Knowledge Reduction in Interval-valued Information Systems

Duoqian Miao, Nan Zhang, Xiaodong Yue
Department of Computer Science and Technology, Tongji University
Shanghai, P.R. China 201804
miaoduoqian@163.com, nan_zhang@yahoo.cn, yswantfly@hotmail.com

Abstract

In this paper, the concept of α -maximal consistent blocks is proposed to formulate the new rough approximations to an arbitrary object set in interval-valued information systems. The α -maximal consistent blocks can provide the simpler discernibility matrices and discernibility functions in reduction of interval-valued information systems. This means that they can provide a more efficient computation for knowledge acquisitions. Numerical examples are employed to substantiate the conceptual arguments.

1. Introduction

Rough set theory (RST), proposed by Pawlak [7] in 1982, is a mathematical approach to deal with imprecision, vagueness and uncertainty in data analysis. It has been successfully applied in many fields, such as data classification, medical diagnosis, pattern recognition, image processing, decision analysis, process control, information retrieval, and conflict analysis.

Due to rampant existence of interval numbers in information systems in real life, the research on combination of rough sets and interval numbers becomes necessary [21-23]. Yao [21] presented a model for the interval set by using the lower and upper approximations. Yao and Liu [23] introduced the generalized decision logic in interval-set-valued information tables which is an extension of decision logic studied by Pawlak. Leung *et al.* [2] defined the concept of misclassification rates and discovered the classification rules from the interval-valued information systems. Qian *et al.* [8] investigated a dominance relation in interval-valued information systems, and proposed attribute reductions of interval ordered information systems that eliminate the redundant information. Rebolledo [9] presented the interval qualitative models based on rough set theory which are more compact and precise than ordinary qualitative models.

Rough set models in interval-valued information systems are also called grey-rough set models in the research of grey system theory. Yang *et al.* [20] introduced a new model based on grey systems to unify fuzzy sets and rough sets in one model, and also proved that fuzzy sets and rough sets are special cases of the proposed grey sets. Crisp, fuzzy, grey and rough mathematical models are reviewed systematically in literature [15]. In [13], Wu *et al.* defined a tolerance relation in grey information systems, and got the decision rules from the decision tables. Based on the equivalence class $[x]_{GR}$,Yamaguchi *et al.* [16] suggested a grey-rough set model which is the special case of grey-rough set model proposed in [13]when $\alpha = 1$ and the tolerance relation actually is the equivalence relation $[x]_{GR}$ .In [17], Yamaguchi *et al.* investigated grey lattice operations, and presented a new grey-rough set model which is based on the grey lattice relation instead of an equivalence relation in classical rough set theory. Li *et al.* [5] introduced a method to resolve supplier's selection problems in combination of rough set theory and grey system theory. In [14], Wu *et al.* suggested a new model of grey-rough set in terms of the tolerance relation, and compared the restricted tolerance relation with the grey tolerance relation in grey information decision tables. Wu and Liu [12] introduced the real formal concept analysis about grey-rough set theory by using grey numbers, and proposed a grey-rough set approach to Galois lattices reductions. An interval data reduction of attributes was proposed in [18] .Yamaguchi *et al.* investigated [19] a method of decision rules extraction in grey information systems.

The rest of the paper is organized as follows. The fundamental notions in interval-valued information systems are introduced in section 2. In section 3, approximations based on α -maximal consistent blocks to a set are redefined. We suggest the knowledge reduction based on α -maximal consistent blocks, and

Proc. 8th IEEE Int. Conf. on Cognitive Informatics (ICCI'09)
G. Baciu, Y. Wang, Y.Y. Yao, W. Kinsner, K. Chan & L.A. Zadeh (Eds.)
978-1-4244-4642-1/09/$25.00 ©2009 IEEE

discover the latent knowledge in an IvIS in section 4. The whole paper is summarized in section 5.

2. Fundamental notions in interval-valued information systems

2.1 Interval-valued information systems

In real life, the data in an information system is not restricted to the discrete value. So, an Interval-valued Information System (IvIS) is proposed as follows:

Definition 1[2]. Let $\zeta=(U,AT,V,f)$ denote an information system called an Interval-valued Information System (IvIS), where $U=\{u_1,u_2,...,u_n\}$ is a non-empty finite set called the universe of discourse, $AT=\{a_1,a_2,...,a_m\}$ is non-empty finite set of m attributes, such that $a_k(u_i)=[l_i^k,u_i^k]$, $l_i^k \le u_i^k$ for all $i=1,2,...,n$ and $k=1,2,...,m$. V is a set of values. f is called the information function as $f:U\times AT\rightarrow V$.

Example 1. Table1 is an IvIS $\zeta=(U,AT,V,f)$, where $U=\{u_1,u_2,...,u_{10}\}$, $AT=\{a_1,a_2,a_3,a_4,a_5\}$, the attribute value $a_k(u_i)=[l_i^k,u_i^k]$ is an interval number.

2.2 Similarity rates

In this section, we will define the similarity rates in an IvIS. First, we introduce some basic operators of interval computing [17], such as intersection, union, and complement.

(1) **Intersection** $a_k(u_i)\cap a_k(u_j)$:

$$a_k(u_i)\cap a_k(u_j)=\begin{cases}\left[l_i^k,u_i^k\right] & \left[l_i^k,u_i^k\right]\subseteq\left[l_j^k,u_j^k\right]\\ \left[l_j^k,u_j^k\right] & \left[l_j^k,u_j^k\right]\subseteq\left[l_i^k,u_i^k\right]\\ \left[l_i^k,u_j^k\right] & l_i^k\in\left[l_j^k,u_j^k\right]\wedge u_j^k\in\left[l_i^k,u_i^k\right]\\ \left[l_j^k,u_i^k\right] & l_j^k\in\left[l_i^k,u_i^k\right]\wedge u_i^k\in\left[l_j^k,u_j^k\right]\\ \varnothing & \text{otherwise}\end{cases}.$$

(2) **Union** $a_k(u_i)\cup a_k(u_j)$:

$$a_k(u_i)\cup a_k(u_j)=\left[min(l_i^k,l_j^k),max(u_i^k,u_j^k)\right].$$

(3) **Complement** $\overline{a_k(u_i)}$:

$$\overline{a_k(u_i)}=(-\infty,l_i^k)\cup(u_i^k,+\infty).$$

Several kinds of relations between different numbers are shown in Fig.1.

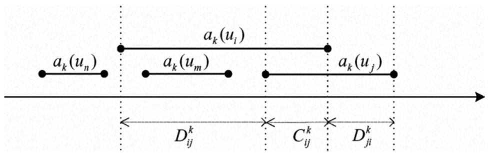

Fig.1. Interval numbers

In an interval-valued information system, we propose the concept of similarity rates as follows:

Definition 2. Let $\zeta=(U,AT,V,f)$ be an IvIS. For $\forall u_i,u_j\in U$, $a_k\in A$, the similarity rate between $a_k(u_i)$ and $a_k(u_j)$ is defined as follows:

$$\alpha_{ij}^k=min\left\{\beta_{ij}^k,\beta_{ji}^k\right\}.$$

Table 1

An interval-valued information system

	a_1	a_2	a_3	a_4	a_5
u_1	[3.12,4.56]	[5.76,6.64]	[7.92, 9.21]	[1.14,3.21]	[8.27,10.13]
u_2	[4.07,5.18]	[6.31,7.20]	[8.01,9.37]	[1.75,3.86]	[9.08,10.49]
u_3	[4.26,5.37]	[5.03,5.91]	[7.87,9.23]	[1.64,3.75]	[7.40, 8.52]
u_4	[3.00,4.44]	[5.83,6.71]	[7.01,8.21]	[1.20,3.30]	[7.85, 9.71]
u_5	[3.77,4.86]	[6.09,6.97]	[8.13,9.54]	[1.96,4.00]	[8.97,10.25]
u_6	[4.21,5.30]	[6.20,7.11]	[8.05,9.50]	[3.11,4.98]	[8.85,10.13]
u_7	[2.97,4.39]	[5.63,6.51]	[8.15,9.43]	[1.50,3.49]	[8.03, 9.97]
u_8	[4.39,5.48]	[6.14,7.02]	[8.07,9.48]	[1.80,3.88]	[9.02,10.30]
u_9	[2.05,3.14]	[6.27,7.15]	[7.98,9.32]	[1.69,3.68]	[8.89,10.25]
u_{10}	[1.15,2.35]	[5.68,6.56]	[9.01,9.89]	[0.12,1.19]	[6.41, 7.52]

Where

$$\beta_{ij}^k=\begin{cases}0 & [l_i^k,u_i^k]\cap[l_j^k,u_j^k]=\varnothing\\ min\left\{\dfrac{min\{u_i^k-l_j^k,u_j^k-l_i^k\}}{u_i^k-l_i^k},1\right\} & \text{otherwise}\end{cases}.$$

For a given similarity rate α, if $\alpha_{ij}^k\le\alpha$, then there exists the attribute $a_k\in A$ such that $a_k(u_i)$ and $a_k(u_j)$ can be separated. If $\alpha_{ij}^k>\alpha$, then there does not exist the attribute $a_k\in A$ such that $a_k(u_i)$ and $a_k(u_j)$ can be separated. Therefore, we claim that $a_k(u_i)$ and $a_k(u_j)$ can not be distinguished within α.

2.3 α-maximal consistent blocks

Definition 3[2]. Let $\zeta=(U,AT,V,f)$ be an IvIS. For a given similarity rate $\alpha\in[0,1]$, $A\subseteq AT$. The α-tolerance relation T_A^α is defined as follows:

$$T_A^\alpha=\{(u_i,u_j)\in U\times U:\alpha_{ij}^k>\alpha,\forall a_k\in A\}.$$

Where α_{ij}^k is introduced by Definition 2.

It is clear that the relation T_A^α is reflexive and symmetric, but not transitive. Based on the concept of T_A^α, the α-tolerance class corresponding to T_A^α is defined as follows:

Definition 4[2, 13, 14]. Let $\zeta=(U,AT,V,f)$ be an IvIS. $S_A^\alpha(u_i)$, called α-tolerance class for u_i with respect to T_A^α, is given as

$$S_A^\alpha(u_i)=\{u_j\in U:(u_i,u_j)\in T_A^\alpha\}.$$

Next, we introduce the α-maximal consistent block in interval-valued information systems.

Let $\zeta=(U,AT,V,f)$ be an IvIS, $A\subseteq AT$. For $\forall u_i$, $u_j\in M$, $M\subseteq U$ satisfying $(u_i,u_j)\in T_A^\alpha$, then M is the α-tolerance class in an IvIS. If $\forall u_m\in U-M$, there exists $u_i\in M$ satisfying $(u_i,u_m)\notin T_A^\alpha$. Here, M is called the α-Maximal Consistent Block (α-MCB), and $M_{AT}^\alpha(u_i)$ is the α-maximal consistent block for the object u_i with respect to T_A^α. It is the extension of application of the original maximal consistent blocks [3].

Let us continue to have $S^\alpha(AT)$ and $\xi^\alpha(AT)$, as

$$S^\alpha(AT)=\{S_{AT}^\alpha(u_1),S_{AT}^\alpha(u_2),....,S_{AT}^\alpha(u_n)\},\tag{1}$$

$$\xi^\alpha(AT)=\{M_{AT}^\alpha(u_1),M_{AT}^\alpha(u_2),....,M_{AT}^\alpha(u_n)\}.\tag{2}$$

Example 2. In an interval-valued information system given by Table 1, if $\alpha=0.7$, we can have the Boolean matrix corresponding of $T_{AT}^{0.7}$ as follows:

$$T_{AT}^{0.7}=\begin{pmatrix}1&0&0&0&0&0&1&0&0&0\\0&1&0&0&1&0&0&1&0&0\\0&0&1&0&0&0&0&0&0&0\\0&0&0&1&0&0&0&0&0&0\\0&1&0&0&1&0&0&0&0&0\\0&0&0&0&0&1&0&0&0&0\\1&0&0&0&0&0&1&0&0&0\\0&1&0&0&0&0&0&1&0&0\\0&0&0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&0&0&1\end{pmatrix}.$$

And then we obtain $S^{0.7}(AT)$ as follows:

$S^{0.7}(AT)=\{S_{AT}^{0.7}(u_1),S_{AT}^{0.7}(u_2),....,S_{AT}^{0.7}(u_{10})\}$.

$S_{AT}^{0.7}(u_1)=\{u_1,u_7\}$, $S_{AT}^{0.7}(u_2)=\{u_2,u_5,u_8\}$, $S_{AT}^{0.7}(u_3)=\{u_3\}$,

$S_{AT}^{0.7}(u_4)=\{u_4\}$, $S_{AT}^{0.7}(u_5)=\{u_2,u_5\}$, $S_{AT}^{0.7}(u_6)=\{u_6\}$,

$S_{AT}^{0.7}(u_7)=\{u_1,u_7\}$, $S_{AT}^{0.7}(u_8)=\{u_2,u_8\}$,

$S_{AT}^{0.7}(u_9)=\{u_9\}$, $S_{AT}^{0.7}(u_{10})=\{u_{10}\}$.

From the formula (2), we have

$\xi^{0.7}(AT)=\{M_{AT}^{0.7}(u_1),M_{AT}^{0.7}(u_2),....,M_{AT}^{0.7}(u_{10})\}$.

$M_{AT}^{0.7}(u_1)=\{u_1,u_7\}$, $M_{AT}^{0.7}(u_2)=\{u_2,u_5\}$,

$M_{AT}^{0.7}(u_3)=\{u_3\}$, $M_{AT}^{0.7}(u_4)=\{u_4\}$, $M_{AT}^{0.7}(u_6)=\{u_6\}$,

$M_{AT}^{0.7}(u_8)=\{u_2,u_8\}$, $M_{AT}^{0.7}(u_9)=\{u_9\}$, $M_{AT}^{0.7}(u_{10})=\{u_{10}\}$.

$\xi_{AT}^\alpha(u_i)$ is the collection of $M\in\xi^\alpha(A)$ containing the object u_i, where $i,j\le n$.

$\xi_{AT}^{0.7}(u_1)=\{M_{AT}^{0.7}(u_1)=\{u_1,u_7\}\}$,

$\xi_{AT}^{0.7}(u_2)=\{M_{AT}^{0.7}(u_2)=\{u_2,u_5\},M_{AT}^{0.7}(u_8)=\{u_2,u_8\}\}$,

$\xi_{AT}^{0.7}(u_3)=\{M_{AT}^{0.7}(u_3)=\{u_3\}\}$,

$\xi_{AT}^{0.7}(u_4)=\{M_{AT}^{0.7}(u_4)=\{u_4\}\}$,

$\xi_{AT}^{0.7}(u_5)=\{M_{AT}^{0.7}(u_2)=\{u_2,u_5\}\}$,

$\xi_{AT}^{0.7}(u_6)=\{M_{AT}^{0.7}(u_6)=\{u_6\}\}$,

$\xi_{AT}^{0.7}(u_7)=\{M_{AT}^{0.7}(u_1)=\{u_1,u_7\}\}$,

$\xi_{AT}^{0.7}(u_8)=\{M_{AT}^{0.7}(u_8)=\{u_2,u_8\}\}$,

$\xi_{AT}^{0.7}(u_9)=\{M_{AT}^{0.7}(u_9)=\{u_9\}\}$,

$\xi_{AT}^{0.7}(u_{10})=\{M_{AT}^{0.7}(u_{10})=\{u_{10}\}\}$.

In what follows, we will analyze the relations between $S_A^\alpha(u_i)$ and $\xi_A^\alpha(u_i)$.

Property 1. Let $\zeta=(U,AT,V,f)$ be an interval-valued information system, $A\subseteq AT$, $M\subseteq U$. Then we will have $M\in\xi^\alpha(A)$ *iff* $M=\bigcap_{u_i\in M}S_A^\alpha(u_i)$.

Proof . (1) "$\Leftarrow$". If $M=\bigcap_{u_i\in M}S_A^\alpha(u_i)$, then $M\in\xi^\alpha(A)$. If not, we suppose $M\cup\{u_j\}$ is a maximal consistent block with respect to A, $u_j \notin M$. For arbitrary $u_i \in M$, we obtain $u_j \in S_A^\alpha(u_i)$. Thus, $u_j \in \bigcap_{u_i \in M} S_A^\alpha(u_i)$. Because $u_j \notin M$, so $M \neq \bigcap_{u_i \in M} S_A^\alpha(u_i)$. This is contrary to $M = \bigcap_{u_i \in M} S_A^\alpha(u_i)$. Thus, if $M = \bigcap_{u_i \in M} S_A^\alpha(u_i)$, then $M \in \xi^\alpha(AT)$. (2) "$\Rightarrow$". If $M \in \xi^\alpha(A)$, then $M = \bigcap_{u_i \in M} S_A^\alpha(u_i)$. If not exist $M = \bigcap_{u_i \in M} S_A^\alpha(u_i)$, we suppose $M \cup \{u_j\} = \bigcap_{u_i \in M} S_A^\alpha(u_i)$, $u_j \notin M$. Then, we have $u_j \in \bigcap_{u_i \in M} S_A^\alpha(u_i)$. It means $M \cup \{u_j\}$ is also a consistent block. This is contrary to $M \in \xi^\alpha(A)$. So, if $M \in \xi^\alpha(A)$, then $M = \bigcap_{u_i \in M} S_A^\alpha(u_i)$. This completes the proof. □

Property 2. Let $\zeta = (U, AT, V, f)$ be an interval-valued information system, $A \subseteq AT$, then we have

$$S_A^\alpha(u_i) = \cup\{M \in \xi_A^\alpha(u_i)\}.$$

Proof . (1) According to the definition of $S_A^\alpha(u_i)$ and $\xi_A^\alpha(u_i)$, for $\forall u_j \in S_A^\alpha(u_i)$, It is clear that $u_j \in \cup\{M \in \xi_A^\alpha(u_i)\}$. So $S_A^\alpha(u_i) \subseteq \cup\{M \in \xi_A^\alpha(u_i)\}$. (2) Suppose $\forall u_j \in M$, $M \in \xi_A^\alpha(u_i)$, so $(u_i, u_j) \in T_A^\alpha$. Therefore $u_j \in S_A^\alpha(u_i)$, $\cup\{M \in \xi_A^\alpha(u_i)\} \subseteq S_A^\alpha(u_i)$. From the analysis above, we have $S_A^\alpha(u_i) = \cup\{M \in \xi_A^\alpha(u_i)\}$. This completes the proof. □

3. Approximations to a set in an IvIS

The lower and upper approximations to a subset X of U in an IvIS in [2, 13, 14] are listed as follows:

Definition 5. Let $\zeta = (U, AT, V, f)$ be an interval-valued information system, $A \subseteq AT$.

$$\underline{APR}_A(X) = \{u_i \in U : S_A^\alpha(u_i) \subseteq X\},$$

$$\overline{APR}_A(X) = \{u_i \in U : S_A^\alpha(u_i) \cap X \neq \varnothing\}.$$

Property 3. Let $\zeta = (U, AT, V, f)$ be an interval-valued information system, $A \subseteq AT$, $X \subseteq U$. Then

$$\underline{APR}_A(X) = \{u_i \in U : S_A^\alpha(u_i) \subseteq X\} = \{u_i \in X : S_A^\alpha(u_i) \subseteq X\},$$

$$\overline{APR}_A(X) = \{u_i \in U : S_A^\alpha(u_i) \cap X \neq \varnothing\} = \cup\{S_A^\alpha(u_i) : u_i \in X\}.$$

In this section, we will redefine the concepts of lower and upper approximations to a subset X in an IvIS, and compare the proposed definition with Definition 5 to illustrate the advantages of the new definition.

Definition 6. Let $\zeta = (U, AT, V, f)$ be an interval-valued information system, for $\forall X \subseteq U$, $A \subseteq AT$, we define the lower and upper approximations to X as follows:

$$\underline{apr}_A(X) = \cup\{M_A^\alpha(u_i) : M_A^\alpha(u_i) \subseteq X\},$$

$$\overline{apr}_A(X) = \cup\{M_A^\alpha(u_i) : M_A^\alpha(u_i) \cap X \neq \varnothing\}.$$

The positive region, boundary region and negative region in an IvIS can be expressed respectively as:

$$pos_A(X) = \underline{apr}_A(X),$$

$$bnr_A(X) = \overline{apr}_A(X) - \underline{apr}_A(X),$$

$$neg_A(X) = U - pos_A(X) \cup bnr_A(X).$$

Property 4. Let $\zeta = (U, AT, V, f)$ be an interval-valued information system, $A \subseteq AT$, $X \subseteq U$. We have

$$\underline{APR}_A(X) \subseteq \underline{apr}_A(X),\ \overline{APR}_A(X) = \overline{apr}_A(X).$$

Proof. First, we will prove $\underline{APR}_A(X) \subseteq \underline{apr}_A(X)$.

According to Definition 5 and Property 2, we have

$$\underline{APR}_A(X) = \{u_i \in U : S_A^\alpha(u_i) \subseteq X\} = \{u_i \in X : S_A^\alpha(u_i) \subseteq X\} = \{u_i \in X : \cup\{M \in \xi_A^\alpha(u_i)\} \subseteq X\}.$$

$$\underline{apr}_A(X) = \cup\{M \in \xi^\alpha(A) : M \subseteq X\}.$$

Thus, we can obtain $\underline{APR}_A(X) \subseteq \underline{apr}_A(X)$.

Next we prove $\overline{APR}_A(X) = \overline{apr}_A(X)$.

$$\overline{APR}_A(X) = \{u_i : S_A^\alpha(u_i) \cap X \neq \varnothing\} = \{u_i \in U : \cup\{M \in \xi_A^\alpha(u_i)\} \cap X \neq \varnothing\} = \cup\{M \in \xi_A^\alpha(u_i) : M \cap X \neq \varnothing\} = \cup\{M \in \xi^\alpha(A) : M \cap X \neq \varnothing\} = \overline{apr}_A(X).$$

This completes the proof. □

Property 4 indicates that higher approximation accuracy is achieved with respect to A by Definition 6 than by Definition 5. We give the example as follows:

Example 3. In an interval-valued information system given by Table 1, let $X = \{u_2, u_4, u_5\}$, we obtain the approximation accuracies of X according to Definition 5 and Definition 6 as follows:

$$\underline{APR}_{AT}(X) = \{u_i \in U : S_{AT}^{0.7}(u_i) \subseteq X\} = \{u_4, u_5\},$$

$$\overline{APR}_{AT}(X) = \{u_i \in U : S_{AT}^{0.7}(u_i) \cap X \neq \varnothing\} = \{u_2, u_4, u_5, u_8\};$$

$$\underline{apr}_{AT}(X) = \cup\{M_{AT}^{0.7}(u_i) : M_{AT}^{0.7}(u_i) \subseteq X\} = \{u_2, u_4, u_5\},$$

$$\overline{apr}_{AT}(X) = \cup\{M_{AT}^{0.7}(u_i) : M_{AT}^{0.7}(u_i) \cap X \neq \varnothing\} = \{u_2, u_4, u_5, u_8\}.$$

According to Definition 5 and 6, the approximation accuracies of X are calculated as:

$\gamma_{AT} = \frac{|\underline{APR}_{AT}(X)|}{|\overline{APR}_{AT}(X)|} = \frac{1}{2}$, $\gamma'_{AT} = \frac{|\underline{apr}_{AT}(X)|}{|\overline{apr}_{AT}(X)|} = \frac{3}{4}$.

Where $|S|$ denotes the cardinality of set *S*.

Obviously, we can obtain $\gamma_{AT} < \gamma'_{AT}$. Example 3 also indicates the validity of Property 4.

4. Rough reduction in an IvIS

Attribute reduction plays an important role in Pawlak rough set theory. The notion of a reduct is essential for analyzing an information table, which is a minimum subset of attributes that provides the same descriptive or classification ability as the entire set of attributes. In other words, attributes in a reduct are sufficient and necessary in an information system. In recent years, many types of knowledge reduction have been researched. Kryszkiewicz [1] proposed a type of attribute reduction that only eliminates the information which is not essential in incomplete information systems. Miao *et al.* and Wang *et al.* [6,11] gave greedy algorithms of knowledge reductions by using the mutual information and conditional information entropy as the heuristic information respectively. Yao and Zhao [24] introduced attribute reduction in decision-theoretic rough set models.

In this section, an approach to attribute reduction in an IvIS will be established and some relative examples will be given. Now, we present the definition of a reduct in an IvIS as follows:

Definition 7. Given an interval-valued information system $\zeta = (U, AT, V, f)$, an attribute set A is a reduct of ζ if it satisfies the following two conditions:

(1) $\xi^{\alpha}(A) = \xi^{\alpha}(AT)$.

(2) For any attribute set $B \subset A \subset AT$,

$$\xi^{\alpha}(B) \neq \xi^{\alpha}(AT).$$

The set of all reducts with respect to AT in ζ is denoted by $red^{\alpha}(AT)$.

Definition 8. Let $\zeta = (U, AT, V, f)$ be an interval-valued information system, $A_1, A_2, ..., A_t$ are all the reducts of ζ. We call a_i the core attribute in ζ *iff* $a_i \in A_1 \cap A_2 \cap ... \cap A_t$. $cor^{\alpha}(AT)$ is the set of a_i.

Based on the concept of discernibility matrix and discernibility function [10], we have the following theorem to compute all reducts in an IvIS.

Theorem 1. Let $\zeta = (U, AT, V, f)$ be an IvIS, $A \subseteq AT$ is a reduct of ζ *iff* $\wedge A$ is a prime implicant of discernibility function $\psi^{\alpha}(AT)$ which can be expressed as follows:

$$\psi^{\alpha}(AT) = \bigwedge_{\{M^{\alpha}_{AT}(u_i), M^{\alpha}_{AT}(u_j)\} \in \xi^{\alpha}(AT) \times \xi^{\alpha}(AT)} \alpha_{AT}\{M^{\alpha}_{AT}(u_i), M^{\alpha}_{AT}(u_j)\}.$$

Where

$$\alpha_{AT}\{M^{\alpha}_{AT}(u_i), M^{\alpha}_{AT}(u_j)\} = \bigwedge_{(u_i,u_j) \in M^{\alpha}_{AT}(u_i) \times M^{\alpha}_{AT}(u_j)} \vee \alpha(u_i, u_j),$$

$$\alpha(u_i, u_j) = \{a_k \in AT : \alpha^k_{ij} \leq \alpha\}.$$

Proof.

Let

$$\mathbb{C} = \{A \subseteq AT : \forall (M^{\alpha}_{AT}(u_i), M^{\alpha}_{AT}(u_j)) \in \xi^{\alpha}(AT) \times \xi^{\alpha}(AT), M^{\alpha}_{AT}(u_i) \neq M^{\alpha}_{AT}(u_j); \forall (u_i, u_j) \in M^{\alpha}_{AT}(u_i) \times M^{\alpha}_{AT}(u_j), A \cap \alpha(u_i, u_j) \neq \varnothing, \alpha(u_i, u_j) \neq \varnothing\}.$$

$$\mathbb{R} = \{A \subseteq AT : \xi^{\alpha}(A) = \xi^{\alpha}(AT)\}.$$

"$\mathbb{C} \Rightarrow \mathbb{R}$". (1) Suppose $M \in \xi^{\alpha}(A)$, $u_i, u_j \in M$. It implies that the arbitrary attribute in attribute set A can not distinguish the object u_i and u_j, then $A \cap \alpha(u_i, u_j) = \varnothing$. Suppose $M - \{u_j\} \in \xi^{\alpha}(AT)$, $M \notin \xi^{\alpha}(AT)$. According to the definition of $\mathbb{C}$, we can have $A \cap \alpha(u_i, u_j) \neq \varnothing$ which is contrary to $A \cap \alpha(u_i, u_j) = \varnothing$. So, if $M \in \xi^{\alpha}(A)$, then $M \in \xi^{\alpha}(AT)$. Therefore, $\xi^{\alpha}(A) \subseteq \xi^{\alpha}(AT)$. (2) Next we prove $\xi^{\alpha}(AT) \subseteq \xi^{\alpha}(A)$. Suppose $M \in \xi^{\alpha}(AT)$, $u_i \in M$ and $u_j \notin M$. By the definition of $\mathbb{C}$, $A \cap \alpha(u_i, u_j) \neq \varnothing$. We suppose $M \notin \xi^{\alpha}(A)$, and $M \cup \{u_j\} \in \xi^{\alpha}(A)$ which notes that for the arbitrary attribute in attribute set A can not distinguish the attribute u_i and u_j, so $A \cap \alpha(u_i, u_j) = \varnothing$ which is contrary to $A \cap \alpha(u_i, u_j) \neq \varnothing$. So, we can obtain $\xi^{\alpha}(AT) \subseteq \xi^{\alpha}(A)$. By $\xi^{\alpha}(A) \subseteq \xi^{\alpha}(AT)$ and $\xi^{\alpha}(AT) \subseteq \xi^{\alpha}(A)$, $\xi^{\alpha}(AT) = \xi^{\alpha}(A)$ holds.

"$\mathbb{R} \Rightarrow \mathbb{C}$".

Suppose $\{M^{\alpha}_{AT}(u_i), M^{\alpha}_{AT}(u_j)\} \in \xi^{\alpha}(AT) \times \xi^{\alpha}(AT)$, $M^{\alpha}_{AT}(u_i) \neq M^{\alpha}_{AT}(u_j)$, $(u_i, u_j) \in M^{\alpha}_{AT}(u_i) \times M^{\alpha}_{AT}(u_j)$. It implies the attribute set AT can distinguish the object u_i and u_j, thus $AT \cap \alpha(u_i, u_j) \neq \varnothing$. Because of $\xi^{\alpha}(AT) = \xi^{\alpha}(A)$, it means that A provides the same descriptive or classification ability as *AT*. Thus, we can obtain $A \cap \alpha(u_i, u_j) \neq \varnothing$.

From the discussion above, $\mathbb{C} \Leftrightarrow \mathbb{R}$ holds. $\mathbb{C} \Leftrightarrow \mathbb{R}$ shows that $A \subseteq AT$ is a reduct of ζ *iff* $\wedge A$ is a prime implicant of discernibility function ψ^{α}. □

In some applications of reduction in interval-valued information systems, we just need identify some objects from the rest. So, the concept of a reduct for an object of *U* in an IvIS will be introduced as follows:

Definition 9. Given an interval-valued information system $\zeta=(U,AT,V,f)$, an attribute set A is a reduct of ζ for u_i, if it satisfies the following two conditions:

(1) $\xi_A^\alpha(u_i)=\xi_{AT}^\alpha(u_i)$.

(2) For any attribute set $B\subset A\subset AT$,

$$\xi_B^\alpha(u_i)\neq\xi_{AT}^\alpha(u_i).$$

The set of all reducts of u_i with respect to AT in ζ is denoted by $red_{AT}^\alpha(u_i)$.

Definition 10. Let $\zeta=(U,AT,V,f)$ be an interval-valued information system, and $A_1,A_2,...,A_t$ are all reducts of ζ for u_i. We call a_i the core attribute for u_i in ζ *iff* $a_i\in A_1\cap A_2\cap...\cap A_t$. $cor_{AT}^\alpha(u_i)$ is the set of a_i for u_i.

Theorem 2. Let $\zeta=(U,AT,V,f)$ be an interval-valued information system, $A\subseteq AT$ is a reduct of ζ for u_i *iff* $\wedge A$ is a prime implicant of discernibility function $\psi_{AT}^\alpha(u_i)$ which is given as follows:

$$\psi_{AT}^\alpha(u_i)=\bigwedge_{M_{AT}^\alpha(u_j)\in\xi^\alpha(AT)-\xi_{AT}^\alpha(u_i)}\alpha_{AT}\left\{u_i,M_{AT}^\alpha(u_j)\right\}$$

$$\bigwedge_{\left\{M_{AT}^\alpha(u_m),M_{AT}^\alpha(u_n)\right\}\in\xi_{AT}^\alpha(u_i)\times\xi_{AT}^\alpha(u_i)}\alpha_{AT}\left\{M_{AT}^\alpha(u_m),M_{AT}^\alpha(u_n)\right\}.$$

Where

$$\alpha_{AT}\left\{u_i,M_{AT}^\alpha(u_j)\right\}=\bigwedge_{u_j\in M_{AT}^\alpha(u_j)}\vee\alpha(u_i,u_j),$$

$$\alpha_{AT}\left\{M_{AT}^\alpha(u_m),M_{AT}^\alpha(u_n)\right\}=\bigwedge_{u_m,u_n\in M_{AT}^\alpha(u_m)\times M_{AT}^\alpha(u_n)}\vee\alpha(u_m,u_n),$$

$$\alpha(u_i,u_j)=\left\{a_k\in AT:\alpha_{ij}^k\le\alpha\right\},$$

$$\alpha(u_m,u_n)=\left\{a_k\in AT:\alpha_{mn}^k\le\alpha\right\}.$$

Proof:

Let

$$\mathbb{C}=\left\{A\subseteq AT:\forall M'\in\xi^\alpha(AT)-\xi_{AT}^\alpha(u_i),\right.$$
$$\forall M_{AT}^\alpha(u_m),M_{AT}^\alpha(u_n)\in\xi_{AT}^\alpha(u_i),\ M_{AT}^\alpha(u_m)\neq M_{AT}^\alpha(u_n),$$
$$\forall(u_m,u_n)\in M_{AT}^\alpha(u_m)\times M_{AT}^\alpha(u_n),\ A\cap\alpha(u_m,u_n)\neq\varnothing,$$
$$\left.\alpha(u_m,u_n)\neq\varnothing\right\},$$
$$\mathbb{R}=\left\{A\subseteq AT:\xi_A^\alpha(u_i)=\xi_{AT}^\alpha(u_i)\right\}.$$

"$\mathbb{C}\Rightarrow\mathbb{R}$". (1) If $M\in\xi_A^\alpha(u_i)$, $u_i,u_j\in M$. It implies $A\cap\alpha(u_i,u_j)=\varnothing$. Suppose $M\notin\xi_{AT}^\alpha(u_i)$, $M-\{u_j\}\in\xi_{AT}^\alpha(u_i)$. We can have $u_j\in M'$, $M'\in\xi_{AT}^\alpha(u_i)$, $M\neq M'$, or $u_j\in M''$, $M''\in\xi^\alpha(AT)-\xi_{AT}^\alpha(u_i)$. Case 1: $u_j\in M'$. According to $A\cap\alpha(u_m,u_n)\neq\varnothing$ in definition $\mathbb{C}$, we can have $A\cap\alpha(u_i,u_j)\neq\varnothing$ There is a conflict. So, if $M\in\xi_A^\alpha(u_i)$, then $M\in\xi_{AT}^\alpha(u_i)$. Thus, $\xi_A^\alpha(u_i)\subseteq\xi_{AT}^\alpha(u_i)$. Case 2: $u_j\in M''$. According to $A\cap\alpha(u_i,u_j)\neq\varnothing$ in definition $\mathbb{C}$, we can have $A\cap\alpha(u_i,u_j)\neq\varnothing$ which is contrary to $A\cap\alpha(u_i,u_j)=\varnothing$. Thus, if $M\in\xi_A^\alpha(u_i)$, then $M\in\xi_{AT}^\alpha(u_i)$. Therefore, $\xi_A^\alpha(u_i)\subseteq\xi_{AT}^\alpha(u_i)$. (2) Next we prove $\xi_{AT}^\alpha(u_i)\subseteq\xi_A^\alpha(u_i)$. If $M\in\xi_{AT}^\alpha(u_i)$, and $u_i\in M$, $u_j\notin M$. It implies $A\cap\alpha(u_i,u_j)\neq\varnothing$. Suppose $M\cup\{u_j\}\in\xi_A^\alpha(u_i)$ it means any attribute in attribute set A can not distinguish the object u_i and u_j, thus $A\cap\alpha(u_i,u_j)=\varnothing$ which is contrary to $A\cap\alpha(u_i,u_j)\neq\varnothing$. So, if $M\in\xi_{AT}^\alpha(u_i)$ then $M\in\xi_A^\alpha(u_i)$. We obtain $\xi_{AT}^\alpha(u_i)\subseteq\xi_A^\alpha(u_i)$.

"$\mathbb{R}\Rightarrow\mathbb{C}$". If $u_i\in M$, $u_j\notin\xi_{AT}^\alpha(u_i)$, then we can obtain $AT\cap\alpha(u_i,u_j)\neq\varnothing$. Because $\xi_A^\alpha(u_i)=\xi_{AT}^\alpha(u_i)$, so $A\cap\alpha(u_i,u_j)\neq\varnothing$. If $\forall M_{AT}^\alpha(u_m)$, $M_{AT}^\alpha(u_n)\in\xi_{AT}^\alpha(u_i)$, $\forall(u_m,u_n)\in M_{AT}^\alpha(u_m)\times M_{AT}^\alpha(u_n)$, $M_{AT}^\alpha(u_m)\neq M_{AT}^\alpha(u_n)$. According to $\xi_{AT}^\alpha(u_i)=\xi_A^\alpha(u_i)$, it is clear that $A\cap\alpha(u_m,u_n)\neq\varnothing$, $\alpha(u_m,u_n)\neq\varnothing$.

From the discussion above, we can have $\mathbb{C}\Leftrightarrow\mathbb{R}$ which shows that $A\subseteq AT$ is a reduct of ζ for u_i *iff* $\wedge A$ is a prime implicant of discernibility function $\psi_{AT}^\alpha(u_i)$. This completes the proof. □

Example 4. Based on Theorem 1 and 2, we can obtain $\psi^{0.7}(AT)$ and $\psi_{AT}^{0.7}(u_i)$ of ζ given by Table 1, as follows:

$$\psi^{0.7}(AT)=a_1\wedge a_3\wedge a_4\wedge(a_2\vee a_5)$$
$$=(a_1\wedge a_2\wedge a_3\wedge a_4)\vee(a_1\wedge a_3\wedge a_4\wedge a_5),$$
$$red^{0.7}(AT)=\left\{\{a_1,a_2,a_3,a_4\},\{a_1,a_3,a_4,a_5\}\right\},$$
$$cor^{0.7}(AT)=\{a_1,a_2,a_3,a_4\}\cap\{a_1,a_3,a_4,a_5\}=\{a_1,a_3,a_4\};$$
$$\psi_{AT}^{0.7}(u_1)=a_3\wedge(a_1\vee a_2\vee a_5)$$
$$=(a_1\wedge a_3)\vee(a_2\wedge a_3)\vee(a_3\wedge a_5),$$
$$red_{AT}^{0.7}(u_1)=\{\{a_1,a_3\},\{a_2,a_3\},\{a_3,a_5\}\};$$
$$\psi_{AT}^{0.7}(u_2)=((a_2\vee a_5)\wedge a_4)\wedge a_1$$
$$=(a_1\wedge a_2\wedge a_4)\vee(a_1\wedge a_4\wedge a_5),$$
$$red_{AT}^{0.7}(u_2)=\left\{\{a_1,a_2,a_4\},\{a_1,a_4,a_5\}\right\};$$
$$\psi_{AT}^{0.7}(u_3)=a_2\vee a_5,\ red_{AT}^{0.7}(u_3)=\left\{\{a_2\},\{a_5\}\right\};$$
$$\psi_{AT}^{0.7}(u_4)=a_3,\ red_{AT}^{0.7}(u_4)=\left\{\{a_3\}\right\};$$
$$\psi_{AT}^{0.7}(u_5)=(a_2\vee a_5)\wedge a_4,\ red_{AT}^{0.7}(u_5)=\left\{\{a_2,a_4\},\{a_4,a_5\}\right\};$$
$$\psi_{AT}^{0.7}(u_6)=a_4,\ red_{AT}^{0.7}(u_6)=\left\{\{a_4\}\right\};$$
$$\psi_{AT}^{0.7}(u_7)=a_3\wedge(a_1\vee a_2\vee a_5)$$
$$=(a_1\wedge a_3)\vee(a_2\wedge a_3)\vee(a_3\wedge a_5),$$
$$red_{AT}^{0.7}(u_7)=\left\{\{a_1,a_3\},\{a_2,a_3\},\{a_3,a_5\}\right\};$$
$$\psi_{AT}^{0.7}(u_8)=(a_2\vee a_5)\wedge a_4,\ red_{AT}^{0.7}(u_8)=\left\{\{a_2,a_4\},\{a_4,a_5\}\right\};$$
$$\psi_{AT}^{0.7}(u_9)=a_1,\ red_{AT}^{0.7}(u_9)=\{a_1\};$$
$$\psi_{AT}^{0.7}(u_{10})=a_1\wedge a_3\wedge a_4\wedge a_5,\ red_{AT}^{0.7}(u_{10})=\left\{\{a_1,a_3,a_4,a_5\}\right\}.$$

5. Conclusions

In this paper, we present the new framework of knowledge reduction in an interval-valued information system based on α -maximal consistent blocks. Numerical examples are also employed to substantiate the conceptual arguments. Contributions of this paper are summarized as follows:

- The research on combination of rough sets and interval numbers (grey numbers) is reviewed systematically in this paper.
- New concept of the similarity rates is defined to measure the two interval numbers.
- The α -maximal consistent blocks are used in an IvIS for formulating approximations to an arbitrary object set in an IvIS with higher accuracy.
- The α -maximal consistent blocks in an IvIS can provide the simpler discernibility matrices and discernibility functions in reduction of an IvIS. Finally, the latent knowledge is discovered from an IvIS by knowledge reduction.

6. Acknowledgement

The authors would like to acknowledge the National Natural Science Foundation of China (No. 60775036, No.60475019) and the Research Foundation of the Doctoral Program of Higher Education of China (No.20060247039) for its partial support to the research.

References

[1] Kryszkiewicz, M. Rough set approach to incomplete information systems, *Information Sciences*, **112**, 39-49, 1998.

[2] Leung, Y. , Fischer , M.M. , Wu , W. Z. and Mi , J. S.A rough set approach for the discovery of classification rules in interval-valued information systems, *International Journal of Approximate Reasoning*, **47** , 233-246, 2008.

[3] Leung, Y. and Li, D.Y. Maximal consistent block technique for rule acquisition in incomplete information systems, *Information Sciences*, **153**, 85 -106, 2003.

[4] Leung, Y., Wu , W.Z. and Zhang ,W.X. Knowledge acquisition in incomplete information systems: A rough set approach, *European Journal of Operational Research*, **168**, 164-180, 2006.

[5] Li, G.D. , Yamaguchi, D. and Nagai, M. A grey-based rough decision-making approach to supplier selection, *Mathematical and Computer Modelling*, **46**, 573-581 , 2007.

[6] Miao, D.Q. and Hu, G.R. A heuristic algorithm for reduction of knowledge, *Chinese Journal of Computer Research and Development*, **36**, 681-684, 1999.

[7] Pawlak, Z. Rough sets, *International Journal of Information and Computer Sciences*, **11**, 341-356, 1982.

[8] Qian, Y.H., Liang, J.Y. and Dang, C.Y. Interval ordered information systems, *Computer and Mathematics with Applications*, **56**, 1994-2009, 2008.

[9] Rebolledo, M. Rough intervals—enhancing intervals for qualitative modeling of technical systems, *Artificial Intelligence*, **170**, 667–685, 2006.

[10] Skowron, A. and Rauszer, C. The discernibility matrices and functions in information systems, in: Slowinski, R. (Eds.), *Intelligence Decision Support: Handbook of Applications and Advances of Rough Set Theory*, Kluwer Academic Publisher, Dordrecht, 1992.

[11] Wang, G.Y., Yu, H. and Yang, D.C. Decision table reduction based on conditional information entropy, *Chinese Journal of Computers*, **25**, 759-766 , 2002.

[12] Wu, Q. and Liu, Z.T. Real formal concept analysis based on grey-rough set theory, *Knowledge-based systems*, **22**, 38-45, 2009.

[13] Wu ,S.X., Huang ,Z.Y., Luo, D.L. and Liu, S.F. A grey rough set model based on (α,β) -grey similarity relation, *Proceedings of 2007 IEEE International Conference on Grey Systems and Intelligent Services*, 18-20, 2007.

[14] Wu, S.X. , Shi, S.H. , Liu, S.F. and Li, M.Q. Study of grey rough set model based on tolerance relation, *Proceedings of 9th International Conference on Control, Automation, Robotics and Vision*, 1-6, 2006.

[15] Yamaguchi, D., Li, G.D. , Chen, L.C. and Nagai, M. Review Crisp, Fuzzy, Grey and Rough Mathematical Models, *Proceedings of 2007 International Conference on Grey Systems and Intelligence Services*, 547-522, 2007.

[16] Yamaguchi, D., Li, G. D. and Nagai, M. On the combination of rough set theory and grey theory based on grey lattice operations, *Proceedings of 5th International Conference on Rough Sets and Current Trends in Computing* , LNAI 4259, 507–516, 2006.

[17] Yamaguchi, D. , Li, G.D. and Nagai, M. A grey-based rough approximation model for interval data processing, *Information Sciences*, **177**, 4727-4744, 2007.

[18] Yamaguchi, D. , Li, G.D. and Nagai, M. A Grey-Rough Set Approach for Interval Data Reduction of Attributes, *Proceedings of International Conference of Rough Sets and Emerging Intelligent Systems Paradigms*, LNCS 4585, 400-410, 2007.

[19] Yamaguchi, D. , Li, G.D. , Mizutani, K. , Akabane, T., Nagai, M. and Kitaoka, M. Decision rule extraction and reduction based on grey lattice classification, *Proceedings of 4th International Conference on Machine Learning and Applications*, 31–36, 2005.

[20] Yang, Y.J., John, R. and Chiclana, F. Grey sets: a unified model for fuzzy sets and rough sets, *Proceedings of the 5th International Conference on Recent Advances in Soft Computing*, 348-353, 2004.

[21] Yao,Y.Y. Interval-set algebra for qualitative knowledge representation, *Proceedings of 5th International Conference Computing and Information*, 370–374, 1993.

[22] Yao, Y.Y. and Li, X. Comparison of rough-set and interval-set models for uncertain reasoning, *Fundamenta Informaticae*, **27**, 289-298, 1996.

[23] Yao, Y.Y. and Liu, Q. A generalized decision logic in interval-set-valued information tables, *Proceedings of the 7th International Workshop on Rough Sets, Fuzzy Sets, Data Mining, and Granular-Soft Computing*, LNAI 1711, 285-293, 1999.

[24] Yao, Y.Y. and Zhao, Y. Attribute reduction in decision-theoretic rough set models, *Information Sciences*, **178**, 3356-3373, 2008.