

Application of Binary System to Information Table Based on Rough Set

Momo Li¹, Yu Shang¹, Xiaoyong Guo², Junli Lu¹, Yanbin Feng¹, Yanqing Duan^{3*}

1.School of Mathematics and Computer Science
Yunnan Nationalities University, Kunming 650031, P.R.China

2.Department of Mathematics and Science
Lincang Teachers' College, Lincang, 677000, P.R.China

3.Hongyun Honghe Tobacco (Group) Co., Ltd.,
Kunming 650202

*Corresponding author: E-mail: dyanqing@yahoo.com.cn

Abstract—First the traditional objects-attribute information table converted into a objects-attribute-attribute value table by using binary. Then two new concept is defined in this paper: concept vector instead of elementary sets (the equivalence classes of R) and knowledge matrix instead of equivalence relation of R by using binary system. Put forward a new solution to the lower and upper approximations support subset on rough set by using concept vector and knowledge matrix. Examples also are given.

Keywords-binary; information table; concept vector; knowledge matrix

I. INTRODUCTION

The problem of imperfect knowledge has been tackled for a long time by philosophers, logicians and mathematicians. Recently it became also a crucial issue for computer scientists, particularly in the area of artificial intelligence. There are many approaches to the problem of how to understand and manipulate imperfect knowledge. Rough Set Theory, as a new kind of data analysis tool for handling was introduced by Polish scientist Pawlak in 1982 is one attempt to this problem. The theory has attracted attention of many researchers and practitioners all over the world, who contributed essentially to its development and applications. Rough set theory has found many interesting applications. The rough set approach seems to be of fundamental importance to AI and cognitive sciences, especially in the areas of machine learning, knowledge acquisition, decision analysis, knowledge discovery from databases, expert systems, inductive reasoning and pattern recognition. The main advantage of rough set theory in data analysis is that it does not need any preliminary or additional information about data.

Rough set theory is based on the assumption that we have some additional information (knowledge, data) about elements of a universe of discourse. Elements that exhibit the same information are indiscernible (similar) and form blocks that can be understood as elementary granules of knowledge about the universe. These granules are called elementary sets (concepts), or and can be considered as elementary building blocks of knowledge. For the application of the rough set theory on reasoning from data, how fast and accurate solution each kind of rough set algorithm is quite important.

II. APPLYING BINARY TO THE INFORMATION TABLE AND DECISION TABLE

Data are often presented as a table, columns of which are labeled by attributes, rows by objects of interest and entries of the table are attribute values. Such tables are known as information systems, attribute-value tables or information tables. If we distinguish in an information table two classes of attributes, called condition and decision (action) attributes. This information table is called decision table. Each row of a decision table determines a decision rule, which specifies decisions (actions) that should be taken when conditions pointed out by condition attributes are satisfied. In the book *Rough Set Theory and Method*, all kinds of algorithms on decision table of rough set have been given, for example the simple classification algorithm P, approximate algorithm on L, lower approximation algorithm R, and intersection set algorithm I, etc. The calculation is essentially based on the intersection, union, complement of set, and equivalence relation. The calculation of the intersection of set and dependency of Attributes are very large, each element of set will frequently calculate with the others. Here we find that for a decision table or information table, assumed that we studied objects and attributes is relatively fixed, and attribute value is in limited within the scope of the each attribute, each object only a corresponding property values. Each object has only one attribute value on the corresponding attribute. That means, to a specific property value, the object is only have two probabilities: one is equal to this value and the other is not equal. Thus we can use one and zero to represent these two probabilities.

For example, here is part of the table CTR(Car Test Results)which published in "Popular Science".

TABLE I. PART OF THE CTR

object s	condition attribute									decision attribute y
	x1	x2	x3	x4	x5	x6	x7	x8	x9	
u1	c	6	y	E	m	h	h	a	m	m
u2	c	6	n	E	m	m	h	ma	m	m
u3	c	6	n	E	m	h	h	ma	m	m

u4	c	4	y	E	m	h	h	ma	l	h
u5	c	6	n	E	m	m	m	ma	m	m
u6	c	6	n	B	m	m	m	a	he	lo
u7	c	6	n	E	m	m	h	ma	he	lo
u8	s	4	n	B	sm	h	lo	ma	l	h
u9	c	4	n	B	sm	h	lo	ma	m	m
u10	c	4	n	B	sm	h	m	a	m	m

In Table 1, x1 is the condition attribute, have two values (c, s), so $U/x1 = \{V11, V12\}$, including $V11 = \{u1, u2, u3, u4, u5, u6, u7, u9, u10\}$, $V12 = \{u8\}$, if converts binary is expressed as:

$V11 = \{111111011\}$, $V12 = \{000000100\}$. In addition, if the attribute have only two values one can converse by another, just like V12 is equal to the converse by V11.

Thus, the Table 1 transfers into the Table 2.

TABLE II. THE OBJECT-ATTRIBUTE-ATTRIBUTE VALUE TABLE OF CTR

attribute	attribute value	objects									
		u1	u2	u3	u4	u5	u6	u7	u8	u9	u10
x1	c	1	1	1	1	1	1	1	0	1	1
	s	0	0	0	0	0	0	0	1	0	0
x2	4	0	0	0	1	0	0	0	1	1	1
	6	1	1	1	0	1	1	1	0	0	0
x3	y	1	0	0	1	0	0	0	0	0	0
	n	0	1	1	0	1	1	1	1	1	1
x4	E	1	1	1	1	1	0	1	0	0	0
	B	0	0	0	0	0	1	0	1	1	1
x5	m	1	1	1	1	1	1	1	0	0	0
	sm	0	0	0	0	0	0	0	1	1	1
x6	h	1	0	1	1	0	0	0	1	1	1
	m	0	1	0	0	1	1	1	0	0	0
x7	h	1	1	1	1	0	0	1	0	0	0
	m	0	0	0	0	1	1	0	0	0	1
x8	lo	0	0	0	0	0	0	0	1	1	0
	a	1	0	0	0	0	1	0	0	0	1
x9	ma	0	1	1	1	1	0	1	1	1	0
	m	0	0	0	0	0	0	0	0	0	0
x10	l	0	0	0	0	0	0	0	0	0	0
	he	0	0	0	0	0	0	0	0	0	0
y	m	1	1	1	0	1	0	0	0	1	1
	h	0	0	0	1	0	0	0	1	0	0
z	lo	0	0	0	0	0	1	1	0	0	0

III. CONCEPT VECTOR AND KNOWLEDGE MATRIX

From the transferred Table 2 we can see, each one row representing a equivalence relationship. In order to better

describe such a combination, we call each row as concept vector and all concept vectors about specific attribute are called as knowledge matrix. And a knowledge base consists of all knowledge matrices. As the number of the objects and the attributes are the same, the rank and the columns of the concept vector and the knowledge matrix are also the same. Thus a information or decision table are transfer to a table only expressed by one and zero. The following calculation are based on this one zero table. Because the bit operation is the fastest way as to computers' process speed, the algorithms will be far more simplified. We can use the bit and operator and or operator instead of calculating intersection and union of the set. For example, in the Table 1, the concept vector of $U/x2=\{V11, V22\}$, $U/x3=\{V31, V32\}$,..., $U/y=\{W11, W12, W13\}$, is the following:

$$V11 = (1110111000), V22 = (0001000111), V31 = (1001000000), V32 = (0110111111)$$

$$W11 = (1110100011), W12 = (0001000100), W13 = (0000011000).$$

The knowledge base is just consists of all knowledge matrixes just like block matrix.

$$s = \begin{pmatrix} 1111111011 \\ 000000100 \\ 1110111000 \\ 0001000111 \\ \dots \dots \\ 1110100011 \\ 0001000100 \\ 0000011000 \end{pmatrix} = \begin{pmatrix} V1 \\ V2 \\ \dots \dots \\ W \end{pmatrix}$$

Now, we can simplify our calculation on information or decision table based on the matrix.

3.1 Applying concept vector on calculate the support subset

Set W as a division on U, for classification U/a, we define the support subset of is

$W^{(U/a)} = \cup_{V \in U/a, V \subseteq W} V$, symbolically $S_a(W)$. $S_a(W)$ is regard as the lower approximation support subset, the degree of

$S_a(W)$ is : $spt_a W = \frac{1}{|U|} (S_a(W))$. The lower approximation support subset for classification U/a is defined as

$$W^{(U/a)} = \cup_{V \in U/a, V \cap W \neq \emptyset} V$$

$$acc_a(W) = \frac{1}{|W^{(U/a)}|} (|W^{(U/a)}|)$$

is the approximation degree of

the attribute a to W. To calculate the $W^{(U/a)} = \cup_{V \in U/a, V \subseteq W} V$

based on concept vector is just one concept vectors operate and with the others. If we get the same concept vector as the classification U/a, this concept vector is belong to the support subset of W.

An example will be given below:

$$U/a = \{V1, V2, V3\} = \{\{u1, u2, u3\}, \{u4, u5\}, \{u6, u7, u8\}\}$$

transfer to the concept vector: $V1=(11100000), V2=(00011000), V3=(00001111)$.

$U/w = \{W1, W2, W3, W4\} = \{\{u1, u2\}, \{u3\}, \{u4, u5\}, \{u6, u7, u8\}\}$, transfer to the concept vector: $W1 = \{11000000\}$, $W2 = \{00100000\}$, $W3 = \{00011000\}$, $W4 = \{00000111\}$,

because

$V1 \& W1 \neq V1$, $V1 \& W2 \neq V1$, $V1 \& W3 \neq V1$, $V1 \& W4 \neq V1$, $V1$ is not the support subset of W ; Similarly, $V2 \& W1 \neq V2$, $V2 \& W2 \neq V2$, $V2 \& W3 = V2$, $V2 \& W4 \neq V2$, so $V2$ is. Also,

$V3 \& W1 \neq V3$, $V3 \& W2 \neq V3$, $V3 \& W3 = V3$, $V3$ is also the subset of W .

The same way we also can use concept subset to simplification calculation of the upper approximation support subset. If the one concept vectors operate and with the others' result is not zero vector, the concept vector is belong to the upper approximate support subset.

3.2 Applying knowledge matrix on multi-conditions support subset

If we will calculate the multi-condition support subset, we will use knowledge matrix to simplify the calculation. Here we use the concept vector as mentioned above $U/x1$ and $U/x2$: $V11, V12, V21, V22$.

$$\begin{aligned} V11 &= (111111101100100111011) & V12 &= (000000010011011000100) \\ V21 &= (111011100000000010000) & V22 &= (00010001111111101111) \end{aligned}$$

They are consisted of $V1$ and $V2$:

$$\begin{aligned} V1 &= \begin{pmatrix} 111111101100100111011 \\ 000000010011011000100 \end{pmatrix}, \\ V2 &= \begin{pmatrix} 111011100000000010000 \\ 00010001111111101111 \end{pmatrix} \end{aligned}$$

Then solving $U/x1x2$ is to operate and with the two matrixes. And the matrix operator *and* is similar to the matrix multiplication, each line of $V1$ respectively operates *and* with all lines of $V2$. Eventually we get one matrix with four lines. Here is an example ($V2'$ is the transpose of $v2$):

$$\begin{aligned} V1 \& V2' &= \begin{pmatrix} 1111111011 & 0010011101 & 1 & 0 \\ 0000000100 & 1101100010 & 0 & 0 \end{pmatrix} \& \begin{pmatrix} 1110111000 & 0000001000 & 0 \\ 0001000111 & 1111110111 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1110111000 & 0000001000 & 0 \\ 0001000011 & 0010010101 & 1 \\ 0000000000 & 0000000000 & 0 \\ 0000000100 & 1101100010 & 0 \end{pmatrix} \end{aligned}$$

Thus, $U/x1x2 = \{\{u1, u2, u3, u5, u6, u7, u17\}, \{u4, u9, u10, u13, u16, u18, u2, u21\}, \{u8, u11, u12, u14, u15, u19\}\}$.

IV. CONCLUSIONS

When using the concept vector and knowledge matrix to represent equivalence class and knowledge base, we find a great of advantage on the simplified calculation. In this way for a fixed decision table will be converted into a corresponding binary knowledge matrix. Thus simplify calculation on information or decision table. However when the table is to belong to dynamic update, or once you add a new attribute value or element, that can change vector column rank and knowledge base's row rank, we have to calculation once again. Whether we can realize the incremental calculation based on the previously result is depended on the following further study in the future.

REFERENCES

- [1] Z.Pawlak. Rough sets[J].International Journal of Computer and Information Sciences, 1, 1982,pp.341-356.
- [2] W.H. Xu, X.Y. Zhang, J.M Zhong and W.X. Zhang. Attribute reduction in ordered information systems based on evidence theory. *Knowledge and Information Systems*. 2010, 25: 69-184
- [3] Shaobo Deng, Min Li. The simple-discernibility matrix in rough sets. Proc. SPIE 8349, *Fourth International Conference on Machine Vision (ICMV 2011): Machine Vision, Image Processing, and Pattern Analysis*, 834904 (January 11, 2012).
- [4] HaiTao Wu. A New Discernibility Matrix Based on Distribution Reduction. Proceedings of the International Symposium on Intelligent Information Systems and Applications (IISA'09). 2009, pp. 390-393.
- [5] N. Zhong, A. Skowron, and S. Ohsuga, Eds. New Direction in Rough Sets, Data Mining, and Granular-Soft Computing. *Springer*, 1999.
- [6] L. Polkowski and A. Skowron, Eds., Rough Sets and Current Trends in Computing. Lecture Notes in Artificial Intelligence 1424, *Springer*, 1998.
- [7] L. Polkowski, S. Tsumoto, and T. Y. Lin, Eds., Rough Set Methods and Applications – New Developments in Knowledge Discovery in Information Systems. *Springer*, 2000.
- [8] Z.Pawlak. Rough Sets -- Theoretical Aspects of Reasoning about Data[M]. Boston: Kluwer *Academic Publishers*, 1991.
- [9] Iwinski T B. Algebraic approach to rough sets[J]. *Bulletin of Polish Academy of Science, Mathematics*, 1987, 35: 673-683.
- [10] Hu X. Knowledge Discovery in Database: an attribute-oriented rough set Approach[D]. *University of Regina, Canada*, 1995.
- [11] L. Polkowski and A. Skowron, Eds., Rough Sets in Knowledge Discovery. Vol. 1–2, *Springer*, 1998.
- [12] L. Polkowski: Rough Sets, Mathematical Foundations, Advances in Soft Computing, Physica – Verlag, *A Springer-Verlag Company*, 2002
- [13] A. Skowron et al: Rough set perspective on data and knowledge, Handbook of Data Mining and Knowledge Discovery (W. Klösgen, J. Żytkow eds.), *Oxford University Press*, 2002, 134-149.
- [14] Felix R Ushio T. Rough Sets - based Machine Learning Using a Binary Discernibility Matrix[c]. [s.l.]: *IPMM*, 1999. 299-305.