# Using Special Structured Fuzzy Measure to Represent Interaction Among IF-THEN Rules

Xi-Zhao Wang and Jun Shen

Machine Learning Center, Faculty of Mathematics and Computer Science,
Hebei University, Baoding, 071002, China
wangxz@mail.hbu.edu.cn, shenjun79225@126.com

**Abstract.** When fuzzy IF-THEN rules are used to approximate reasoning, interaction exists among rules. Handling the interaction based on a non-integral can lead to an improvement of reasoning accuracy but the determination of non-linear integral usually needs to solve a linear programming problem with too many parameters when the rules are a little many. That is, the number of parameters increases exponentially with the number of rules. This paper proposes a new approach to denoting the interaction by a 2-additive fuzzy measure which replaces the general set function of the old non-linear integral approach. The number of parameters determined in the new approach is greatly less than the number of parameters in the old approach. Compared with the old approach, the new one has a little loss of accuracy but the new approach reduces the number of parameters from an exponential to polynomial quantity.

## 1 Introduction

Fuzzy IF-THEN rules are widely used in expert systems to represent fuzzy and uncertain concepts. Given a set of FPRs and an observed fact, FPR reasoning is used to draw an approximate conclusion by matching the observed fact against the set of FPRs. Many researchers have investigated this fundamental issue in fuzzy reasoning ([1-6]).

It is important to find the interaction among rules. Interaction can help domain experts discover new knowledge existing among rules. With respect to a given consequent, knowing and modelling the enhancing or resisting effect among rules learned from data is helpful to maintain the rule base. By discovering the interaction among the rules and then applying it to fuzzy reasoning, it is expected to improve the reasoning accuracy. In [1], we can see the WFPR (Weighted Fuzzy Production Rule) fails to apply to such a situation that interaction exists among the rules. In order to handle this situation, the authors propose to use a non-linear integral tool. The interaction among the rules is considered as the non-additive set function, and the classification is computed by using the integral model. Such a handling of interaction in fuzzy IF-THEN rule reasoning can lead to a well understanding of the rule base, and also can lead to an improvement of reasoning accuracy.

Reference [1] investigates how to determine from the given data the non-additive set function which cannot be specified by domain experts. The main problem of the

D.S. Yeung et al. (Eds.): ICMLC 2005, LNAI 3930, pp. 459 – 466, 2006.

proposed approach to interaction handling is too many parameters when the number of the rules increases. The approach is not appropriate in many situations due to the exponentially increasing complexity. In order to solving this problem, we could consider the special structured fuzzy measure to replace the general set function. There are many fuzzy measures that have the special structure, for example, the belief fuzzy measure, the plausibility fuzzy measure, $\lambda -$ fuzzy measure, 2-additive fuzzy measure and so on. In [8, 9], the authors once used genetic algorithms to determine the $\lambda -$ fuzzy measure. In this paper, we propose using 2-additive fuzzy measure to represent the interaction among the rules such that the representation parameters can be reduced greatly. The parameter reduction is at the cost of some loss of accuracy.

## 2 Background on Fuzzy Measures and Integrals

### 2.1 Fuzzy Measures

**Definition 1.** A fuzzy measure $\mu$ defined on $X$ is a set function $\mu : p(\mathrm{X}) \rightarrow [0,1]$, satisfying the following axioms:

$$(1)\ \mu(\varnothing)=0, \mu(\mathrm{X})=1$$

$$(2)\ \mathrm{A} \subseteq \mathrm{B} \Rightarrow \mu(\mathrm{A}) \leq \mu(\mathrm{B})$$

A fuzzy measure can be additive, supper-additive, or sub-additive. Let $X=\{R_1,R_2,\cdots,R_n\}$ be a set of rules with the same consequent. We regard A, B as two subsets of rules. $\mu(A)$ is the importance of the subset $A$ .The additive means that there is no interaction among two subsets; the sub-additive and super-additive mean that there exists interaction. The sub-additive indicates that the two sets of rules are resisting each other; super-additive means that the two sets of rules are enhancing each other.

### 2.2 $k$ -Additive Fuzzy Measures

**Definition 2.** A Pseudo-Boolean function is a real valued function $f : \{0,1\}^n \rightarrow \Re$. A fuzzy measure can be viewed as a particular case of the pseudo-Boolean function, defined for any $A \subset X$ , such that $A$ is equivalent to a $\{x_1, x_2, ..., x_l\}$ in $\{0,1\}^l$ , where $x_i = 1$ if $i \in \mathrm{A}$ . It can be shown that any pseudo-Boolean can be expressed as a multi-linear polynomial in L variables, that is $f(x)=\sum_{T\subset L}\left\{a(T)\prod_{i\in T}x_i\right\}$ with $a(T) \in R$ and $x \in \{0,1\}^l$ . The coefficients $a(T)$, $T \subset X$ can be interpreted as the Mobius transform of a set function. We note $\mu_i = a_i$ , $\mu_i = \mu(\{i\}), a_i = a(\{i\})$.

**Definition 3.** A fuzzy measure $\mu$ defined on $X$ is said to be $k$ -order additive if its corresponding Pseudo-Boolean function is a multi-linear polynomial of degree $k$ , i.e., $a(T)=0$, $\forall\, T$ such that $|T|>K$, and there exits at least one $T$ of $k$ elements such that $a(T)\neq 0$.

For any $K\subset X$ and $|K|\geq 2$, with $x_i=1$ if $i\in K$, $x_i=0$ otherwise. The 2-additive fuzzy measure is defined by: $\mu(K)=\sum_{i=1}^{n}a_i x_i+\sum_{\{i,j\}\subseteq X}a_{ij}x_i x_j$. The fact that $\mu_i=a_i$ for all $i$, we can get the following expression:

$$\mu_{ij}=\mu(\{x_i,x_j\})=a_i+a_j+a_{ij}=\mu_i+\mu_j+a_{ij}.$$

The general formula for 2-additive fuzzy measure is

$$\mu(K)=\sum_{i\in K}a_i+\sum_{\{i,j\}\subseteq K}a_{ij}=\sum_{\{i,j\}\subseteq K}\mu_{ij}-(|K|-2)\sum_{i\in K}\mu_i$$

for any $K\subset X$ and $|K|\geq 2$. For example

$$\mu_{ijk}=\mu(\{x_i,x_j,x_k\})=\mu_i+\mu_j+\mu_k+a_{ij}+a_{ik}+a_{jk}.$$

The 2-additive fuzzy measure is determined by the coefficients $\mu_i$ and $\mu_{ij}$, only $L(L+1)/2$ coefficients $\mu_i$ and $\mu_{ij}$ have to be determined from training data. The coefficients for all other subsets $K\subset X$ and $|K|\geq 2$ are calculated from $\mu_i$ and $\mu_{ij}$. In order to obtain a monotonic fuzzy measure, the coefficients $\mu_i$ and $\mu_{ij}$ must satisfy particular conditions. The monotonicity which constraints on the coefficients of the 2-additive fuzzy measure can be formulated as follows:

$$\sum_{j\in K}\mu_{ij}-\sum_{j\in K}\mu_j-(L-2)\mu_i\geq 0,\forall i\in X,K\subseteq X\setminus\{i\} \qquad (1)$$

where $|X|=L$, to obtain a fuzzy measure normalized to the interval $[0,1]$, the coefficients $\mu_i$ and $\mu_{ij}$ must also satisfy the normalization condition, for $k=2$, we have $\mu(X)=\sum_{i\in X}\mu_i+\sum_{\{i,j\}\subseteq X}a_{ij}=1$.

In order to ensure the monotonicity and normalization, 2-additive fuzzy measure satisfies constrains:

$$a(\varnothing)=0,$$

$$\sum_{i\in L}a_i+\sum_{\{i,j\}\subseteq L}a_{ij}=1,\ a_i\geq 0,\forall i\in L,$$

$$a_i+\sum_{j\in T}a_{ij}\geq 0,\forall i\in L,\forall T\subseteq L\setminus\{i\}.$$

The concept of 2-additive fuzzy measure provides a trade-off between richness and complexity of fuzzy measure.

### 2.3 Fuzzy Integral

Fuzzy integral is a type of integrals of a real function with respect to a fuzzy measure. There have been several definitions of fuzzy integrals. We restrict our discussion here to the Choquet integral.

**Definition 4.** Let $X = \{x_1, x_2, ..., x_n\}$, $\mu$ be a fuzzy measure defined on the power of $X$, $f$ be a function from $X$ to $[0,1]$. The Choquet integral of $f$ with respect to $\mu$ is defined by

$$(C)\int_X f d_\mu = C_\mu\left(f(x_1),...,f(x_n)\right) = \sum_{i=1}^{n}\left(f\left(x_{(i)}\right) - f\left(x_{(i-1)}\right)\right)\mu(A_i)$$

where the subscript $(i)$ indicates that the indices have been permuted so that $0 \le f\left(x_{(1)}\right) \le ... f\left(x_{(n)}\right) \le 1$, also $A_i = \left(x_{(1)}, ..., x_{(n)}\right)$, and $f\left(x_0\right) = 0$.

## 3 Using 2-Additive Fuzzy Measure to Represent Interaction Among Rules

Yeung et al. determined the set function by solving a linear programming problem in [1]. The main problem of the proposed approach to interaction is too many parameters when the number of rules increases. At this stage, the approach is not yet appropriate for many cases due to the exponential complexity. Actually, it is not feasible to implement in the real world. In this paper we propose to use 2-additive fuzzy measure to replace the general set function for the interaction handling. In [1], the number of the parameters of a set function is $2^m + 2^{M-m} - 2$. If we use 2-additive fuzzy measure instead of the set function, we only need to determine $m \times (m+1)/2 + (M-m) \times (M-m+1)/2$ parameters.

Let us recall in detail the question given in [1]. Suppose that, using some learning techniques, we have already extracted M fuzzy rules from the N examples. The FPR form is IF (attribute-value) THEN (Class), where the class is either C1 or C2. The extracted M rules are categorized into two groups, $S_1 = \{R_i, i = 1, 2, ..., m\}$ and $S_2 = \{R_i, i = m+1, m+2, ..., M\}$, one leading to the consequent C1 and the other leading to C2. The N examples are also classified into two parts, as follows: $T_1 = \{e_i, i = 1, 2, ..., n\}$ and $T_2 = \{e_i, i = n+1, n+2, ...N\}$. The actual classification of the examples within $T_1$ is C1, and within $T_2$ is C2. Noting the definitions of $S_1$, $S_2$ and $T_1$, $T_2$, we hope that the following inequalities hold: $x_{i1} > x_{i2}$ for $i = 1, 2, ..., n$, $x_{i1} < x_{i2}$ for $i = n+1, n+2, ..., N$.

We may numerically determine the fuzzy measure which is unknown by using the optimisation criterion of reasoning accuracy. Let $\mu_1$, $\mu_2$ be two set functions defined on $S_1 = \{R_i, i = 1, 2, ..., m\}$ and $S_2 = \{R_i, i = m+1, m+2, ..., M\}$.

Suppose that matching degree functions of $e_i$ matching $S_1$ and $S_2$ are

$$f_{i1} = \left(SM_i^{(1)}, SM_i^{(2)}, \ldots SM_i^{(n)}\right) \text{ and } f_{i2} = \left(SM_i^{(m+1)}, SM_i^{(m+2)}, \ldots SM_i^{(M)}\right),$$

where $SM_i^{(j)}$ is the result of $e^i$ matching the $j$-th rule $R_j$ $(i = 1, 2, \ldots, N; j = 1, 2, \ldots, M)$ then $x_{i1} = (C)\int_{S_1} f_{i1} d_{\mu_1}$ and $x_{i2} = (C)\int_{S_2} f_{i2} d_{\mu_2}$.

Therefore we have the following inequalities:

$$(C)\int_{S_1} f_{i1} d_{\mu_1} > (C)\int_{S_2} f_{i2} d_{\mu_2} \text{ for } i = 1, 2, \ldots, n,$$

$$(C)\int_{S_1} f_{i1} d_{\mu_1} < (C)\int_{S_2} f_{i2} d_{\mu_2} \text{ for } i = n+1, n+2, \ldots, N,$$

subject to $0 \le \mu_1, \mu_2 \le 1$.

If we use 2-additive fuzzy measure instead of the set function in [1], the problem of solving the inequalities (13)-(16) in [1] can be transformed into the following linear programming problem (2).

$$\text{Minimize: } \xi_1 + \xi_2 + \ldots + \xi_n$$

$$\begin{aligned}
\text{Subject to } & \sum_{i=1}^{m}\sum_{j=i}^{m} b_{ij}^k a_{ij} + \sum_{i=m+1}^{M}\sum_{j=i}^{M} b_{ij}^k a_{ij} + \xi_k > 0 \quad K = 1, 2, \ldots N \\
& \sum_{i\in L_1} a_{ii} + \sum_{\{i,j\}\subseteq L_1} a_{ij} = 1 \quad L_1 = \{1, 2, \ldots, m\}; \\
& \sum_{i\in L_2} a_{ii} + \sum_{\{i,j\}\subseteq L_2} a_{ij} = 1 \quad L_2 = \{m+1, m+2, \ldots, M\}; \\
& a_{ii} \ge 0 \ \forall i \in L_1, \forall i \in L_2; \\
& a_{ii} + \sum_{j\in T_1} a_{ij} \ge 0 \quad \forall i \in L_1, \forall T_1 \subseteq L_1 \backslash \{i\}; \\
& a_{ii} + \sum_{j\in T_2} a_{ij} \ge 0 \ \forall i \in L_2, \forall T_2 \subseteq L_2 \backslash \{i\}; \\
& \xi_i > 0 \ i = 1, 2, \ldots N
\end{aligned} \tag{2}$$

where $a_{ii}$ is the measure of $R_i$, $\mu_{ij}$ is the measure of $\{R_i, R_j\}$,

$$L_1 = \{1, 2, \ldots, m\}, L_2 = \{m+1, m+2, \ldots M\} \quad \mu_i = a_i = a_{ii},$$
$$\mu_{ij} = a_i + a_j + a_{ij} = a_{ii} + a_{jj} + a_{ij}.$$

In this section, with respect to our particular issues of handling interaction among rules, we only need to determine the values of $\mu_i$ and $\mu_{ij}$ , that is $a_{ii}$ and $a_{ij}$ . Once we determined the value of $a_{ii}$ and $a_{ij}$ , the other value of the composed rules can be expressed by $a_{ii}$ and $a_{ij}$ . The details about no interaction case are presented in [1].

When we use 2-additive fuzzy measure to replace the set function, the advantage is that the number of parameters is reduced from an exponential to polynomial quantity with the increasing number of rules; the disadvantage is that the accuracy may be lower than the set function. Because 2-additive fuzzy measure satisfies monotocinity, it only considers the enhancing-effect as the interaction exists among rules, but the resisting-effect cannot be considered. The set function not only expresses the enhancing-effect but also resisting-effect properly, but it has too much computational complexity.

## 4 Experimental Simulation

We chose 5 widely used machine learning classification problems to verify advantages of our method. The five databases employed for experiments are obtained from [7]. We conduct our experiments as follows. Each database is randomly split into two parts. One part is used for training while the remaining is used for testing.

**Table 1.** The comparative of the reasoning accuracy of globally weighted, the set function, 2-additive fuzzy measure

| Database | Globally weighted | | The set function | | 2-additive fuzzy measure | |
|---|---|---|---|---|---|---|
| | Training accuracy | Testing accuracy | Training accuracy | Testing accuracy | Training accuracy | Testing accuracy |
| Glass Identification | 60% | 58.47% | 70.47% | 61.54% | 66.67% | 60% |
| Pima India diabetes | 75.06% | 73.6% | 78.4% | 74.46% | 77.47% | 74.03% |
| Rice taste | 84.9% | 84.4% | 86.3% | 87.5% | 84.9% | 87.5% |
| Mango | 70.96% | 72% | 71.55% | 72% | 72.41% | 72% |
| Wine | 95.97% | 77.78% | 95.97% | 79.63% | 95.97% | 77.78% |

From Table1, we can summarize the following experimental conclusion:

1) To a certain degree, the amount of training and testing accuracy improvement depends on the concrete structure of database.

2) Of the five databases, the accuracy of using 2-additive fuzzy measure to represent the interaction is between the weighted fuzzy reasoning and reasoning based on the set function.

3) In Table 1, the Mango leaf data 's testing accuracy is almost same in three methods. This implies that the rules extracted from the database have not interactive effect. Wine data shows that the learning accuracy does not improve significantly. This implies that the rules extracted from the databases have little interactive effect. In these

situations the handling of interaction among rules can be replaced with handling based on weights.

4) From Rice Taste and Glass Identification databases, we can see that the interaction exists among the rules and the accuracy of using 2-additive fuzzy measure to represent the interaction is between the weighted fuzzy reasoning and the reasoning based on the set function.

From Table 1, we see that the interaction can be ignored in some databases. We also see that 2-additive fuzzy measure can replace the set function in some cases and to some extent.

## 5 Conclusions

The number of the determined parameters by using a set function is too large when there are many rules. It is not feasible to implement in the real world. This paper proposes using 2-additive fuzzy measure to replace the set function for the interaction handling. The main advantage is that the parameters can be reduced greatly. The disadvantage is that the reasoning accuracy of using the 2-additive to represent the interaction is little lower than using the set function in some cases. To balance the complexity (i.e., the number of parameters to be determined) and the reasoning accuracy, this paper seems to have given an appropriate trade-off by replacing the general set function with the 2-additive fuzzy measure.

## Acknowledgement

This research is supported by the National Natural Science Foundation of China (NSFC 60473045 / 60573069); the Natural Science Foundation of Hebei Province (603137).

## References

1. Daniel Yeung, Xi-Zhao Wang, Handling Interaction in Fuzzy Production Rule Reasoning, IEEE Transactions on Systems, Man and Cybernetics, Part B: Cybernetics, Volume: 34, No.5, October 2004, pp. 1-9
2. M. Grabisch, The representation of importance and interaction of features by fuzzy measures, Pattern Recognition Letters., vol.17,pp.567-575,1996.
3. D.S.Yeung and E.C.C.Tsang, A comparative study on similarity based fuzzy reasoning methods, IEEE Trans, Systems, Man, Cybernetics, part B, vol.27, pp. 216-227, Apr. 1997.
4. D.S.Yeung and E.C.C.Tsang, A multilevel weighted fuzzy reasoning algorithm for expert systems, IEEE Transactions on System. Man and Cybernetics, vol. 28, pp.149-158, Apr. 1998.
5. D. S. Yeung and E. C. C. Tsang, Weighted fuzzy production rules, Fuzzy Sets and Systems, vol. 88, No 3, pp. 299-313, 1997
6. Shyi-Ming Chen, Weighted Fuzzy Reasoning Using Weighted Fuzzy Petri Nets, IEEE Transactions on Knowledge and Data Engineering, Volume: 14, No.2, March/April 2002, pp. 386-397

7. UCI Repository of Machine Learning Databases and Domain Theories[Online]. Available: ftp://ftp.ics.uci.edu/pub/machine-learningdatabases
8. W. Wang, Z. Wang, and G. J. Klir, “Genetic algorithms for determining fuzzy measure from data,” J. Intell. Fuzzy System, vol. 6, no. 2, pp.171–183, 1998.
9. Z. Wang, K.-S. Leung, and J. Wang, “A genetic algorithm for determining nonadditive set functions in information fusion,” Fuzzy Sets Systems, vol. 102, no. 3, pp. 463–469, 1999.