

ROUGH SETS AND DECISION TABLES

Z. Pawlak

Institute of Computer Science
Polish Academy of Sciences
P.O. Box 22
00-901 Warsaw, PKiN, Poland

1. Introduction

We show in this article that the concept of the rough set (see Pawlak (1982 and 1985)) can be used as a basis for the decision tables theory (see Pollack, Hicks and Harrison (1971)). The ideas introduced in this paper have been applied to the implementation of cement kiln control algorithm (see Mrózek (1984)) and showed considerable practical advantages to other methods.

2. Decision tables

In this section we give a formal definition of a decision table which will be used throughout this paper.

A decision table is a system

$$S = (\text{Univ}, \text{Att}, \text{Val}, f)$$

where:

Univ - is a finite set of states, called the universe

Att = Con \cup Dec - is the set of attributes; Con - is the set of conditions attributes and Dec - is the set of decisions attributes.

Val = $\bigcup_{a \in \text{Att}} \text{Val}_a$, where Val_a is the set of values of an attribute $a \in \text{Att}$ (domain of a).

$f : \text{Univ} \times \text{Att} \rightarrow \text{Val}$ - is a total function, called the decision function, such that $f(x,a) \in \text{Val}_a$ for every $x \in \text{Univ}$ and $a \in \text{Att}$.

A decision rule in S is a function $f_x : \text{Att} \rightarrow \text{Val}$, such that $f_x(a) = f(x,a)$ for every $x \in \text{Univ}$ and $a \in \text{Att}$.

If f_x is a decision rule in S then f_x/Con and f_x/Dec are called conditions and decisions of the decision rule f_x respectively.

A decision rule f_x in S is deterministic (consistent) if for every $y \in \text{Univ}$, $y \neq x$ $f_x/\text{Con} = f_y/\text{Con}$ implies $f_x/\text{Dec} = f_y/\text{Dec}$; otherwise the decision rule f_x is nondeterministic (inconsistent).

A decision table S is deterministic (consistent) if all its decision rules are deterministic; otherwise the decision table S is nondeterministic (inconsistent).

A decision table $S^- = (X, \text{Att}, \text{Val}^-, f^-)$ is said to be a X-restriction of the decision table $S = (\text{Univ}, \text{Att}, \text{Val}, f)$, if $X \subseteq \text{Univ}$, $f^- = f/X \times \text{Att}$ and $\text{Val}^- = \{v \in \text{Val} : \bigvee_{x \in X} f_x(a) = v\}$

An example of a decision table is shown below:

Univ	a	b	c	d	e
1	1	0	2	2	0
2	0	1	1	1	2
3	2	0	0	1	1
4	1	1	0	2	2
5	1	0	2	0	1
6	2	2	0	1	1
7	2	1	1	1	2
8	0	1	1	0	1

Fig. 1

In the decision table $\text{Univ} = \{1, 2, \dots, 8\}$, $\text{Con} = \{a, b, c\}$, $\text{Dec} = \{d, e\}$ and domains of all attributes are equal $\text{Val} = \{0, 1, 2\}$.

3. Rough sets

Let $S = (\text{Univ}, \text{Att}, \text{Val}, f)$ be a decision table and let $a \subseteq \text{Att}$, $x, y \in \text{Univ}$.

With every subset of attributes $A \subseteq \text{Att}$ we associate the equivalence relation \tilde{A} defined thus

$$(x, y) \in \tilde{A} \quad \text{iff} \quad f_x(a) = f_y(a) \quad \text{for every } a \in A.$$

If $(x, y) \in \tilde{A}$ we say that x and y are indiscernible with respect to \tilde{A} in S (A - indiscernible) and \tilde{A} is called an indiscernibility relation in S . Equivalence classes of the indiscernibility relation \tilde{A} are called A-elementary sets in S and the family of an equivalence classes of \tilde{A} is denoted by A^* .

Let $A \subseteq \text{Att}$ and $X \subseteq \text{Univ}$ in a decision table $S = (\text{Univ}, \text{Att}, \text{Val}, f)$.

By A-lower (A-upper) approximation of X in S we mean the sets

$$\begin{aligned}\underline{A}X &= \{x \in \text{Univ} : [x]_A \subseteq X\} \\ \overline{A}X &= \{x \in \text{Univ} : [x]_A \cap X \neq \emptyset\}.\end{aligned}$$

Let $\text{Bn}_A(X) = \overline{A}X - \underline{A}X$ will be called A-boundary of X in S .

We shall use also the following definitions:

A - positive region of set X is the set $\underline{A}X$;

A - doubtful region of set X is the set $\text{Bn}_A(X)$;

A - negative region of set X is the set $\text{Neg}_A X = \text{Univ} - \overline{A}X$.

If $\overline{A}X = \underline{A}X$ we say that set X is A-definable in S ; otherwise set X is A-nondefinable in S .

Nondefinable sets will be called also rough sets in S .

The number

$$\alpha_A(X) = \frac{\text{card } \underline{A}X}{\text{card } \overline{A}X}$$

will be called the accuracy of the X with respect to A in S , and the number

$$\rho_A(X) = 1 - \alpha_A(X)$$

will be called the roughness of the set X with respect to A in S .

Let us notice that each subset of attributes $A \subseteq \text{Att}$ in a decision table $S = (\text{Univ}, \text{Att}, \text{Val}, f)$ defines uniquely the topological space $T_S = (\text{Univ}, \text{Def}_A(S))$, where $\text{Def}_A(S)$ is the family of all A -definable sets in S , and the lower and upper approximations are interior and closure in the topological space T_S , thus approximations have the following properties:

- 1) $\underline{A}X \subseteq X \subseteq \overline{A}X$
- 2) $\underline{A}\emptyset = \overline{A}\emptyset = \emptyset$; $\underline{A}\text{Univ} = \overline{A}\text{Univ} = \text{Univ}$
- 3) $\underline{A}(X \cup Y) \supseteq \underline{A}X \cup \underline{A}Y$
- 4) $\overline{A}(X \cup Y) = \overline{A}X \cup \overline{A}Y$
- 5) $\underline{A}(X \cap Y) = \underline{A}X \cap \underline{A}Y$
- 6) $\overline{A}(X \cap Y) \subseteq \overline{A}X \cap \overline{A}Y$
- 7) $\underline{A}(-X) = -\overline{A}(X)$
- 8) $\overline{A}(-X) = -\underline{A}(X)$

Moreover in this topological space we have the following two properties:

- 9) $\underline{AA}X = \overline{\overline{AA}}X$
 10) $\overline{\overline{AA}}X = \underline{AA}X$

From the topological view the rough sets can be classified as follows:

- a) Set X is roughly A-definable in S if $\underline{AX} \neq \emptyset$ and $\overline{AX} \neq \text{Univ}$,
 b) Set X is internally A-nondefinable in S if $\underline{AX} = \emptyset$ and $\overline{AX} \neq \text{Univ}$,
 c) Set X is externally A-nondefinable in S if $\overline{AX} = \text{Univ}$ and $\underline{AX} \neq \emptyset$,
 d) Set X is totally A-nondefinable in S if $\underline{AX} = \emptyset$ and $\overline{AX} = \text{Univ}$.

4. Dependency of attributes

Let $S = (\text{Univ}, \text{Att}, \text{Val}, f)$ be a decision table, $F = \{X_1, X_2, \dots, X_n\}$, where $X_i \subseteq \text{Univ}$, a family of subsets of Univ and $A \subseteq \text{Att}$.

By A-lower (A-upper) approximation of F in S we mean the families

$$\underline{AF} = \{\underline{AX}_1, \underline{AX}_2, \dots, \underline{AX}_n\}$$

$$\overline{AF} = \{\overline{AX}_1, \overline{AX}_2, \dots, \overline{AX}_n\}.$$

The A-positive region of a family F is the set

$$\text{Pos}_A(F) = \bigcup_{X_i \in F} \underline{AX}_i.$$

The A-doubtful region of a family F is the set

$$\text{Bn}_A(F) = \bigcup_{X_i \in F} \text{Bn}_A X_i.$$

The A-negative region of a family F is the set

$$\text{Neg}_A(F) = \text{Univ} - \bigcup_{X_i \in F} \overline{AX}_i.$$

The number

$$\gamma_A(F) = \frac{\text{card Pos}_A(F)}{\text{card } U}$$

will be called the quality at the approximation of F by A in S , and the number

$$\beta_A(F) = \frac{\text{card Pos}_A(F)}{\sum_{X_i \in F} \text{card } \overline{AX}_i}$$

will be called the accuracy at approximation of F by A in S .

Let $B, C \subseteq \text{Att}$ be two subsets of attributes in $S = (\text{Univ}, \text{Att}, \text{Val}, f)$, and k real number such that $0 \leq k \leq 1$.

We say that C depends in a degree k on B in S , in symbols $B \xrightarrow{k} C$, if $k = \gamma_B(C^*)$.

If $k = 1$ we say that C totally depends on B in S and we write also $B \rightarrow C$ instead of $B \xrightarrow{1} C$.

If $0 < k < 1$ we say that C roughly depends on B in S .

If $k = 0$ we say that C is totally independent on B in S .

The following properties are valid:

Property 1

A decision table $S = (\text{Univ}, \text{Att}, \text{Val}, f)$ is deterministic iff $\text{Con} \rightarrow \text{Dec}$ in S .

A decision table $S = (\text{Univ}, \text{Att}, \text{Val}, f)$ is called roughly deterministic if $\text{con} \xrightarrow{k} \text{Dec}$ and $0 < k < 1$.

A decision table $S = (\text{Univ}, \text{Att}, \text{Val}, f)$ is called totally nondeterministic if $\text{Con} \xrightarrow{0} \text{Dec}$.

Property 2

The following properties are true:

- 1) $\text{Con} \xrightarrow{1} \text{Dec}$ in $S/\text{Pos}_{\text{Con}}(\text{Dec}^*)$
- 2) $\text{Con} \xrightarrow{0} \text{Dec}$ in $S/\text{Bn}_{\text{Con}}(\text{Dec}^*)$.

Note. The above property says that every decision table can be decomposed into two parts (possibly empty) such that one is deterministic and the second totally nondeterministic.

It is easy to compute that in the decision table shown in Fig. 1,

$$\text{Pos}_{\text{Con}}(\text{Dec}^*) = \{3, 4, 6, 7, \}, \quad \text{Bn}_{\text{Con}}(\text{Dec}^*) = \{1, 2, 5, 8, \},$$

$\gamma_{\text{Con}}(\text{Dec}^*) = 0,5$, i.e. $\text{Con} \xrightarrow{0,5} \text{Dec}$, that is to mean that the decision table is roughly deterministic and can be decomposed into the following two decision tables

Univ	a	b	c	d	e
3	2	0	0	1	1
4	1	1	0	2	2
6	2	2	0	1	1
7	2	1	1	1	2

Fig. 2

Univ	a	b	c	d	e
1	1	0	2	1	0
2	0	1	1	1	2
5	1	0	2	0	1
8	0	1	1	0	1

Fig. 3

The decision table shown in Fig. 2 is totally deterministic and the table shown in Fig. 3 is totally nondeterministic.

Let us also notice that $\beta_{\text{Con}}(\text{Dec}^*) = 1/3$ which is the ratio of all deterministic decisions to all possible decisions in the table. The meaning of the number $\beta_{\text{Con}}(\text{Dec}^*)$ is obvious.

5. Reduction of attributes

Let $S = (\text{Univ}, \text{Att}, \text{Val}, f)$ be a decision table and let $A \subseteq \text{Att}$.

Set A is independent in S if for every $B \subset A$, $\tilde{B} \supset \tilde{A}$.

Set A is dependent in S if there exists $B \subset A$ such that $\tilde{B} = \tilde{A}$.

Set $B \subseteq A$ is a reduct of A in S if B is the maximal independent set in S .

Subset $B \subseteq A$ is a reduct of A with respect to $C \subseteq \text{Att}$ in S if B is an independent subset of A such that $\gamma_B(C^*) = \gamma_A(C^*)$ (or $\text{Pos}_B(C^*) = \text{Pos}_A(C^*)$).

Let us notice that if $A = C$ the reduct of A with respect to C coincide with the reduct of A .

Property 3

If $A \xrightarrow{k} B$ in S and C is a reduct of A , or reduct of A with respect to B in S , then $C \xrightarrow{k} B$,

In particular if C is a reduct of conditions attributes Con in a decision table S and $\text{Con} \xrightarrow{k} \text{Dec}$, then $C \xrightarrow{k} \text{Dec}$. This is to mean that we can simplify the decision table by reducing the set of conditions attributes.

For example in the decision table shown in Fig. 1 the only reduct of conditions attributes is $C = \{a, b\}$, thus we can simplify the decision table as shown below:

Univ	a	b	d	e
1	1	0	2	0
2	0	1	1	2
3	2	0	1	1
4	1	1	2	2
5	1	0	0	1
6	2	2	1	0
7	2	1	1	2
8	0	1	0	1

Fig. 4

and consequently the decision table can be decomposed as follows

Univ	a	b	d	e
3	2	0	1	1
4	1	1	2	2
6	2	2	1	1
7	2	1	1	2

Fig. 5

Univ	a	b	d	e
1	1	0	2	0
2	0	1	1	2
5	1	0	0	1
8	0	1	0	1

Fig. 6

It can be easily seen that the set of decision attributes in the table is independent.

We can also define the approximate reduct (or approximate reduct with respect to a subset C) in the following way:

Let $0 < \varepsilon < 1$ be a real number and let $B \subseteq A \subseteq \text{Att}$ in a decision table $S = (\text{Univ}, \text{Att}, \text{Val}, f)$.

Subset B of A is a ε -reduct of A in S if B is independent in S and $\gamma_B(A^*) = 1 - \varepsilon$.

Subset B of A is a ε -reduct of A in S with respect to $C \subseteq \text{Att}$ if B is independent in S and $\gamma_B(C^*) = \gamma_A(C^*) = \varepsilon$.

Directly from these definitions we have

Property 4

If B is a ε -reduct of A in S then $B \xrightarrow{1-\varepsilon} A$.

Property 5

If B is a ε -reduct of A in S with respect to $C \subseteq \text{Att}$, and $A \stackrel{k}{\rightarrow} C$, then $B \stackrel{k-\varepsilon}{\rightarrow} C$.

In particular if $\text{Con} \stackrel{k}{\rightarrow} \text{Dec}$ in S and $C \subseteq \text{Con}$ is a ε -reduct of Con in S , then $C \stackrel{k-\varepsilon}{\rightarrow} \text{Dec}$. That is to mean that we can reduce the set of conditions attributes, in such a way that the degree of dependency between decisions and conditions attributes is decreased by the constant ε .

6. Cost of decision table

With every decision table $S = (\text{Univ}, \text{Att}, \text{Val}, f)$ and subset of conditions attributes $B \subseteq \text{Con}$, we associate cost $C_{S,B}$, or in short C_B , when S is understood defined thus:

$$C_B = \sum_{a \in B} c_1(a) + \sum_{x \in \text{Bn}_B(\text{Dec}^*)} c_2(x),$$

where

$$c_1 : \text{Con} \rightarrow \mathbb{R}^+, \quad c_2 : \text{Univ} \rightarrow \mathbb{R}^+$$

and \mathbb{R}^+ is the set of nonnegative reals; $c_1(a)$ - is the cost of measuring the value of attribute a and $c_2(x)$ - is the cost of imprecise classification of x , due to smaller set of attributes (imprecise decision in state x).

There is of course trade-off between c_1 and c_2 and we can minimize the total cost C_B by proper choosing of conditions attributes.

The problem will be discussed in some details in a subsequent paper.

7. Decision nets

Many problems require not one decision table, but a set of different decision tables connected in a net in such a way that if some conditions are satisfied, the decision making process can be switched from one decision table to another one.

In order to do so assume that each decision table is labelled by a "name" of the decision table, and the set of decision attributes in each decision table contains one specific attribute - values of which are names of decision tables. Thus each condition in the decision table

specifies also next decision table to be used, to make a decision.

With each set of such decision tables we can associate a directed graph, nodes of which are labelled by decision tables, and are labelled by conditions, which determine transfers between corresponding tables.

For example consider three decision tables A,B,C as shown in Fig. 7.

A					B					C				
u	a	b	c	N	u	a	b	c	N	u	a	b	c	N
1	0	1	2	A	1	1	2	2	B	1	0	2	2	C
2	2	1	1	B	2	0	0	0	A	2	1	0	1	C
3	0	0	0	A	3	2	1	1	A	3	2	1	0	A
4	1	1	2	C	4	2	0	2	B					
					5	1	2	1	C					

Fig. 7

N - is the "next table" attribute.

The graph associated with this set of tables is shown in Fig. 8.

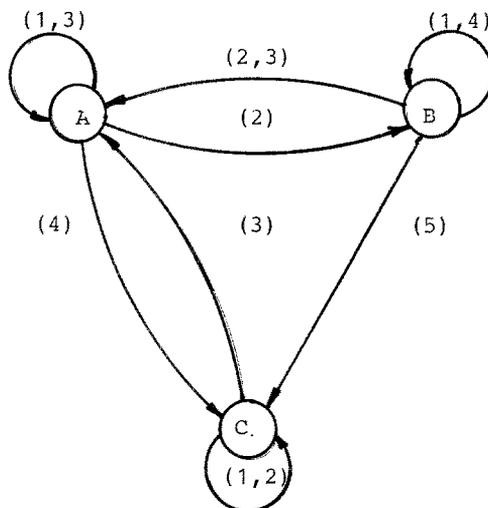


Fig. 8

For the sake of simplicity we labelled branches of the graph not by conditions but corresponding states.

Several theoretical problems arise in connection with decision nets, for example the problem of consistency (the decision not shown in Fig. 8 is inconsistent!) but we shall not discuss these problems

in these paper.

Acknowledgement. Thanks are due to dr. A. Skowron for critical remarks.

References

Mrózek, A., (1984). Information Systems and Control Algorithms, Bull. Polish Acad. Sci. (to appear)

Fawlak, Z., (1982). Rough Sets. International Journal of Information and Computer Sciences, 11(5), 341-356

Fawlak, Z., (1985). Decision tables and decision algorithms. Bull. Polish Acad. Sci. (to appear)

Pollack, S., Hicks, H., and Harrison, W., (1971). Decision Tables: Theory and Practice. Wiley and Sons, Inc. New York.