Fundamentals of Rough Set Theory and Its Applications

C.-C. Chan
Department of Computer Science
University of Akron
Akron, OH 44325-4003
USA
chan@cs.uakron.edu
And
Department of Information Science and Applications
Asia University
500 Liufeng Rd.
WuFeng, Taichung, 41354
Taiwan
chan@asia.edu.tw

Outline

- Basic Concepts of Rough Sets
  - Relationship to Fuzzy Sets
- Rough Multi-Sets and Multi-Decision Tables
  - Relationship to Dempster-Shafer's Theory of Evidence
- Rough Sets and Bayes' Theorem
- Learning Rules from Examples
  - LEM2
  - RLEM2
  - BLEM2
- Software Tools
- Concluding Remarks

Introduction

- Rough set theory was introduced by Zdzislaw Pawlak in 1982
- Representative Publications:
  - Z. Pawlak, "Rough Sets", *International Journal of Computer and Information Sciences*, Vol.11, 341-356 (1982).
  - Z. Pawlak, *Rough Sets - Theoretical Aspect of Reasoning about Data*, Kluwer Academic Pubilishers (1991).
- RST is useful for reasoning about knowledge of objects represented by attributes (features).
- Fundamental Assumptions:
  - Objects are represented by values of attributes.
  - Objects with the same information are indiscernible.

Basic Concepts

- Approximation Space

  An approximation space is a pair (U, R) where

  U is a nonempty finite set called the universe and

  R is an equivalence relation defined on U.

- Information System

  An information system is a pair $\boldsymbol{S} = (U, A)$, where

  U is a nonempty finite set called the universe and

  A is a nonempty finite set of attributes, i.e., $a: U \rightarrow V_a$ for $a \in A$, where V_a is called the domain of a.

- Decision Table (Data Table)

  A decision table is a special case of information systems,

  $S = (U, A = \mathrm{C} \cup \{d\})$,

  where attributes in C are called **condition attributes** and d is a designated attribute called the **decision attribute**.

Table 1. Example of Decision Table.

STUDENT	Category	Major	Birth_Place	Grade
1	PhD	History	Detroit	A
2	MS	Chemistry	Akron	A
3	MS	History	Detroit	C
4	BS	Math	Detroit	B
5	BS	Chemistry	Akron	C
6	PhD	Computing	Cleveland	A
7	BS	Chemistry	Cleveland	C
8	PhD	Computing	Akron	A

$\overline{R}$

Approximations of Sets

Let S = (U, R) be an approximation space and X be a subset of U.

The lower approximation of X by R in S is defined as

$\underline{R}X = \{ e \in U \mid [e] \subseteq X\}$ and

The upper approximation of X by R in S is defined as

$\overline{R}X = \{ e \in U \mid [e] \cap X \neq \emptyset\}$

where [e] denotes the equivalence class containing e.
[e] is called elementary set.

A subset X of U is said to be R-definable in S

if and only if $\underline{R}X = \overline{R}X$

The boundary set $BN_R(X)$ is defined as $\overline{R}X - \underline{R}X$

A set X is rough in S if its boundary set is nonempty.

Accuracy of Approximations

$$\alpha_B(X) = \frac{|\underline{B}(X)|}{|\overline{B}(X)|}$$

where S = (U, A), B $\subseteq$ A and X $\subseteq$ U
|X| denotes the cardinality of X

If $\alpha_B(X) = 1$, *then X* is *crisp* with respect to *B*.

If $\alpha_B(X) < 1$, then *X* is *rough* with respect to *B.*

U	Category	Major	Birth_Place	Grade
1	PhD	History	Detroit	A
2	MS	Chemistry	Akron	A
3	MS	History	Detroit	C
4	BS	Math	Detroit	B
5	BS	Chemistry	Akron	C
6	PhD	Computing	Cleveland	A
7	BS	Chemistry	Cleveland	C
8	PhD	Computing	Akron	A

U\{Category} = {{1, 6, 8}, {2, 3}, {4, 5, 7}}
U\{Major} = {{1, 3}, {2, 5, 7}, {4}, {6, 8}}
U\{Birth_Place} = {{2, 5, 8}, {1, 3, 4}, {6, 7}}
U\{Grade} = {{1, 2, 6, 8}, {4}, {3, 5, 7}}

Let X = [(Grade, A)] = {1, 2, 6, 8}

Let B = {Major, Birth_Place}

U\B= {{1, 3}, {2, 5}, {4}, {6}, {7}, {8}}

$\underline{B}(X) = \{6, 8\}$

$\overline{B}(X) = \{1, 2, 3, 5, 6, 8\}$

$BN_B(X) = \{1, 2, 3, 5\}$

Accuracy:
$\alpha_B(X) = 2/6 = 1/3$

Dependency of Attributes

Let *C* and *D* be subsets of *A*. We say that *D* depends on *C* in a degree *k* ($0 \leqq k \leqq 1$) denoted by $C \rightarrow_k D$ if

$$k = \gamma(C, D) = \sum_{X \in U/D} \frac{|\underline{C}(X)|}{|U|}.$$

$$k = \gamma(C, D) = \frac{|POS_C(D)|}{|U|}$$

where $POS_C(D) = \bigcup_{X \in U/D} \underline{C}(X),$ called *C*-positive region of *D*.

If $k = 1$ we say that *D* depends totally on *C*.
If $k < 1$ we say that *D* depends partially (in a degree *k*) on *C*.

U	Category	Major	Birth_Place	Grade
1	PhD	History	Detroit	A
2	MS	Chemistry	Akron	A
3	MS	History	Detroit	C
4	BS	Math	Detroit	B
5	BS	Chemistry	Akron	C
6	PhD	Computing	Cleveland	A
7	BS	Chemistry	Cleveland	C
8	PhD	Computing	Akron	A

U\{Category} = {{1, 6, 8}, {2, 3}, {4, 5, 7}}
U\{Major} = {{1, 3}, {2, 5, 7}, {4}, {6, 8}}
U\{Birth_Place} = {{2, 5, 8}, {1, 3, 4}, {6, 7}}
U\{Grade} = {{1, 2, 6, 8}, {4}, {3, 5, 7}}

C = {Major, Birth_Place}

D = {Grade}

U\C = {{1, 3}, {2, 5}, {4}, {6}, {7}, {8}}

U\D = {{1, 2, 6, 8}, {4}, {3, 5, 7}}

$\underline{C}(\{1, 2, 6, 8\}) = \{6, 8\}$, $\underline{C}(\{4\}) = \{4\}$, $\underline{C}(\{3, 5, 7\}) = \{7\}$

$C \rightarrow_k D = |POS_C(D)| / |U| = 4/8 = 1/2$

Dispensable and Indispensable Attributes

Let S = (*U*, *A* = C ∪ D) be a decision table.

Let c be an attribute in C.

Attribute c is dispensable in S if $POS_C(D) = POS_{(C-\{c\})}(D)$

Otherewise, c is indispensable.

A decision table S is independent if all attributes in C are indispensable.

Reducts and Core

Let S = (*U*, *A* = C $\cup$ D) be a decision table.

A subset R of C is a reduct of C, if

$POS_R(D) = POS_C(D)$ and

S' = (U, R $\cup$ D) is independent, i.e., all attributes in R are indispensable in S'.

Core of C is the set of attributes shared by all reducts of C.

$$CORE(C) = \cap RED(C)$$

where RED(C) is the set of all reducts of C.

U	Category	Major	Birth_Place	Grade
1	PhD	History	Detroit	A
2	MS	Chemistry	Akron	A
3	MS	History	Detroit	C
4	BS	Math	Detroit	B
5	BS	Chemistry	Akron	C
6	PhD	Computing	Cleveland	A
7	BS	Chemistry	Cleveland	C
8	PhD	Computing	Akron	A

C = {Category, Major, Birth_Place}

D = {Grade}

U\C = {{1}, {2}, {3}, {4}, {5}, {6}, {7}, {8}}

C1 = {Category, Major}, C2 = {Category, Birth_Place}

C3 = {Major, Birth_Place}

U\C1 = {{1}, {2}, {3}, {4}, {5, 7}, {6, 8}}

U\C2 = {{1}, {2}, {3}, {4}, {5}, {6}, {7}, {8}}

U\C3 = {{1, 3}, {2, 5}, {4}, {6}, {7}, {8}}

U\{Category} = {{1, 6, 8}, {2, 3}, {4, 5, 7}}
U\{Major} = {{1, 3}, {2, 5, 7}, {4}, {6, 8}}
U\{Birth_Place} = {{2, 5, 8}, {1, 3, 4}, {6, 7}}
U\{Grade} = {{1, 2, 6, 8}, {4}, {3, 5, 7}}

$POS_C(D)$ = {1, 2, 3, 4, 5, 6, 7, 8}

$POS_{C1}(D)$ = {1, 2, 3, 4, 5, 6, 7, 8}

$POS_{C2}(D)$ = {1, 2, 3, 4, 5, 6, 7, 8}

$POS_{c3}(D)$ = {4, 6, 7, 8}

C1 and C2 are reducts of C.

C3 is not a reduct of C.

The Core of C is {Category}

Rough Membership Function [Pawlak & Skowron, 1994]

Let S = (U, A), $B \subseteq A$ and $X \subseteq U$

Then the rough membership function μ^B_X for X is a mapping from U to [0, 1],

$$\mu^B_X : U \rightarrow [0, 1]$$

For all e in U, the degree of e belongs to X in light of the set of attributes B is defined as

$$\mu^B_X(e) = |B(e) \cap X| / |B(e)|$$

where B(e) denotes the block containing e.

Properties of rough membership function:

P1: $\mu^B_X(e) = 1$ iff e in $B_*(X)$

P2: $\mu^B_X(e) = 0$ iff e in $U - B^*(X)$

P3: $0 < \mu^B_X(e) < 1$ iff e in $BN_B(X)$

P4: $\mu^B_{U-X}(e) = 1 - \mu^B_X(e)$ for any e in U

P5: $\mu^B_{X\cup Y}(e) \geqq \max(\mu^B_X(e), \mu^B_y(e))$ for any e in U

P6: $\mu^B_{X\cap Y}(e) \leqq \min(\mu^B_X(e), \mu^B_y(e))$ for any e in U

where $B_*(X)$ is the lower approximation of X in B and $B^*(X)$ is the upper approximation of X in B.

Fuzzy Sets and Rough Sets:

Let U be a domain of objects.
A fuzzy set X defined on U is characterized by a membership function μ_X:

$\mu_X: U \rightarrow [0, 1]$

Let A and B be two fuzzy sets, and

$\mu_{A \cap B} = \min(\mu_A, \mu_B)$

$\mu_{A \cup B} = \max(\mu_A, \mu_B)$

Let S = (U, R) be an approximation space and X be a subset of U.
Define

$$\mu_X(e) = \begin{cases} 1 & \text{if } e \text{ in } \underline{R}X \\ 1/2 & \text{if } e \text{ in } BN_R(X) \\ 0 & \text{if } e \text{ in } -\overline{R}X \end{cases}$$

where –X is the complement of X.

Then, the rough membership function cannot be extended to the fuzzy union and intersection of sets.

In general:

$$\underline{R}(X \cup Y) \supseteq \underline{R}X \cup \underline{R}Y \qquad \text{and} \qquad \overline{R}(X \cap Y) \subseteq \overline{R}X \cap \overline{R}Y.$$

The rough membership function will reduce to fuzzy set when

$$\underline{R}(X \cup Y) = \underline{R}X \cup \underline{R}Y \qquad \text{and} \qquad \overline{R}(X \cap Y) = \overline{R}X \cap \overline{R}Y$$

Rough MultiSets and MultiSet Decision Tables

Related Concepts:
Rough Sets and Information Systems [Pawlak, 1982]
Rough MultiSets and Information Multisystems [Grzymala-Busse, 1987]
Multiset Decision Tables [Chan 2001, 2004]

- Pawlak, Z., "Rough sets: basic notion," *Int. J. of Computer and Information Science* 11, 344-56, (1982).

- Grzymala-Busse, J.W., "Learning from examples based on rough multisets," *Proc. of the 2nd Int. Symposium on Methodologies for Intelligent Systems*, Charlotte, North Carolina, October 14-17, 325-332, (1987).

- Chan, C.-C., "Distributed incremental data mining from very large databases: a rough multiset approach," *Proc. the 5th World Multi-Conference on Systemics, Cybernetics and Informatics, SCI 2001*, Orlando, Florida, July 22-25, (2001), 517-522.
- Chan, C.-C., "Learning rules from very large databases using rough multisets," *LNCS Transactions on Rough Sets*, (J. Peters, A. Skowron, J.W. Grzymala-Busse, B. Kostek, R.W. Swiniarski, M. Szczuka Eds.), Vol. 1, pp. 59 – 77, Springer –Verlag Berlin Heidelberg, 2004.

U	A	B	C	E	F	D
1	0	0	1	0	0	1
2	1	1	1	0	0	1
3	0	1	0	0	0	2
4	1	0	0	0	1	1
5	0	1	0	0	0	2
6	1	0	0	0	1	3
7	0	0	0	1	1	3
8	1	1	1	1	1	1
9	0	0	0	1	1	3
10	0	0	1	0	0	1
11	1	1	1	0	0	2
12	1	1	1	1	1	2
13	1	1	0	1	1	3
14	1	1	0	0	1	3
15	0	0	1	1	1	1
16	1	1	0	1	1	2
17	0	0	0	1	1	3
18	0	0	0	0	0	2
19	0	0	0	0	0	2
20	1	1	1	0	0	3
21	1	1	0	0	1	2
22	0	0	1	0	1	1
23	1	1	1	0	0	2
24	0	0	1	1	1	2
25	1	0	1	0	1	1
26	1	0	1	0	1	3
27	1	0	1	0	1	2
28	1	1	1	1	0	3

A	B	C	E	F	D	W
0	0	0	0	0	2	**2**
0	0	0	1	1	3	**3**
0	0	1	0	0	1	**2**
0	0	1	0	1	1	**1**
0	0	1	1	1	1	**1**
0	0	1	1	1	2	**1**
0	1	0	0	0	2	**2**
1	0	0	0	1	1	**1**
1	0	0	0	1	3	**1**
1	0	1	0	1	1	**1**
1	0	1	0	1	2	**1**
1	0	1	0	1	3	**1**
1	1	0	0	1	2	**1**
1	1	0	0	1	3	**1**
1	1	0	1	1	2	**1**
1	1	0	1	1	3	**1**
1	1	1	0	0	1	**1**
1	1	1	0	0	2	**2**
1	1	1	0	0	3	**1**
1	1	1	1	0	3	**1**
1	1	1	1	1	1	**1**
1	1	1	1	1	2	**1**

The projection of the multirelation onto the set P of attributes $\{A, B, C, E, F\}$ is shown in Table 3.

Table 3. An information multisystem

A	B	C	E	F	W
0	0	0	0	0	2
0	0	0	1	1	3
0	0	1	0	0	2
0	0	1	0	1	1
0	0	1	1	1	2
0	1	0	0	0	2
1	0	0	0	1	2
1	0	1	0	1	3
1	1	0	0	1	2
1	1	0	1	1	2
1	1	1	0	0	4
1	1	1	1	0	1
1	1	1	1	1	2

Table 4. A sub-multiset X of $\tilde{P}$.

A	B	C	E	F	W
0	0	1	0	0	2
1	1	1	0	0	1
1	0	0	0	1	1
1	1	1	1	1	1
0	0	1	1	1	1
0	0	1	0	1	1
1	0	1	0	1	1

Table 5. P-lower approximation of X.

A	B	C	E	F	W
0	0	1	0	0	2
0	0	1	0	1	1

Table 6. P-upper approximation of X.

A	B	C	E	F	W
0	0	1	0	0	2
1	1	1	0	0	4
1	0	0	0	1	2
1	1	1	1	1	2
0	0	1	1	1	2
0	0	1	0	1	1
1	0	1	0	1	3

The classification of P generated by attribute D in S:

consists of three sub-multisets which are given Tables 7, 8, and 9.

Table 7. Sub-multiset of the multipartition D_P with $D = 1$.

A	*B*	*C*	*E*	*F*	*W*
0	0	1	0	0	2
0	0	1	0	1	1
0	0	1	1	1	1
1	0	0	0	1	1
1	0	1	0	1	1
1	1	1	0	0	1
1	1	1	1	1	1

Table 8. Sub-multiset of the multipartition D_P with $D = 2$.

A	*B*	*C*	*E*	*F*	*W*
0	0	0	0	0	2
0	0	1	1	1	1
0	1	0	0	0	2
1	0	1	0	1	1
1	1	0	0	1	1
1	1	0	1	1	1
1	1	1	0	0	2
1	1	1	1	1	1

Table 9. Sub-multiset of the multipartition D_P with $D = 3$

A	*B*	*C*	*E*	*F*	*W*
0	0	0	1	1	3
1	0	0	0	1	1
1	0	1	0	1	1
1	1	0	0	1	1
1	1	0	1	1	1
1	1	1	0	0	1
1	1	1	1	0	1

Multiset Decision Table

A	B	C	E	F	W	D_1	D_2	D_3	w_1	w_2	w_3
0	0	0	0	0	**2**	0	1	0	0	2	0
0	0	0	1	1	**3**	0	0	1	0	0	3
0	0	1	0	0	**2**	1	0	0	2	0	0
0	0	1	0	1	**1**	1	0	0	1	0	0
0	0	1	1	1	**2**	1	1	0	1	1	0
0	1	0	0	0	**2**	0	1	0	0	2	0
1	0	0	0	1	**2**	1	0	1	1	0	1
1	0	1	0	1	**3**	1	1	1	1	1	1
1	1	0	0	1	**2**	0	1	1	0	1	1
1	1	0	1	1	**2**	0	1	1	0	1	1
1	1	1	0	0	**4**	1	1	1	1	2	1
1	1	1	1	0	**1**	0	0	1	0	0	1
1	1	1	1	1	**2**	1	1	0	1	1	0

Condition attributes $C = \{A, B, C, E, F\}$ and decision attribute D.

The Boolean vector is denoted by [$D1$, $D2$, $D3$], and the integer vector is denoted by [$w1$, $w2$, $w3$]. Note that $W = w1 + w2 + w3$ on each row.

Table 11. *C*-lower approximation of *D1.*

A	B	C	E	F	W
0	0	1	0	0	2
0	0	1	0	1	1

Table 12. *C*-upper approximation of D_1.

A	B	C	E	F	W
0	0	1	0	0	2
0	0	1	0	1	1
0	0	1	1	1	2
1	0	0	0	1	2
1	0	1	0	1	3
1	1	1	0	0	4
1	1	1	1	1	2

To determine the partition of boundary multisets, we use the following two steps.

Step 1. Identify rows with $D1 + D2 + D3 > 1$, we have the following multiset in table form:

Table 13. Elements in the boundary sets.

A	B	C	E	F	W	D_1	D_2	D_3
0	0	1	1	1	**2**	1	1	0
1	0	0	0	1	**2**	1	0	1
1	0	1	0	1	**3**	1	1	1
1	1	0	0	1	**2**	0	1	1
1	1	0	1	1	**2**	0	1	1
1	1	1	0	0	**4**	1	1	1
1	1	1	1	1	**2**	1	1	0

To determine the partition of boundary multisets, we use the following two steps.

Step 1. Identify rows with *D1* + *D2* + *D3* > *1*, we have the following multiset in table form:

Table 13. Elements in the boundary sets.

A	*B*	*C*	*E*	*F*	*W*	D_1	D_2	D_3
0	0	1	1	1	**2**	1	1	0
1	0	0	0	1	**2**	1	0	1
1	0	1	0	1	**3**	1	1	1
1	1	0	0	1	**2**	0	1	1
1	1	0	1	1	**2**	0	1	1
1	1	1	0	0	**4**	1	1	1
1	1	1	1	1	**2**	1	1	0

Step 2. Grouping the above table in terms of *D1*, *D2*, and *D3*, we have the following blocks in the partition.

A	***B***	***C***	***E***	***F***	***W***	D_1	D_2	D_3
0	0	1	1	1	**2**	1	1	0
1	1	1	1	1	**2**	1	1	0

A	***B***	***C***	***E***	***F***	***W***	D_1	D_2	D_3
1	0	0	0	1	2	1	0	1

A	***B***	***C***	***E***	***F***	***W***	D_1	D_2	D_3
1	1	0	0	1	2	0	1	1
1	1	0	1	1	2	0	1	1

A	*B*	*C*	*E*	*F*	*W*	D_1	D_2	D_3
1	0	1	0	1	3	1	1	1
1	1	1	0	0	4	1	1	1

Rough Sets and Dempter-Shafer's Theory of Evidence

The relationship between rough set theory and Dempster-Shafer's theory of evidence was first shown in [Grzymala-Busse, 1987] and further developed in [Skowron, A. and J. Grzymala-Busse, 1994].

The concept of partition of boundary sets was introduced in [Skowron, A. and J. Grzymala-Busse, 1994].

The basic idea is to represent an expert's classification on a set of objects in terms of lower approximations and a partition on the boundary set.

Grzymala-Busse, J.W., "Rough set and Dempster-Shafer approaches to knowledge acquisition under uncertainty - a comparison," *manuscript*, (1987).

Skowron, A. and J. Grzymala-Busse, "From rough set theory to evidence theory." in *Advances in the Dempster-Shafer Theory of Evidence*, edited by R. R. Yager, J. Kacprzyk, and M. Fedrizzi, 193-236, John Wiley & Sons, Inc, New York, (1994).

Table 18. Grouping over D_1, D_2, D_3 and sum over W.

D_1	D_2	D_3	W
1	0	0	3
0	1	0	4
0	0	1	4
0	1	1	4
1	0	1	2
1	1	0	4
1	1	1	7

Let $\Theta = \{1, 2, 3\}$.

Table 19. The bpa derived from Table 2.

X	{1}	{2}	{3}	{1, 2}	{1, 3}	{2, 3}	{1, 2, 3}
m(X)	3/28	4/28	4/28	4/28	2/28	4/28	7/28

Rough Sets and Bayes' Theorem

[Pawlak 1999, 2002]
Rough Membership Function [Pawlak & Skowron, 1994]

Let $S = (U, A)$, $B \subseteq A$ and $X \subseteq U$
Then the rough membership function $\mu_B X$ for X is a mapping from U to [0, 1],

$$\mu_B X: U \rightarrow [0, 1]$$

For all e in U, the degree of e belongs to X in light of the set of attributes B is defined as

$$\mu_B X(e) = |B(e) \cap X| / |B(e)|$$

where B(e) denotes the block containing e.

Let $C \rightarrow D$ be a decision rule and let e in U, then

Support of $C \rightarrow_e D$ is defined as

$$supp_e(C, D) = |C(e) \cap D(e)|$$

Strength of $C \rightarrow_e D$:

$$\sigma_e(C, D) = |C(e) \cap D(e)| \; / \; |U| = supp_e(C, D) \; / \; |U|$$

Certainty of $C \rightarrow_e D$:

$$cer_e(C, D) = |C(e) \cap D(e)| \; / \; |C(e)|$$

Inverse decision rule:
Let $C \rightarrow_e D$ be a decision rule, then $D \rightarrow_e C$ is an inverse decision rule of $C \rightarrow_e D$.

Coverage of $C \rightarrow_e D$: (certainty of the $D \rightarrow_e C$)

$$cove(C, D) = |D(e) \cap C(e)| \; / \; |D(e)|$$

Learning Rules from Examples

- LEM2 [Chan, 1989, 1991]
  - Basic idea is to learn rules from lower and upper approximations
  - Work in incremental and non-incremental modes
- RLEM2 [Chan, 2001]
  - Basic idea is to learn rules from Multiset Decision Tables using extended SQL operators
  - Rules are learned from lower and upper approximations
- BLEM2 [Chan, 2003]
  - Basic idea is to learn rules from lower approximation and partition of boundary set
  - Rules are associated with four factors based on Bayes' theorem and rough sets [Pawlak, 2002]

Generate rules with support, certainty, strength, and coverage
For a decision rule r, $T \rightarrow (d, v)$, derived from a decision table $(U, A=C \cup D)$, we have

support of r, $supp(r) = |[T] \cap [(d, v)]|$,
strength of r, $\sigma(r) = |[T] \cap [(d, v)]| / |U|$,
certainty of r, $cer(r) = |[T] \cap [(d, v)]| / |[T]|$, and
coverage of r, $cov(r) = |[T] \cap [(d, v)]| / |[(d,v)]|$.

where U is a finite set of examples,
C is a set of condition attributes and
D is a singleton set $\{d\}$ of decision attribute,
(d, v) is a decision-value pair,
T is a nonempty set of condition-value pairs,
$[T] = \bigcap_{t \in T} [t]$ is the block of T.

$T \rightarrow (d, v)$ if and only if $[T] \subseteq [(d, v)]$

Set T is called a ***complex*** of (d, v) when $T \rightarrow (d, v)$

Set T is a ***minimal complex*** of (d, v) when T is minimal, i.e., for any t in T, $[T - \{t\}] \not\subset [(d, v)]$

A nonempty set R of minimal complexes of (d, v) is a ***local covering*** of (d, v), if $\cup_{T \in R}$**[T] = [(d, v)] and R is minimal**

The objective is to find a local covering for each decision-value pair in a decision table.

LEM2 and BLEM2

```
procedure LEM2
inputs:  set X, which is a lower or upper approximation of a decision-value pair.
outputs:  a single local covering T of the decision-value pair.
begin
    G := X; //initial target set
    T := ∅;  //final local covering
while G ≠ ∅ do
begin
    T := ∅;  //initial complex
    T(G) := {t | [t] ∩ G ≠ ∅ };  //list of relevant a-v pairs
    while T = ∅ or not ([T] ⊆ X) do
    begin
  select a pair t in T(G) with the highest rank, if a tie occurs, select a pair t in T(G)
  such that |[t] ∩ G| is maximum; if another tie occurs, select a pair t in T(G) with the
  smallest cardinality of [t]; if a further tie occurs, select the first pair;
  T := T ∪ {t};
  G := [t] ∩ G;
  T(G) := {t | [t] ∩ G ≠ ∅ };
  T(G) := T(G) – T;
     end;  //while
    for each t in T do
       if [T – {t}] ⊆ X then T := T – {t};  //LINE 21
    T := T ∪ {T};
    G := X – ∪T ∈ T[T];
end;  //while
for each T ∈ T do
if ∪S ∈ T – {T}[S] = X then T := T – {T};
end; {procedure LEM2}
```

BLEM2:

if [T – {t}] $\subseteq$ G **then** T := T – {t}; **//LINE 21**

LEM2 generates smaller number of rules than BLEM2.

BLEM2 generates rules with support, strength, certainty, and coverage factors

Facilitate the design of Bayesian classifiers using BLEM2 rules

Experimental results showed that they are effective tools for developing efficient inference engines

Tools Based on Rough Set

ROSE/ROSE2
ROugh Set data Explorer

4eMka
Dominance-based Rough Set Approach to Multicriteria Classification

JAMM
A New Decision Support Tool for Analysis and Solving of Multicriteria Classification Problems

http://idss.cs.put.poznan.pl/site/software.html

Rough Set Exploration System

What is RSES?

RSES is a toolkit for analysis of table data running under Windows NT/95/98/2000/XP. It is based on methods and algorithms coming from the area of Rough Sets. It comprises of two general components - the GUI front-end and the computational kernel. The kernel is based on renewed RSESlib library.

Requirements:
PC with 128+ MB RAM.
3 MB of disc space + space occupied by Java VM
Windows NT4/95/98/2000/XP or Linux/i386
Java Runtime Environment (JRE) or Java SDK. We recommend the use of version 1.4.1 or higher.

The system, starting from version 2.0, is distributed as single self-installing bundle (for Windows). The installation procedure is described in User's Guide.
http://logic.mimuw.edu.pl/~rses

Computational algorithms implemented by (in alphabetical order): Jan Bazan, Rafał Latkowski, Nguyen Sinh Hoa, Nguyen Hung Son, Piotr Synak, Arkadiusz Wojna, Marcin Wojnarski and Jakub Wróblewski.

Concluding Remarks

- Classical Rough Set Theory (CRST) introduced by Pawlak has been applied to the development of learning and data reduction algorithms for data mining tasks.
- Extensions of CRST such as Dominance Based Rough Sets (DBRS) further facilitate the development of tools for Multi-Criteria Decision Analysis
- Rough Sets + Genetic Algorithms still have rooms to be developed
- Extension from rough sets to granular computing is undergoing development

Thank You!

References

Z. Pawlak, "Rough Sets", *International Journal of Computer and Information Sciences*, Vol.11, 341-356 (1982).

Z. Pawlak, *Rough Sets - Theoretical Aspect of Reasoning about Data*, Kluwer Academic Pubilishers (1991).

- Grzymala-Busse, J.W., "Rough set and Dempster-Shafer approaches to knowledge acquisition under uncertainty - a comparison," *manuscript*, (1987).
- Skowron, A. and J. Grzymala-Busse, "From rough set theory to evidence theory." in *Advances in the Dempster-Shafer Theory of Evidence*, edited by R. R. Yager, J. Kacprzyk, and M. Fedrizzi, 193-236, John Wiley & Sons, Inc, New York, (1994).
- Pawlak, Z., "Rough sets: basic notion," *Int. J. of Computer and Information Science* 11, 344-56, (1982).
- Grzymala-Busse, J.W., "Learning from examples based on rough multisets," *Proc. of the 2nd Int. Symposium on Methodologies for Intelligent Systems*, Charlotte, North Carolina, October 14-17, 325-332, (1987).
- Chan, C.-C., "Distributed incremental data mining from very large databases: a rough multiset approach," *Proc. the 5th World Multi-Conference on Systemics, Cybernetics and Informatics, SCI 2001*, Orlando, Florida, July 22-25, (2001), 517-522.
- Chan, C.-C., "Learning rules from very large databases using rough multisets," *LNCS Transactions on Rough Sets*, (J. Peters, A. Skowron, J.W. Grzymala-Busse, B. Kostek, R.W. Swiniarski, M. Szczuka Eds.), Vol. 1, pp. 59 – 77, Springer –Verlag Berlin Heidelberg, 2004.
- Chan, C.-C. and S. Santhosh, "BLEM2: Learning Bayes' rules from examples using rough sets," *Proc. NAFIPS 2003, 22nd Int. Conf. of the North American Fuzzy Information Processing Society*, July 24 – 26, 2003, Chicago, Illinois, pp. 187-190.
- Greco, S., Matarazzo, B., Slowinski, R., "Rough approximation of a preference relation by dominance relations," *ICS Research Report* 16/96, Warsaw University of Technology, Warsaw (1996). *European Journal of Operational Research* 117(1), 63–83 (1999).
- Greco, S. Matarazzo, B., Slowinski, R., "A new rough set approach to evaluation of bankruptcy risk," in Zopounidis, C. (Ed.), *Operational Tools in the Management of Financial Risks*. Kluwer Academic Publishers, Dordrecht, pp. 121–136 (1998).
- Greco, S., Matarazzo, B., Slowinski, R., "The use of rough sets and fuzzy sets in MCDM," in Gal, T., Stewart, T., Hanne, T. (Eds.), *Advances in Multiple Criteria Decisions Making*. Kluwer Academic Publishers, Dordrecht, pp. 14.1–14.59 (Chapter 14) (1999).
- Greco, S., Matarazzo, B., Slowinski, R., "Rough sets theory for multicriteria decision analysis," *European Journal of Operational Research* 129(1), 1 – 47 (2001).
- Greco, S., Matarazzo, B., Slowinski, R., "Rough sets methodology for sorting problems in presence of multiple attributes and criteria," European Journal of Operational Research 138(2), 247 – 259 (2002).
- Greco, S., Matarazzo, B., Slowinski, R., Stefanowski, J., "An algorithm for induction of decision rules consistent with the dominance principle," in Ziarko, W. and Yao, Y. (Eds.), RSCTC 2000, LNAI 2005, Springer-Verlag Berlin Heidelberg, pp. 304 – 313 (2001).
- Greco, S., Matarazzo, B., Slowinski, R., Stefanowski, J., Zurawski, M., "Incremental versus non-incremental rule induction for multicriteria classification," Peters, J.F. et al. (Eds.): *Transactions on Rough Sets II, LNCS 3135*, Springer-Verlag Berlin Heidelberg, pp. 33 – 53 (2004).