commerce (CIMCA'06), Sydney, Australia, 2006.
[29] 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.
[30] 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.
[31] 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.
[32] 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.
[33] 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.
[34] 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.
[35] 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.
[36] J. Quinlan, C4. 5: Programs for Machine Learning, Morgan Kaufmann, 1993.
[37] M. Sugeno, Theory of Fuzzy Integrals and its Applications, Ph.D. thesis, Tokyo Institute of Technology, 1974.
[38] M. Sugeno, Fuzzy measures and fuzzy integrals: a survey, Fuzzy Automata and Decision Processes (1977) 89–102.
[39] B. Tripathy, A. Baruah, New type of difference sequence spaces of fuzzy real numbers, Mathematical Modelling and Analysis 14 (3) (2009) 391–397.
[40] B. Tripathy, A. Baruah, Nrlund and riesz mean of sequences of fuzzy real numbers, Applied Mathematics Letters 23 (2010) 651–655.
[41] B. Tripathy, A.J. Dutta, On fuzzy real-valued double sequence spaces, Mathematical and Computer Modelling 46 (9-10) (2007) 1294–1299.
[42] B. Tripathy, A.J. Dutta, Bounded variation double sequence space of fuzzy real numbers, Computers and Mathematics with Applications 59 (2) (2010) 1031–1037.
[43] B. Tripathy, B. Sarma, Sequence spaces of fuzzy real numbers defined by orlicz functionss, Mathematica Slovaca 58 (5) (2008) 621–628.
[44] 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).
[45] Z. Turskis, E. Zavadskas, F. Peldschus, Multi-criteria optimization system for decision making in construction design and management, Inzinerine Ekonomika – Engineering Economics 1 (61) (2009) 7–17.
[46] X.-Z. Wang, Y.-L. He, L.-C. Dong, H.-Y. Zhao, Particle swarm optimization for determining fuzzy measures from data, Information Sciences 181 (19) (2011) 4230–4252.

[47] K. Wong, T. Gedeon, A. Chong, L. Kóczy, Fuzzy signatures for complex decision model, Australian Journal of Intelligent Information Processing Systems 8 (4) (2005) 193–199.
[48] K. Wong, T. Gedeon, L. Kóczy, Construction of fuzzy signature from data: an example of sars pre-clinical diagnosis system, in: Proceedings 2004 IEEE International Conference on Fuzzy Systems, 3, 2004.
[49] Z. Xu, Choquet integrals of weighted intuitionistic fuzzy information, Information Sciences 180 (5) (2010) 726–736.
[50] R. Yager, On ordered weighted averaging aggregation operators in multicriteria decision making, IEEE Transactions on Systems, Man and Cybernetics 18 (1) (1988) 183–190.