Zdzisław Pawlak

Rough sets

Algebraic and topological approach
Zdzisław Pawlak

ROUGH SETS:

Algebraic and topological approach

Warsaw, June 1982
Abstract. Содержание. Streszczenie

In this paper the theory of rough sets is developed in an algebraic and topological framework. Applications of this theory of rough sets are described for constructing decision support systems and learning systems.

Приближенные множества. Алгебраическое и топологическое отношение

В работе развивается теория приближенных множеств в алгебраическом и топологическом отношениях. Представлены применение конструирования учащихся и решаемых систем.

Zbiory przybliżone. Ujęcie algebraicznego i topologicznego

W pracy rozwiniona jest teoria zbiorów przybliżonych w ujęciu algebraicznym i topologicznym. Przedstawione są zastosowania tej teorii w zakresie konstruowania systemów uczących się i systemów decyzyjnych.
INTRODUCTION

This paper is an extended version of [15,16], where the concept of a rough set is studied.

It seems that the concept of the rough set can be of some value in some branches of artificial intelligence such as inductive reasoning, automatic classification, pattern recognition, learning algorithm etc.

Some results concerning rough sets are given in [4,5,6,7,15,16,19,20].

The idea of a rough set is in a way related to that of a fuzzy set [21], concept of alternative set theory [18] and non-standard analysis [17], however there are essential differences between those concepts, but we shall not discuss that problem here.

Thanks are due to prof. W. Marek and prof. E. Orłowska for their helpful comments and remarks.

1. APPROXIMATION SPACE; APPROXIMATIONS

1.1. Basic notions

Let \( U \) be a certain set called the universum, and let \( R \) be an equivalence relation on \( U \). The pair \( A = (U,R) \) will be called an approximation space. We shall call \( R \) an indiscernibility relation. If \( x,y \in U \) and \( (x,y) \in R \) we say that \( x \) and \( y \) are indistinguishable in \( A \).

Subsets of \( U \) will be denoted by \( X, Y, Z \) possibly with indices. The empty set will be denoted by \( 0 \), and the universum \( U \).
will be also denoted by $1$.

Equivalence classes of the relation $R$ will be called elementary sets (atoms) in $A$ or in short elementary sets. The set of all atoms in $A$ will be denoted by $U/R$.

We assume that the empty set is also elementary in every $A$.

Every union (finite) of elementary sets in $A$ will be called a composed set in $A$, or in short a composed set. The family of all composed sets in $A$ will be denoted as $\text{Com}(A)$. Certainly $\text{Com}(A)$ is a Boolean set algebra, i.e., the family of all composed sets is closed under intersection, union, and complement of sets.

Let $X$ be a certain subset of $U$. The least composed set in $A$ containing $X$ will be called the best upper approximation of $X$ in $A$, in symbols $\text{APr}_A(X)$; the greatest composed set in $A$ contained in $X$ will be called the best lower approximation of $X$ in $A$, in symbols $\text{AP}_A(X)$. If $A$ is known instead of $\text{APr}_A(X)$ (or $\text{AP}_A(X)$) we shall write $\text{APr}(X)$ (or $\text{AP}_A(X)$).

Since the intersection of arbitrary number of composed sets is again a composed set therefore we find the best upper approximation of a set exists. Similarly best lower approximation exists.

The set $\text{Bnd}_A(X) = \text{APr}_A(X) \cap \text{AP}_A(X)$ (in short $\text{Bnd}(X)$) will be called boundary of $X$ in $A$.

Sets $\text{Edg}_A(X) = X - \text{AP}_A(X)$ (in short $\text{Edg}(X)$) and $\text{Edg}_A(X) = \text{APr}_A(X) - X$, (in short $\text{Edg}(X)$) are referred to as an internal and an external edge of $X$ in $A$, respectively.

Of course $\text{Bnd}_A(X) = \text{Edg}_A(X) \cup \text{Edg}_A(X)$.

Fig. 1 depicts the notion of an upper and lower approximation in a two dimension approximation space consisting of a rectangle partitioned into elementary squares.

Let us define the membership functions $\mu_A, \nu_A$ (called strong and weak membership, respectively), as follows:

$$
\begin{align*}
&x \in A \quad \text{iff} \quad x \in \text{APr}_A(X) \\
&x \notin A \quad \text{iff} \quad x \notin \text{AP}_A(X)
\end{align*}
$$

If $x \in A$, we say "$x$ surely belongs to $X$ in $A$", and if $x \notin A$, is to mean "$x$ possibly belongs to $X$ in $A$". Thus we can interpret approximations as counterparts of necessity and possibility in modal logic.

1.2. Approximation space and topological space

It is easy to check that the approximation space $A = (U, \Omega)$ defines uniquely the topological space $T(A)$ (in short $T_A$), where $T_A = (U, \text{Com}(A))$, and $\text{Com}(A)$ is the family of all open sets in $T_A$, and $U/R$ is a base for $T_A$.

From the definition of approximations (lower and upper) follows that $\text{Com}(A)$ is both the set of all open and closed sets in $T_A$.

Thus $\text{APr}_A(X)$ and $\text{AP}_A(X)$ can be interpreted as an interior and closure of the set $X$ in the topological space $T_A$, respectively.

If $\text{APr}_A(X) = \text{AP}_A(X)$, for every $x \in U$, then $A = (U, R)$ will be called a discrete approximation space.

One can easily check that if $\tilde{A}$ is a discrete approximation space, then all atoms in $A$ are the same element sets.

Of course a discrete approximation space $A$ generates discrete topological space $T_A$.

Let us notice that, unless $A$ is discrete, $T_A$ is not a Hausdorff space.
1.3. Properties of approximations

From the topological interpretation of the approximation operations, follows that for every \( X, Y \subseteq U \) and every approximation space \( A = (U, R) \), the following properties are valid:

\[
\begin{align*}
(\text{A1}) \quad & A^R_A(X) \supset X \supset A^R_A(X) \\
(\text{A2}) \quad & A^R_A(1) = A^R_A(1) = 1 \\
(\text{A3}) \quad & A^R_A(0) = A^R_A(0) = 0 \\
(\text{A4}) \quad & A^R_A(A^R_A(x)) = A^R_A(A^R_A(x)) = A^R_A(x) \\
(\text{A5}) \quad & A^R_A(A^R_A(x)) = A^R_A(A^R_A(x)) = A^R_A(x) \\
(\text{A6}) \quad & A^R_A(X \cup Y) = A^R_A(X) \cup A^R_A(Y) \\
(\text{A7}) \quad & A^R_A(X \cap Y) = A^R_A(X) \cap A^R_A(Y) \\
(\text{A8}) \quad & A^R_A(X) = - A^R_A(-X) \\
(\text{A9}) \quad & A^R_A(X) = - A^R_A(-X),
\end{align*}
\]

where \(-X\) is an abbreviation for \( U \setminus X \).

Moreover, we have

\[
\begin{align*}
(\text{B1}) \quad & A^R_A(X \cap Y) \subseteq A^R_A(X) \cap A^R_A(Y) \\
(\text{B2}) \quad & A^R_A(X \cup Y) \supseteq A^R_A(X) \cup A^R_A(Y) \\
(\text{B3}) \quad & A^R_A(X \setminus Y) \supseteq A^R_A(X) \setminus A^R_A(Y) \\
(\text{B4}) \quad & A^R_A(X \setminus Y) \subseteq A^R_A(X) \setminus A^R_A(Y)
\end{align*}
\]

The following are counterparts of the law \( X \cup -X = 1 \), for approximations:

\[
\begin{align*}
(\text{C1}) \quad & A^R_A(X) \cup A^R_A(-X) = 1 \\
(\text{C2}) \quad & A^R_A(X) \cup A^R_A(-X) = 1 \\
(\text{C3}) \quad & A^R_A(X) \cup A^R_A(-X) = 1 \\
(\text{C4}) \quad & A^R_A(X) \cup A^R_A(-X) = - \text{Bnd}_A(X)
\end{align*}
\]

Moreover, we have

\[
\begin{align*}
(\text{D1}) \quad & A^R_A(X) \cap A^R_A(-X) = 0 \\
(\text{D2}) \quad & A^R_A(X) \cap A^R_A(-X) = \text{Bnd}_A(X) \\
(\text{D3}) \quad & A^R_A(X) \cap A^R_A(-X) = 0 \\
(\text{D4}) \quad & A^R_A(X) \cap A^R_A(-X) = 0
\end{align*}
\]

The law \( X \cap -X = 0 \) has the following analogues for approximations:

\[
\begin{align*}
(\text{E1}) \quad & A^R_A(X) \cap A^R_A(-X) = 0 \\
(\text{E2}) \quad & A^R_A(X) \cap A^R_A(-X) = \text{Bnd}_A(X) \\
(\text{E3}) \quad & A^R_A(X) \cap A^R_A(-X) = 0 \\
(\text{E4}) \quad & A^R_A(X) \cap A^R_A(-X) = 0
\end{align*}
\]

Let us notice the following equivalence:

\[
X = A^R_A(X) \iff X = A^R_A(X) \iff X \text{ is a composed set in } A.
\]

1.4. Refinement of an approximation space

If our discernment ability increases, that is to say we are able to distinguish smaller clusters of elements in the universe, the corresponding approximation space should have smaller atoms. We can formulate this problem precisely in the following way:

Let \( A = (U, R) \) and \( A' = (U, R') \) be two approximation spaces. If \( R \subseteq R' \), we say that the space \( A' \) is finer than the space \( A \) or that the space \( A \) is coarser than the space \( A' \).

If \( A' \) is finer than \( A \), then

\[
\begin{align*}
& A^R_A(X) \supset A^R_{A'}(X) \\
& A^R_A(X) \subseteq A^R_{A'}(X) \\
& \text{Bnd}_A(X) \subseteq \text{Bnd}_{A'}(X)
\end{align*}
\]

for every \( X \subseteq U \).

We say that an approximation space \( A \) is a refinement of another approximation space \( A' \), if there exist \( k \) \((k > 1)\) such that each atom of \( A' \) is union of \( k \) atoms of the space \( A \).
Of course if \( A' \) is refinement of \( A \) then \( A \) is finer than \( A' \).

Certainly discrete approximation spaces are the finest possible approximation spaces.

1.6. Sample of a set

Let \( A = (U, R) \) be an approximation space and let \( X \subseteq U \). We say that \( Y \ (Y \subseteq X) \) is a lower (upper) sample of \( X \) in \( A \) if \( \overline{A R}_A(Y) = A R_A(X) \) (\( \overline{A R}_A(Y) = \overline{A R}_A(X) \)).

A lower (upper) sample \( Y \) of \( X \) in \( A \) is proper if \( Y \) is the minimal sample of \( X \) in \( A \). Thus each lower (upper) sample of \( X \) in \( A \) has some element in common with every atom of \( \overline{A R}_A(X) \) (\( \overline{A R}_A(X) \)). The proper lower (upper) sample of \( X \) in \( A \) has exactly one element in common with every atom in \( \overline{A R}_A(X) \) (\( \overline{A R}_A(X) \)).

If \( A \) is a discrete approximation space, then for every \( X \subseteq U \), \( X \) is identical with its samples and \( X \) is proper sample of \( X \) in \( A \).

In fact this is a characteristic property of discrete spaces.

1.6. Accuracy of an approximation

In order to express the "quality" of an approximation we introduce some accuracy measure.

Let \( A = (U, R) \) be an approximation space, and let \( X \subseteq U \).

By \( \mu_A(X) \) (\( \mu_A(X) \)) we denote the number of atoms in \( \overline{A R}_A(X) \) (\( \overline{A R}_A(X) \)), and we call \( \mu_A(X) \) (\( \mu_A(X) \)) the internal (external) measure of \( X \) in \( A \).

In what follows we shall consider only sets with finite internal and external measure.

If \( \mu_A(X) = \mu_A(X) \) we say that \( X \) is measurable.

Thus the set \( X \) is measurable in \( A \) if and only if it is measurable in \( X \).

Let \( A = (U, R) \) be an approximation space and let \( X \neq \emptyset \).

By the accuracy of approximation of \( X \) in \( A \) we mean the number.

\[
\eta_A(X) = \frac{\mu_A(X)}{\mu_A(X)}
\]

Obviously \( 0 \leq \eta_A(X) \leq 1 \)

for any approximation space \( A = (U, R) \) and any \( X \subseteq U \).

For any set \( X \) in a discrete approximation space \( A = (U, R) \),

\[
\eta_A(X) = 1,
\]

and this is the greatest possible accuracy.

If \( A = (U, R) \) is a refinement of \( A = (U, R) \), then for every \( X \subseteq U \)

\[
\eta_A(X) \geq \eta_A(X).
\]

Certainly if \( Y' \) and \( Y'' \) are lower and upper proper samples of \( X \) in \( A \), then

\[
\eta_A(X) = \frac{\mu_A(Y')}{\mu_A(Y'')}, \quad \eta_A(X) = \frac{\mu_A(Y')}{\mu_A(Y'')}
\]

1.7. Examples

In this paragraph we depict introduced previously notions by means of simple examples.

1. Let \( R^+ \) be the set of non-negative real numbers, and let \( S \) be the indiscernibility relation on \( R^+ \) defined by the following partition:

\[
(0, 17), (1, 27), (2, 37), \ldots
\]
where \((1,1+1)\), \(i = 0,1,2,\ldots\) denotes a half-opened interval.

The corresponding approximation space will be denoted as \(A = (\mathbb{R}, S)\).

In that approximation space we have for example:

\[
\begin{align*}
\text{Ap}(1) &= \mathbb{R}, \\
\text{Ap}(1) &= (0,1], \\
\text{Ap}(1, \frac{1}{2}, \frac{3}{2}) &= \mathbb{R}, \\
\text{Ap}(1, \frac{1}{2}, \frac{1}{2}) &= (1,3].
\end{align*}
\]

If \(N = \{1,2,3,\ldots\}\) is the set of natural numbers, then:

\[
\begin{align*}
\text{Ap}(N) &= \mathbb{R}, \\
\text{Ap}(N) &= (0,1].
\end{align*}
\]

Thus the set of natural numbers is a disperse set in \(A\), and \(N\) is the proper upper sample of \(\mathbb{R}\).

**Example 2.** Let \(A = (\mathbb{R}, S)\) be an approximation space as in the previous example, and let us consider approximations of an open interval \((0,r)\), where \(n \leq r \leq n+1\) for a certain \(n \geq 0\). From the definition we have:

\[
\begin{align*}
\text{Ap}(0,r) &= \bigcup_{i=0}^{1} (i,i+1) = (0,n], \text{ for } n \geq 1 \text{ and } \emptyset \text{ for } n = 0, \\
\text{Ap}(0,r) &= \bigcup_{i=0}^{n} (i,i+1) = (0,n+1].
\end{align*}
\]

The internal and external measures of \((0,r)\) in \(A\) are:

\[
\begin{align*}
\mu(0,r) &= n, \\
\overline{\mu}(0,r) &= n+1,
\end{align*}
\]

and the accuracy of \((0,r)\) in \(A\) is:

\[
\eta(0,r) = \frac{n}{n+1}.
\]

Thus we can interpret the approximation space \(A = (\mathbb{R}, S)\) as a measurement system, where:

\[
\mathfrak{A}(1,1+1) = \mu_{A}(1,1+1) = 1, \quad i = 0,1,\ldots
\]

is the unit of measurement in \(A\), and \(\eta(0,r)\) is the degree of \((0,r)\) in \(A\).

**Example 3.** Let \(V\) be a finite set called a vocabulary and \(V^{*}\) be the set of all finite sequences over \(V\). Any subset \(W \subseteq V^{*}\) will be called a language over \(V\).

Let \(R \subseteq V^{*} \times V^{*}\) be an indiscernibility relation, and \(A = (V, R)\) be an approximation space defined by \(V\) and \(R\).

A language \(L \subseteq V^{*}\) is recognizable in \(A\) if \(\text{Ap}_{A} = \text{Ap}_{A}(L)\).

The family of all recognizable languages in \(A\), denoted \(\text{Rec}(A)\), is the topology induced by \(A = (V, R)\) and \(\text{Rec}(A)\) is a topology in \(V^{*}/R\).

**Example 4.** Let \(S = \{X,A,V,p\}\) be an information system, where:

- \(X\) is the set of objects
- \(A\) is the set of attributes
- \(V = \bigcup_{i} V_{a}\) - is the set of values of attribute \(a \in A\)
- \(p : X \times A \rightarrow V\) is an information function, \(p_{X} : A \rightarrow V\), \(x \in X\)
- \(p_{X}(a) = p(x,a)\)

for every \(x \in X\) and \(a \in A\).

We define a binary relation \(S\) over \(X\) in the following way:

\[
x S y \quad \text{iff} \quad p(x) = p(y).
\]

Obviously \(S\) is an equivalence relation and \(A = (X, S)\) will be called the approximation space induced by the information system \(S\).

Any subset \(Y \subseteq X\) is called describable in \(S\) iff \(\text{Ap}_{A}(Y) = \text{Ap}_{A}(V)\).

The set of all describable sets in \(S\), denoted as
Des(S), is a zero dimensional topology induced by $S$ on $X$, and the base of the topology is $X/\mathcal{D}$. 

2. ROUGH EQUALITY OF SETS

2.1. Basic definitions

Let $A = (U, R)$ be an approximation space and let $X, Y \subseteq U$. We say that:

(a) The sets $X, Y$ are **roughly bottom - equal** in $A$, in symbols $X \mathcal{R}_A Y$, iff
$$\mathcal{AP}_A(X) = \mathcal{AP}_A(Y).$$

(b) The sets $X, Y$ are **roughly top - equal** in $A$, in symbols $X \mathcal{T}_A Y$, iff
$$\mathcal{AP}_A(X) = \mathcal{AP}_A(Y).$$

(c) The sets $X, Y$ are **roughly equal** in $A$, in symbols $X \mathcal{E}_A Y$, iff
$$X \mathcal{R}_A Y \text{ and } X \mathcal{T}_A Y.$$

It is easy to check that $\mathcal{R}_A, \mathcal{T}_A, \mathcal{E}_A$ are equivalence relations on $P(U)$.

In what follows we shall omit the subscript $A$ if the approximation space $A$ is understood - and write $\simeq, \approx, \equiv$, instead $\mathcal{R}, \mathcal{T}, \mathcal{E}$. 

2.2. Properties of rough equality

For any approximation space $A = (U, R)$ and any $X, Y \subseteq U$ the following properties are true:

- (A1) If $X \simeq Y$, then $X \cap Y \simeq X \equiv Y$
- (A2) If $X \equiv Y$, then $X \cup Y \equiv X \equiv Y$
- (A3) If $X \equiv X'$ and $Y \simeq Y'$, then $X \cup Y \simeq X' \cap Y'$
- (A4) If $X \simeq X'$ and $Y \equiv Y'$, then $X \cap Y \equiv X' \cap Y'$
- (A5) If $X \equiv Y$, then $X = Y$ and $X \equiv 0$
- (A6) $X \simeq Y$ iff $X = Y$
- (A7) If $X \equiv Y$, then $-(X) \equiv Y$
- (A8) If $X \equiv Y$, then $-(X) \equiv Y$
- (A9) If $X \equiv Y$, then $-(X) \equiv Y$
- (A10) If $X \equiv Y$, then $X \cup Y \simeq 1$
- (A11) If $X \equiv Y$, then $X \cup -Y \simeq 1$
- (A12) If $X \equiv Y$, then $X \cap -Y \simeq 0$
- (A13) If $X \equiv Y$, then $X \cap -Y \simeq 0$

Set $X$ will be called **roughly dense** ($r$ - dense) in $X \simeq 1$.

Set $X$ will be called **roughly co-dense** ($r$ - co-dense) in $X \mathcal{E} 0$.

Set $X$ will be called **roughly dispersed** ($r$ - dispersed) in $A$ if $X$ is both $r$-dense and $r$-co-dense in $A$.

One can easily show the following properties:

- (B1) If $X \subseteq Y$ and $Y \equiv 0$, then $X \equiv 0$
- (B2) If $X \subseteq Y$ and $X \equiv 1$, then $Y \equiv 1$
- (B3) If $X \equiv 1$, then $-X \equiv 0$
- (B4) If $X \equiv 0$, then $-X \equiv 1$
- (B5) If $X$ is a $r$-dispersed set, then so is $-X$
- (B6) $X \cap Y \equiv 0$ iff $X \equiv 0$ or $Y \equiv 0$
- (B7) $X \cup Y \equiv 1$ iff $X \equiv 1$ or $Y \equiv 1$
- (B8) If $X, Y$ are both $r$-dense, then $X \mathcal{E} Y$
- (B9) If $X, Y$ are both $r$-co-dense then $X \mathcal{E} Y$
3. ROUGH INCLUSION OF SETS

3.1. Basic definitions

Let $A = (U, R)$ be an approximation space and let $X, Y \subseteq U$.

We introduce the following definitions:

(a) We say that $X$ is roughly bottom included in $Y$, in $A$, in symbols $X \preceq_A Y$, if $\text{Apr}_A(X) \subseteq \text{Apr}_A(Y)$.

(b) We say that $X$ is roughly top included in $Y$, in $A$, in symbols $X \succeq_A Y$, if $\text{Apr}_A(X) \supseteq \text{Apr}_A(Y)$.

(c) We say that $X$ is roughly included (r-included) in $Y$, in $A$, in symbols $X \preceq_{r} Y$ if $X \preceq_A Y$ and $X \succeq_A Y$.

If $A$ is understood instead of $X \preceq_A Y$, $X \succeq_A Y$, we shall write $X \preceq Y$, $X \succeq A$, $X \preceq Y$, respectively.

If $X \preceq_A Y$, $X$ is called rough upper-subset of $Y$ in $A$;

If $X \succeq_A Y$, $X$ is called rough lower-subset of $Y$ in $A$.

One can easily check that all rough inclusions $\preceq$, $\succeq$, $\preceq$, are ordering relations.

3.2. Properties of rough inclusions

It is easy to prove by simple computations that the following properties are true:

- If $X \preceq Y$, then $X \preceq Y$, $X \preceq Y$.
- If $X \preceq Y$ and $Y \preceq X$, then $X \preceq Y$.
- If $X \preceq Y$ and $Y \preceq X$, then $X \preceq Y$.

3.3. Rough powersets

The family of all r-subsets (lower, upper) of $X$ in $A$ will be denoted by $P_A(X) = \{P_A(X), P_A(X)\}$ and will be called rough (lower, upper) powerset of $X$ in $A$.

Thus:

- $P_A(X) = \{Y : Y \preceq_A X\}$
- $P_A(X) = \{Y : Y \succeq_A X\}$
- $P_A(X) = \{Y : Y \preceq_{r} X\}$

It is easy to see that:

- $P(X) \subseteq P_A(X)$
- $P(X) \subseteq P_A(X)$
- $P(X) \subseteq P_A(X)$

for every $A$. 
4. ROUGH SETS

4.1. Basic notions

Let \( A = (U, R) \) be an approximation space, and let \( \overline{\sim}_A, \overline{\sim}_A^* \), \( \overline{\sim}_A^* \), be the induced equivalence relations on \( P(U) \).

Every approximation space \( A = (U, R) \) defines three following approximation spaces:

\[
\overline{\sim}_A^* = (P(U), \overline{\sim}_A^*), \\
\overline{\sim}_A^* = (P(U), \overline{\sim}_A^*), \\
A^* = (P(U), \overline{\sim}_A^*),
\]

in which objects are subsets of \( U \) and the relations \( \overline{\sim}_A, \overline{\sim}_A^* \) are the indiscernibility relations in the corresponding spaces \( \overline{\sim}_A, \overline{\sim}_A^* \), \( A^* \).

The approximation space \( A^*(\overline{\sim}_A, \overline{\sim}_A^*) \) will be called the extension (lower, upper) of \( A \).

Equivalence classes of the relation \( \overline{\sim}_A^* \) will be called rough sets (lower, upper).

Thus, rough set (lower, upper) is a family of subsets of \( U \), which are equivalent with respect to the indiscernibility relation \( \overline{\sim}_A \).

Every approximation space \( A^* \) induces a topology \( \text{Con}(A^*) \), \( \text{Con}(\overline{\sim}_A^*) \), \( \text{Con}(\overline{\sim}_A^*) \) respectively, and consequently the topological spaces

\[
\overline{\sim}_A = (P(U), \text{Con}(\overline{\sim}_A)), \\
\overline{\sim}_A^* = (P(U), \text{Con}(\overline{\sim}_A^*)), \\
\overline{\sim}_A^* = (P(U), \text{Con}(\overline{\sim}_A^*)),
\]

and \( P(U)/\overline{\sim}_A, P(U)/\overline{\sim}_A^*, P(U)/\overline{\sim}_A^* \) are the bases for the corresponding topological spaces.

In other words \( P(U)/\overline{\sim}_A, P(U)/\overline{\sim}_A^*, P(U)/\overline{\sim}_A^* \) are families of equivalence classes of the relations \( \overline{\sim}_A, \overline{\sim}_A^* \), \( \overline{\sim}_A^* \) respectively.

i.e. families of elementary classes in the corresponding spaces \( A^*, \overline{\sim}_A^*, \overline{\sim}_A^* \).

If \( X, Y \subseteq U \), and \( X = Y \) \( \overline{\sim}_A \), \( Y \) \( \overline{\sim}_A \) \( Y \) \( \overline{\sim}_A \), we say that \( Y \) are close (bottom, top) in \( A \), otherwise sets \( X, Y \) are to in \( A \).

That is to say, that if \( X, Y \) are close (bottom, top) they belong to the same equivalence class in the approximation space \( \overline{\sim}_A \) \( \overline{\sim}_A^* \). Thus sets which are close (bottom, top) are in a sense similar and we are unable to distinguish then our approximation space \( A^*(\overline{\sim}_A, \overline{\sim}_A^*) \).

Thus we are able to introduce the upper and lower approximations of a family of subsets of \( U \) in \( \overline{\sim}_A \), \( \overline{\sim}_A^* \), \( \overline{\sim}_A^* \), i.e., if is a certain family of subsets of \( U \) we can introduce the following approximations of \( F \) in \( A^*, \overline{\sim}_A^*, \overline{\sim}_A^* \):

\[
\text{Ap}_A(F), \text{Ap}_{\overline{\sim}_A^*}(F), \text{Ap}_{\overline{\sim}_A^*}(F),
\]

\[
\text{Ap}_{\overline{\sim}_A^*}(F), \text{Ap}_{\overline{\sim}_A^*}(F), \text{Ap}_{\overline{\sim}_A^*}(F),
\]

but we shall not consider this problem here.

There is also another important problem considered in artificial intelligence. Given a family \( F \) of subsets of the universe \( U \), the task is to classify members of \( F \), such that sets in the same equivalence class are according to a certain criterion similar.

In our approach we can formulate the problem as follows:

Let \( A = (U, R) \) be an approximation space and let \( F \subseteq P(U) \), be a certain family (non-empty) of subsets of the universe \( U \).

By \( \overline{\sim}_A \cap F^2 \) (\( \overline{\sim}_A \cap F^2 \), \( \overline{\sim}_A \cap F^2 \)) we mean the restriction of the relation \( \overline{\sim}_A \cap F^2 \), to the family \( F \). Then \( F/\overline{\sim}_A \cap F^2 \) (\( F/\overline{\sim}_A \cap F^2 \), \( F/\overline{\sim}_A \cap F^2 \)) is to mean the family of equivalence classes of the relation \( \overline{\sim}_A \cap F^2 \) restricted to the family \( F \).
42. Extensions of higher order

In a similar way as before we can extend each of the approximation spaces $A^k$, $A^+$, $A^*$ obtaining approximation spaces of higher order.

The $k$-extension (lower, upper) $k \geq 0$, will be defined inductively:

(i) An approximation space $A = (U, R)$ is of order $0$,

(ii) If $A^k$ is the $k$-extension of $A$, then $k+1$-extension (lower, upper) of $A$ are defined as follows:

$$A^{k+1} = (p^{k+1}(U), \leq_k)$$

$$A^{k+1} = (p^{k+1}(U), \geq_k)$$

where $p^k(U)$ is defined as

(a) $p^0(U) = U$,

(b) $p^{k+1}(U) = p(p^k(U))$.

Thus $A^0 = A$, $A^1 = A^+$, etc.

In this way every approximation space $A = (U, R)$ defines uniquely, infinite sequence of approximation spaces of higher or-

ders, allowing to cluster sets, families of sets etc., shall not discuss that problem here.

4.3. Ordering of rough sets

Every approximation space $A = (U, R)$ induces three relations $\leq$, $\leq$, $\leq$ (in short $\leq$, $\leq$, $\leq$) on the families $P(U)/A$, $P(U)/A$, $P(U)/A$ respectively.

Let $A = (U, R)$ be an approximation space, and let $X, Y \subseteq A$ then

(a) $[X]_A \leq [Y]_A \iff X \subseteq Y$

(b) $[X]_A \leq [Y]_A \iff X \supseteq Y$

(c) $[X]_A \leq [Y]_A \iff X \cap Y$

One can easily check that $\leq$, $\leq$, $\leq$ are ordering relations on $P(U)/A$, $P(U)/A$, $P(U)/A$ respectively. Thus

$$p^0(U) = (P(U)/A, \leq)$$

$$p^1(U) = (P(U)/A, \leq)$$

$$p^2(U) = (P(U)/A, \leq)$$

are partially ordered families of sets.

The smallest element in $p^0(U)$ is $[\emptyset]_A$, i.e. the family of all co-dense sets in the topological space $T_A$. The greatest element in $p^0(U)$ is $[U]_A = 1$, i.e. the class consisting only of one set, the universe $U$.

The smallest element in $p^2(U)$ is $[\emptyset]_A = 0$, i.e. the class consisting of only one set, the empty set, the greatest element in $p^2(U)$ is $[1]_A = 1$, i.e. the family of all dense sets in the topological space $T_A$.

The smallest element in $p^3(U)$ is $[\emptyset]_A = 0$, and the greatest element in $p^3(U)$ is $[1]_A = 1$. 

...
Finally let us remark that if an approximation space $\mathcal{A}(U, R)$ has $k$ atoms (elementary sets), then there are $2^k$ equivalence classes in a family $P(U)/\mathcal{A}$, (and $P(U)/\mathcal{A}$), but there are $\sum_{k=0}^{n} \binom{n}{k} 2^{k-1} = 2^k$ equivalence classes in a family $P(U)/\mathcal{A}$.

4.4. Algebraic properties

Let $A = (U, R)$ be an approximation space and let $X \subseteq U$ be a certain subset of the universe $U$. Rough set in $\mathcal{A}(\mathcal{A}, \mathcal{A}^*)$ containing set $X$ will be denoted by $[X]_\mathcal{A} \mathcal{A} (\mathcal{A}, \mathcal{A}^*, \mathcal{A})$ or in short $[X]_\mathcal{A}$, $[X]_\mathcal{A}$, $[X]_\mathcal{A}$, when $\mathcal{A}(\mathcal{A}, \mathcal{A}^*)$ is known.

We introduce three operations on rough sets (lower, upper) denoted as $\lor$, $\land$, $\neg$, and called union, intersection and complement respectively.

Let $A = (U, R)$ be an approximation space and let $X, Y, Z \subseteq U$.

For the approximation space $\mathcal{A} = (P(U), \mathcal{A})$, we define

\[(A1) \; [X]_\mathcal{A} \lor [Y]_\mathcal{A} = \{Z \subseteq U : A(R)(Z) = A(R)(X \cup Y)\}.
\]

\[(A2) \; [X]_\mathcal{A} \land [Y]_\mathcal{A} = \{Z \subseteq U : A(R)(Z) = A(R)(X \cap Y)\}.
\]

\[(A3) \; \neg [X]_\mathcal{A} = \{Z \subseteq U : A(R)(Z) = \neg A(R)(-X)\}.
\]

Directly from the definition of a rough set one can easily obtain that

\[(i) \; [X]_\mathcal{A} \lor [Y]_\mathcal{A} \lor [X \cap Y]_\mathcal{A} \lor [X \cup Y]_\mathcal{A}
\]

\[(ii) \; [X]_\mathcal{A} \land [Y]_\mathcal{A} = [X \cap Y]_\mathcal{A}
\]

\[(iii) \; \neg [X]_\mathcal{A} = \neg [X]_\mathcal{A}
\]

For the approximation space $\mathcal{A}^* = (P(U), \mathcal{A})$ we define

\[(A1) \; [X]_\mathcal{A} \lor [Y]_\mathcal{A} = \{Z \subseteq U : A(R)(Z) = A(R)(X \cup Y)\}.
\]

\[(A2) \; [X]_\mathcal{A} \land [Y]_\mathcal{A} = \{Z \subseteq U : A(R)(Z) = A(R)(X \cap Y)\}.
\]

\[(A3) \; \neg [X]_\mathcal{A} = \{Z \subseteq U : A(R)(Z) = \neg A(R)(-X)\}.
\]

As before, we obtain for this approximation space

\[(i) \; [X]_\mathcal{A} \lor [Y]_\mathcal{A} = [X \cup Y]_\mathcal{A}
\]

\[(ii) \; [X]_\mathcal{A} \land [Y]_\mathcal{A} \lor [X \cap Y]_\mathcal{A} \lor [X \cup Y]_\mathcal{A}
\]

\[(iii) \; \neg [X]_\mathcal{A} = \neg [X]_\mathcal{A}
\]

A family $F$ of subsets of the set $U$ is called an ideal if the following conditions are satisfied:

\[(a) \; \text{if } X \in F \text{ and any } Y \subseteq X, \text{ then } Y \in F.
\]

\[(b) \; \text{if } X \in F \text{ and } Y \in F, \text{ then } X \cup Y \in F.
\]

A family $G$ of subsets of the set $U$ is called a filter if the following conditions are satisfied:

\[(c) \; \text{if } X \in G \text{ and } X \subseteq Y, \text{ then } Y \in G.
\]

\[(d) \; \text{if } X \in G \text{ and } Y \in G, \text{ then } X \cap Y \in G.
\]

Let $A = (U, R)$ be an approximation space, and let $\mathcal{A}^*$ be its lower (upper) extension, and let $P(U)/\mathcal{A} = (P(U)/\mathcal{A})$ be the family of lower, (upper) rough sets generated by $A\mathcal{A}(\mathcal{A})$.

A family $F$ of rough sets in $\mathcal{A}^*$ will be called rough ideal in $\mathcal{A}^*$ if the conditions are satisfied:

\[(1) \; \text{if } [X]_\mathcal{A} \lor [Y]_\mathcal{A} \in F \text{ and } Y \subseteq X, \text{ then } [Y]_\mathcal{A} \in F.
\]

\[(2) \; \text{if } [X]_\mathcal{A} \land [Y]_\mathcal{A} \in F \text{ and } [Y]_\mathcal{A} \in F, \text{ then } [X]_\mathcal{A} \lor [Y]_\mathcal{A} \lor [X \cap Y]_\mathcal{A} \lor [X \cup Y]_\mathcal{A} \in F.
\]

A family $G$ of rough sets in $\mathcal{A}^*$ will be called rough filter in $G$ if the following conditions are satisfied:
Thus each information system $S = \langle X, A, V, \rho \rangle$ induces an approximation space $A = (X, \succeq)$. Suppose we are given the set $Y \subseteq X$ of patients suffering from a certain disease (the set $Y$ can be given by an expert) and we are interested in finding the characteristic symptoms of that disease.

It follows from the previous considerations that we can give those characteristic symptoms only in that case when $Y$ is a composed set in $S$, otherwise we can give only symptoms of lower or upper approximation of $Y$ in the approximation space $A = (X, \succeq)$. In other words if $Y$ is not a composed set in $S$ we are not able to give the characteristic symptoms of the set $Y$, but we can give only the symptoms of patients who surely have the disease $Y$ (symptoms of patients belonging to the lower approximation of $Y$) or the symptoms of patients who possibly have the disease $Y$ (symptoms of patients belonging to the upper approximation of $Y$). Let us remark that we identify here the disease with the set of patients having this disease according to the opinion of a certain expert. Another expert can point out of course different sets of patients having the considered disease.

Of course, in order to get the characteristic symptoms of the disease $Y$, we can use not the whole set $Y$ but its lower and upper sample.

5. EXAMPLES

5.1. Characteristic symptoms

Let us consider an information system $S = \langle X, A, V, \rho \rangle$ as in example 4, section 1.7, and let us assume that $X$ is a set of patients in a certain hospital, $A$ is the set of attributes like temperature, blood pressure etc., $V = U V_A$, where $A$ is the set of values of attributes, and the function $\rho: X \rightarrow V$ describes symptoms of patient $X$.

Obviously patients belonging to the same equivalence class of the relation $S$ have the same symptoms.
In general case of course the answer is not, the student can describe the set Y pointed out by an expert in terms of symptoms only if Y is a composed set in S. Otherwise, the student can give only approximate description of the disease Y, i.e. symptoms of lower and upper approximations of Y in S.

We understand that if Y contains patients having a certain disease, then the set - Y does not contain patients having this disease. That is to say that the expert classifies all patients in two classes Y - Y such that Y contains all patients having a certain disease, and - Y - not having this disease.

Sometimes an expert may be unable to classify patients in two classes, as before, since in some cases he may be unable to classify a patient to the class Y or -Y. That is to say sometimes an expert does not know how to classify some objects. In fact in this case he may classify patients in three classes Y+, Y-, Y0, such that Y+ contains patients who are ill, Y- those who are not ill and in Y0 there are patients about which an expert is unable to decide whether they are ill or not.

The question arises how this not complete classification influences the process of learning?

From the previous consideration follows that if the set Y0 A 0, the learning is not affected by the "incomplete" knowledge of an expert, and a student can get exactly the same results as in the case when the expert classification were complete. Otherwise, i.e. if Y0 is not bottom equal to zero, a student is unable to learn (even approximate) classification properly.

That is to say if the "incompleteness" of knowledge of an expert is small enough it does not affect learning, otherwise the learning is affected.

5.1. The case of many experts

Let us consider an information system as in previous sections, and let us suppose, that we employed k experts, to pick up all patients having a certain disease. Thus we obtain a family F = {X1, X2, ..., Xn} of subsets of X such that the set Xk contains all the patients having the considered disease according the opinion of expert k.

The question arises what is the difference between opinions of experts, or, in other words how to classify opinions of experts in such a way that "close" opinions are in the same class and "remote" opinions are in different classes.

To do this we can use the three natural classifications C₁(F), C₂(F), C₃(F), which have in this case the following meaning:

Each equivalence class of the classification C₁(F), contains all subsets of the family F having the same lower approximations, i.e. sets which are "close" with respect to symptoms which solely occur among patients in all sets in each equivalence class.

Each equivalence class of the classification C₂(F) contains all subsets of the family F having the same upper approximations, i.e. sets which are "close" with respect to symptoms which possibly occur among patients in all sets in each equivalence class.

Each equivalence class in the third classification C₃(F) contains which have the same sure and possible symptoms.

Thus we can cluster opinions (or experts) into natural similarity classes.
Again let us consider information system as before and a family \( F = \{ X_1, X_2, \ldots, X_n \} \) of subsets of \( X \). Suppose that \( F \) has been given by an expert and each \( X_i \) represent, according to his knowledge different disease, i.e. all patients having the disease \( i \), belong to the subset \( X_i \).

The question is whether we are able to distinguish all subset of the family \( F \) by symptoms? or in other words whether we are able to classify all subset of \( F \) in similarity classes, so that in each similarity class we have all subsets of \( F \) which are undistinguishable in the approximation space \( A = (X, S) \).

To solve this problem we can use the three "natural" classifications \( C^A(F), C^X(F), C^S(F) \) as in the previous section.

The meaning of the classification \( C^A(F) \) is that in each equivalence class of \( C^A(F) \) we have all subsets of \( F \) (or diseases) which we are unable to distinguish by means of symptoms available in our information system, and which surely occur in each disease being in the same equivalence class.

The meaning of the classification \( C^X(F) \) and \( C^S(F) \) is obvious.

Thus we can cluster diseases (subsets of the family \( F \)) into classes such that in each equivalence class there are diseases, which we are not able to distinguish by means of symptoms available in the information system \( S \).

### 5.5. Diagnosis

Suppose again that we are given an information system as previously, and the family \( F = \{ X_1, X_2, \ldots, X_k \} \) of subsets of \( X \), determined by an expert such that each \( X_i \) contains all patients having a certain disease.

The problem is the following: given a symptom \( p \)

a) what diseases surely have symptom \( p \)
b) what diseases possibly have symptom \( p \)?

Let \( E_p \) denote an equivalence class of the relation \( S \), defined by the symptom \( p \).

Of course, all diseases \( X_j \in F \) such that \( AP_{AF}(X_j) \supseteq E_p \)

surely have the symptom \( p \), and all diseases \( Y_i \in F \) such that \( AP_{AF}(Y_i) \supseteq E_p \) possibly have the symptom \( p \).

If we classify diseases \( F \) according to the classifications \( C^A(F) \) and \( C^S(F) \), instead of checking whether lower (upper) approximation of each subset \( X_i \) of the family \( F \) contains \( E_p \), we can simply check whether the corresponding classes contain \( E_p \) or not - what considerably simplifies the algorithm.

---

Fig. 1
REFERENCES

[7] Orłowska, E., Languages of approximate information, ICS PAS Reports (1982), No 479

[19] Zakowski, W., On a concept of rough sets, Doktor Mathematicae (to appear)
| CONTENTS |
|---------------------------------|---|
| Introduction                     | 5 |
| 1. Approximation space; approximations | 5 |
| 1.1. Basic notions               | 5 |
| 1.2. Approximation space and topological space | 7 |
| 1.3. Properties of approximations| 8 |
| 1.4. Refinement of an approximation space | 9 |
| 1.5. Sample of a set             | 10|
| 1.6. Accuracy of an approximation| 10|
| 1.7. Examples                    | 11|
| 2. Rough equality of sets        | 14|
| 2.1. Basic definitions           | 14|
| 2.2. Properties of rough equality| 14|
| 3. Rough inclusion of sets       | 16|
| 3.1. Basic definitions           | 16|
| 3.2. Properties of rough inclusions | 16|
| 3.3. Rough powersets             | 17|
| 4. Rough sets                    | 18|
| 4.1. Basic notions               | 18|
| 4.2. Extensions of higher order  | 20|
| 4.3. Ordering of rough sets      | 21|
| 4.4. Algebraic properties        | 22|
| 5. Examples                      | 24|
| 5.1. Characteristic symptoms     | 24|
| 5.2. Learning                    | 25|
| 5.3. The case of many experts    | 27|
| 5.4. Classification              | 28|
| 5.5. Diagnostic                  | 28|
| Bibliography                     | 31|