INFORMATION
TECHNOLOGY JOURNAL

Interval-valued Information Systems Rough Set Model
Based on Multi-granulations

Wei Yang
Department of Computer Science and Technology, Tongji University,
Shanghai, 201804, People's Republic of China

Abstract: This study extends the rough set in interval-valued information systems to the multi-granulation rough set in interval-valued information systems. The lower and upper approximations of a set in interval-valued information systems based on multi-granulations are defined and some basic properties are introduced. In order to substantiate the conceptual arguments numerical examples are given.

Key words: Interval-valued information systems, attribution reduction, rough set, multi-granulations

INTRODUCTION

Pawlak (1982) proposed the rough set theory, which is used to analyze imprecise, vague and uncertain data sets in many fields. Because of the classical rough set is not to use in some cases, such as the incomplete information system and so on. Several extensions on rough set have been proposed, such as a novel rough set approach is proposed by Leung *et al.* (2008). A dominance relation to interval information systems is presented by Qian *et al.* (2008). The knowledge reduction of α-maximal consistent blocks is proposed in interval-valued information systems by Miao *et al.* (2009).

In this study, the rough set of interval-valued information system is extended to multi-granulations rough set based on multi-granulations (Qian and Liang, 2006, 2010). Some interesting properties are discussed based on the multi-granulations rough set and proved two properties. Based on the given maximal intersection rates the reduction is computed.

THE MULTI-GRANULATIONS ROUGH SET IN INTERVAL-VALUED INFORMATION SYSTEMS

The basic notion of an interval-valued information system is cited for facilitating our discussion.

The interval-valued information system: "[L]et $\zeta = (U, AT, V, f)$ denote an information system called an interval-valued information system (IvIS), where $U = [u_1, u_2,\ldots, u_n]$ is a non-empty finite set called the universe of discourse, $AT = [a_1, a_2,\ldots, a_m]$ is non-empty finite set of m attribute, such that $a_k(u_i) = [l_i^k, u_i^k]$, $l_i^k \le u_i^k$ for all $i = 1, 2,\ldots, n$ and $k = 1, 2,\ldots, m$. V is a set of values. f is called the information function as $f:U\times AT\to V$.

Example 1: Table 1 is an IvIS $\zeta = (U, AT, V, f)$, where $U = [u_1, u_2,\ldots, u_n]$, $AT = [a_1, a_2,\ldots, a_m]$, the attribute value $a_k(u_i) = [l_i^k, u_i^k]$ is an interval number (Miao *et al.*, 2009).

Definition 1: [L]et $\zeta = (U, AT, V, f)$ be an IvIS, $X\subseteq U$, $A\subseteq AT$, card(A) = s. The lower and upper approximations to a subset X of U can be defined as follows (Miao and Yang, 2010):

$$\underline{APR}_A(X)=\left\{u_i \in U : S_A^{\beta_1,\beta_2,\ldots,\beta_s}(u_i)\subseteq X\right\}a$$

$$\overline{APR}_A(X)=\left\{u_i \in U : S_A^{\beta_1,\beta_2,\ldots,\beta_s}(u_i)\cap X \neq \varnothing\right\}$$

In the view of multi-granulations, in the interval-valued information systems the $\left(\underline{APR}_A(X),\overline{APR}_A(X)\right)$ is called single granulation rough set.

Definition 2: Let $\zeta = (U, AT, V, f)$ be an IvIS, $X\subseteq U$, A, B$\subseteq$AT, card(A) = s, card(B) = t. The lower and upper approximations to a subset X of U can be defined as follows:

$$\underline{APR}_{A+B}(X)=\left\{u_i \in U : S_A^{\beta_1,\beta_2,\ldots,\beta_s}(u_i)\subseteq X\right.$$
$$\left. or\ S_B^{\beta_1,\beta_2,\ldots,\beta_t}(u_i)\subseteq X\right\}$$
$$\overline{APR}_{A+B}(X)=\sim \underline{APR}_{A+B}(\sim X)$$

In the view of multi-granulations, in the IvIS the $\left(\underline{APR}_{A+B}(X),\overline{APR}_{A+B}(X)\right)$ is called double granulations rough set.

Definition 3: Let $\zeta = (U, AT, V, f)$ be an IvIS, $X\subseteq U$, $A_1, A_2,\ldots, A_m\subseteq AT$. The lower and upper approximations to a subset X of U can be defined as follows:

$$\underline{APR}_{\sum_1^m A_j}(X) = \left\{u_i \in U : \vee S_{A_j}^{\beta_1,\beta_2,...,\beta_k}(u_i) \subseteq X, j \le m\right\} \quad \overline{APR}_{\sum_1^m A_j}(X) = \sim \underline{APR}_{\sum_1^m A_j}(\sim X)$$

In the view of multi-granulations, in the IvIS the $\left(\underline{APR}_{\sum_1^m A_j}(X), \overline{APR}_{\sum_1^m A_j}(X)\right)$ is called multi-granulations rough set.

Proposition 1: Let $\zeta = (U, AT, V, f)$ be an IvIS, $X \subseteq U$, $A_1, A_2, ..., A_m \subseteq AT$. Some properties can be discussed as follows:

$$\underline{APR}_{\sum_1^m A_j}(X) = \cup_1^m \underline{APR}_{A_j}(X)$$

$$\overline{APR}_{\sum_1^m A_j}(X) = \cap_1^m \overline{APR}_{A_j}(X)$$

Proof: If $j>1$, the properties can be proved as follows:

$$\underline{APR}_{\sum_1^m A_j}(X) = \left\{u_i \in U : \vee S_{A_j}^{\beta_1,\beta_2,...,\beta_k}(u_i) \subseteq X, j \le m\right\}$$
$$= \cup\left\{u_i \in U : S_{A_j}^{\beta_1,\beta_2,...,\beta_k}(u_i) \subseteq X\right\}, j \le m = \cup_1^m \underline{APR}_{A_j}(X)$$

$$\overline{APR}_{\sum_1^m A_j}(X) = \sim \underline{APR}_{\sum_1^m A_j}(\sim X)$$
$$= \sim \cup_1^m \underline{APR}_{A_j}(\sim X)$$
$$= \sim \cup_1^m \sim \left(\overline{APR}_{A_j} X\right)$$
$$= \cap_1^m \overline{APR}_{A_j}(X)$$

Proposition 2: Let $\zeta = (U, AT, V, f)$ be an IvIS, $X \subseteq U$, $A_1, A_2, ..., A_m \subseteq AT$. Then:

$$\underline{APR}_{\sum_1^m A_j}(X) \subseteq \underline{APR}_{\cup_1^m A_j}(X)$$

$$\overline{APR}_{\sum_1^m A_j}(X) \supseteq \overline{APR}_{\cup_1^m A_j}(X)$$

Proposition 3: Let $\zeta = (U, AT, V, f)$ be an IvIS, $X \subseteq U$, $A_1, A_2, ..., A_m \subseteq AT$, $X_1 \subseteq X_2 \subseteq ... X_n \subseteq U$. Then:

$$\underline{APR}_{\sum_1^m A_j}(X_1) \subseteq \underline{APR}_{\sum_1^m A_j}(X_2) \subseteq \cdots \subseteq \underline{APR}_{\sum_1^m A_j}(X_n)$$

$$\overline{APR}_{\sum_1^m A_j}(X_1) \subseteq \overline{APR}_{\sum_1^m A_j}(X_2) \subseteq \cdots \subseteq \overline{APR}_{\sum_1^m A_j}(X_n)$$

Example 2: We can consider the interval-valued information system given in Table 1 and assume the maximal intersection rates (Miao and Yang, 2010) $\beta_1, \beta_2, ..., \beta_m = 0.9$ are given with respect to all the attributes.

For $X = \{u_1, u_2, u_4, u_8, u_9\}$

$$S_{\alpha_4}^{0.9}(u_1) = \{u_1, u_4\}$$
$$S_{\alpha_4}^{0.9}(u_2) = \{u_2, u_3, u_5, u_8, u_9\}$$
$$S_{\alpha_4}^{0.9}(u_3) = \{u_2, u_3, u_8, u_9\}$$
$$S_{\alpha_4}^{0.9}(u_4) = \{u_1, u_4\}$$
$$S_{\alpha_4}^{0.9}(u_5) = \{u_2, u_5, u_8\}$$
$$S_{\alpha_4}^{0.9}(u_6) = \{u_6\}$$
$$S_{\alpha_4}^{0.9}(u_7) = \{u_3, u_4, u_7, u_9\}$$

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]

$$S_{\alpha_4}^{0.9}(u_8) = \{u_2, u_3, u_5, u_8, u_9\}$$
$$S_{\alpha_4}^{0.9}(u_9) = \{u_2, u_3, u_7, u_8, u_9\}$$
$$S_{\alpha_4}^{0.9}(u_{10}) = \{u_{10}\}$$
$$\underline{APR}_{\alpha_4}(X) = \{u_1, u_4\}$$
$$\overline{APR}_{\alpha_4}(X) = \{u_1, u_2, u_3, u_4, u_5, u_7, u_8, u_9\}$$
$$S_{\alpha_5}^{0.9}(u_1) = \{u_1, u_7\}$$
$$S_{\alpha_5}^{0.9}(u_2) = \{u_2\}$$
$$S_{\alpha_5}^{0.9}(u_3) = \{u_3\}$$
$$S_{\alpha_5}^{0.9}(u_4) = \{u_4, u_7\}$$
$$S_{\alpha_5}^{0.9}(u_5) = \{u_1, u_2, u_5, u_6, u_8, u_9\}$$
$$S_{\alpha_5}^{0.9}(u_6) = \{u_1, u_5, u_6, u_9\}$$
$$S_{\alpha_5}^{0.9}(u_7) = \{u_7\}$$
$$S_{\alpha_5}^{0.9}(u_8) = \{u_2, u_5, u_8, u_9\}$$
$$S_{\alpha_5}^{0.9}(u_9) = \{u_1, u_5, u_6, u_8, u_9\}$$
$$S_{\alpha_5}^{0.9}(u_{10}) = \{u_{10}\}$$
$$\underline{APR}_{\alpha_5}(X) = \{u_2\}$$
$$\overline{APR}_{\alpha_5}(X) = \{u_1, u_2, u_4, u_5, u_6, u_8, u_9\}$$
$$\underline{APR}_{\alpha_4 + \alpha_5}(X) = \{u_1, u_2, u_4\}$$
$$\overline{APR}_{\alpha_4 + \alpha_5}(X) = \{u_1, u_2, u_4, u_5, u_8, u_9\}$$

So, $\underline{APR}_{\alpha_4 + \alpha_5}(X) = \underline{APR}_{\alpha_4}(X) \cup \underline{APR}_{\alpha_5}(X)$

$\overline{APR}_{\alpha_4 + \alpha_5}(X) = \overline{APR}_{\alpha_4}(X) \cap \overline{APR}_{\alpha_5}(X)$

$$S_{\{\alpha_4,\alpha_5\}}^{0.9}(u_1) = \{u_1\}$$
$$S_{\{\alpha_4,\alpha_5\}}^{0.9}(u_2) = \{u_2\}$$
$$S_{\{\alpha_4,\alpha_5\}}^{0.9}(u_3) = \{u_3\}$$
$$S_{\{\alpha_4,\alpha_5\}}^{0.9}(u_4) = \{u_4\}$$
$$S_{\{\alpha_4,\alpha_5\}}^{0.9}(u_5) = \{u_2, u_5, u_8\}$$
$$S_{\{\alpha_4,\alpha_5\}}^{0.9}(u_6) = \{u_6\}$$
$$S_{\{\alpha_4,\alpha_5\}}^{0.9}(u_7) = \{u_7\}$$
$$S_{\{\alpha_4,\alpha_5\}}^{0.9}(u_8) = \{u_2, u_5, u_8, u_9\}$$
$$S_{\{\alpha_4,\alpha_5\}}^{0.9}(u_9) = \{u_8, u_9\}$$
$$S_{\{\alpha_4,\alpha_5\}}^{0.9}(u_{10}) = \{u_{10}\}$$
$$\underline{APR}_{\{\alpha_4,\alpha_5\}}(X) = \{u_1, u_2, u_4, u_9\}$$
$$\overline{APR}_{\{\alpha_4,\alpha_5\}}(X) = \{u_1, u_2, u_4, u_5, u_8, u_9\}$$

So, $\underline{APR}_{\alpha_4 + \alpha_5}(X) \subseteq \underline{APR}_{\{\alpha_4,\alpha_5\}}(X)$

$\overline{APR}_{\{\alpha_4,\alpha_5\}}(X) \subseteq \overline{APR}_{\alpha_4 + \alpha_5}(X)$.

ATTRIBUTE REDUCTION

Definition 4: Let $\zeta = (U, AT, V, f)$ be an IvIS, $A \subseteq AT$, $B \subset A$, card(A) = s, card(B) = t, card(AT) = m, $A = \{a_1, a_2, ..., a_s\}$, $B = \{b_1, b_2, ..., b_t\}$, $AT = \{c_1, c_2, ..., c_m\}$, $X \subseteq U$, $A^{\beta_1,\beta_2,...,\beta_t}$ is a reduction with respect to the given $a_1, a_2, ..., a_m$ in ζ, if and only if:

$\forall X \subseteq U, \underline{APR}_{\sum_i^s b_i}(X) = \underline{APR}_{\sum_i^m c_i}(X)$ and $\forall B \subset A, \exists X \subseteq U, \underline{APR}_{\sum_i^t b_i}(X) \neq \underline{APR}_{\sum_i^m c_i}(X)$

Definition 5: Let $\zeta = (U, AT, V, f)$ be an IvIS, $A \subseteq AT$, $B \subset A$, card(A) = s, card(B) = t, card(AT) = m, $A = \{a_1, a_2, \ldots, a_s\}$, $B = \{b_1, b_2, \ldots, b_t\}$, $AT = \{c_1, c_2, \ldots, c_m\}$, $X \subseteq U$, $A^{\beta_1, \beta_2, \ldots, \beta_s}$ is a reduction of ζ for u with respect to the given $a_1, a_2, \ldots, a_m$ in ζ, if and only if:

$u \in X, \forall X \subseteq U, \underline{APR}_{\sum_i^s b_i}(X) = \underline{APR}_{\sum_i^m c_i}(X)$ and $\forall B \subset A, u \in X, \exists X \subseteq U, \underline{APR}_{\sum_i^t b_i}(X) \neq \underline{APR}_{\sum_i^m c_i}(X)$

Example 3: We can consider the interval-valued information system given in Table 1 and assume the maximal intersection rates $\beta_1, \beta_2, \ldots, \beta_m = 0.9$ are given with respect to all the attributes. It is easy to obtain the reductions:

$A^{0.9} = (a_1, a_3, a_4, a_5)$
$A_1^{0.9}(u_1) = \{a_1, a_3\}, A_2^{0.9}(u_1) = \{a_1, a_5\}$
$A_1^{0.9}(u_2) = \{a_1\}, A_2^{0.9}(u_2) = \{a_2\}$
$A_1^{0.9}(u_3) = \{a_2\}, A_2^{0.9}(u_3) = \{a_5\}$
$A^{0.9}(u_4) = \{a_3\}$
$A^{0.9}(u_5) = \{a_1\}$
$A^{0.9}(u_6) = \{a_4\}$
$A^{0.9}(u_7) = \{a_5\}$
$A^{0.9}(u_8) = \{a_1\}$
$A^{0.9}(u_9) = \{a_1\}$
$A_1^{0.9}(u_{10}) = \{a_1\}, A_2^{0.9}(u_{10}) = \{a_3\}, A_3^{0.9}(u_{10}) = \{a_4\}$,
$A_4^{0.9}(u_{10}) = \{a_5\}$

CONCLUSION

In this study, we extend the multi-granulations rough set theory to interval-valued information system. The extension of classical rough set model is an important direction of research. We proposed the rough set model based on multi-granulation in interval-valued information system and discussed some properties on the model. The reduction is defined with the new model. In order to substantiate the conceptual arguments numerical examples are given.

REFERENCES

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

Miao, D., N. Zhang, X. Yue, 2009. Know reduction in interval-valued information systems. Proceedings of the 8th IEEE International Conference on Cognitive Informatics, June 15-17, 2009, Kowloon, Hong Kong, pp: 320-327.

Miao, D.Q. and W. Yang, 2010. Generalized relation based knowledge discovery in interval-valued information systems. Proceedings of the IEEE International Conference on Granular Computing, August 14-16, 2010, USA., pp: 743-748.

Pawlak, Z., 1982. Rough sets. Int. J. Comput. Inform. Sci., 11: 341-356.

Qian, Y.H. and J.Y. Liang, 2006. Rough set method based on Multi-granulations. Proceedings of the 5th IEEE International Conference on Cognitive Informatics, Volume 1, July 17-19, 2006, Beijing, China, pp: 297-304.

Qian, Y.H., J.Y. Liang and C.Y. Dang, 2008. Interval ordered information systems. Comput. Math. Appli., 56: 1994-2009.

Qian, Y., J. Liang, Y. Yao and C. Dang, 2010. MGRS: A Multi-granulation rough set. Inform. Sci., 180: 949-970.