PRACE IPI PAN • ICS PAS REPORTS

Zdzisław Pawlak

Rough sets

Power set hierarchy

470

March 1982
WARSZAWA

INSTYTUT PODSTAW INFORMATYKI POLSKIEJ AKADEMII NAUK
INSTITUTE OF COMPUTER SCIENCE POLISH ACADEMY OF SCIENCES
00-901 WARSAW, P.O. Box 22, POLAND

Zdzisław Pawlak

ROUGH SETS

Power set hierarchy

470

Warsaw, April 1982

Pracę zgłosił Andrzej Blikle

Mailing address: Zdzisław Pawlak
Institute of Computer Science
Polish Academy of Sciences
00-901 Warsaw, PKiN
P.O. Box 22

ISSN 0138-0648

P r i n t e d   a s   a   m a n u s c r i p t
N a   p r a w a c h   r ę k o p i s u.

Nakład 600 egz. Ark. wyd. 0,90; ark. druk. 1,75. Papier offset. kl. III, 70 g, 70 x 100. Oddano do druku w kwietniu 1982 r. W.D.N. Zam. nr

Abstract . Streszczenie . Содержание

In this paper we consider rough equality of sets and rough inclusion of sets. A family of roughly equal sets is called rough set. Some properties of rough sets are given. The idea of a rough set can be used as a basis for a "topological cluster analysis" which can be considered as a counterpart of a "metric cluster analysis".

Zbiory przybliżone - hierarchia podzbiorów

W pracy rozważana jest przybliżona równość oraz zawieranie zbiorów. Rodzina zbiorów w przybliżeniu równych jest nazywana zbiorem przybliżonym. Badano niektóre własności zbiorów przybliżonych. Idea zbioru przybliżonego może być użyta jako podstawa dla "topologicznej analizy skupień", którą można uważać za odpowiednik "metrycznej analizy skupień".

Нечеткие множества - иерархия подмножеств

В работе рассматривается нечеткое равенство а также отношение содержания множеств. Семейство множеств приблизительно равных называется нечетким множеством. Исследовано некоторые свойства нечетких множеств. Идея нечеткого множества может быть использована в качестве основы для "топологического анализа сосредоточенности", которую можно считать как эквивалент "метрического анализа сосредоточенности".

INTRODUCTION

The aim of this paper is to give some properties of rough sets, intorduced in [7] and investigated in [2], [3], [4], [5], [8], [9], [10], [11], [14].

The rough set concept can be of some importance, primarily in some branches of artificial intelligence, such as inductive reasoning, automatic classification, pattern recognition, learning algorithms etc.

The idea of a rough set could be placed on a more general setting, leading to a fruitful further area of research and applications in fields like theory of classification, cluster analysis, measurement theory, taxonomy and others.

The key to the presented approach is provided by the exact mathematical formulation of the concept of approximative (rough) equality of sets in a given approximation space; an approximation space is understood as pair (U,R), where U is a certain set called universum, and $R \subset U \times U$ is an indiscernibility relation. We assume through this paper that R is an equivalence relation.

Some ideas underlaying the outlined theory are common with fuzzy set theory [15], alternative set theory [13], non standard analysis [12], however we are primarily aiming at creating mathematical foundations for artificial intelligence, and not a new set theory or analysis.

Some application of the presented ideas will be published elsewhere.

Thanks are due to prof. E. Orłowska and prof. W. Marek for fruitful discussions.

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 indicernibility 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 Com(A). Certainly Com(A) is a Boolean algebra, i.e. the family of all composed set 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 $\overline{Apr}_A(X)$; the greatest composed set in A contained in X, will be called the best lower approximation of X in A, in symbols $\underline{Apr}_A(X)$.
If A is known instead of $\overline{Apr}_A(X)$ $(\underline{Apr}_A(X))$ we shall write $\overline{Apr}(X)$ $(\underline{Apr}(X))$.

The set $Bnd_A(X) = \overline{Apr}_A(X) - \underline{Apr}_A(X)$ (in short Bnd(X)) will be called boundary of X in A.

Sets $\underline{Edg}_A(X) = X - \underline{Apr}_A(X)$ (in short $\underline{Edg}(X)$) and $\overline{Edg}_A(X) = \overline{Apr}_A(X) - X$, (in short $\overline{Edg}(X)$) are refered to as an internal and an external edge of X in A respectively.

Of course $Bnd_A(X) = \underline{Edg}_A(X) \cup \overline{Edg}_A(X)$.

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

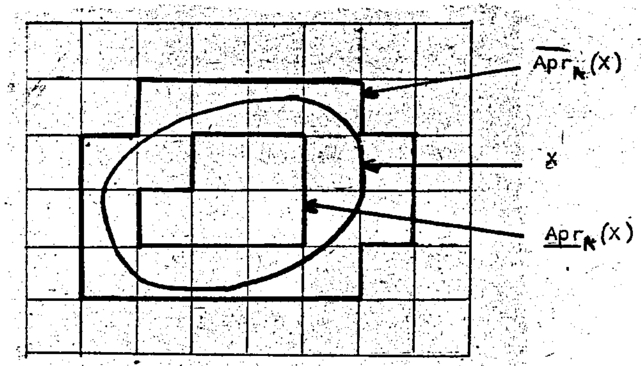

Fig. 1

Let us define two membership functions $\underline{\in}_A, \overline{\in}_A$ (called strong and weak membership respectively), as follows:

$x \underline{\in}_A X$ iff $x \in \underline{Apr}_A(X)$

$x \overline{\in}_A X$ iff $x \in \overline{Apr}_A(X)$

If $x \underline{\in}_A X$ we say "x surely belongs to X in A", and $x \overline{\in}_A X$ is to mean "X possibly belongs to X in A". Thus we can interprete approximations as counterparts of necessity and possibility in modal legre.

Let $A = (U,R)$ be an approximation space and let $F = \{X_1, X_2, \ldots\}$, $X_i \subset U$, be a certain family of subsets of U. The lower and upper approximation of the family F in A (in symbols $\underline{Apr}_A(F)$, $\overline{Apr}_A(F)$) are defined as

$$\underline{Apr}_A(F) = \{\underline{Apr}_A(X_1), \underline{Apr}_A(X_2), \ldots\}$$

$$\overline{Apr}_A(F) = \{\overline{Apr}_A(X_1), \overline{Apr}_A(X_2), \ldots\}$$

respectively.

1.2. Approximation space and topological space

It is easy to check that the approximation space $A = (U,R)$ defines uniquely the topological space $T(A)$ (in short T_A), where $T_A = (U, Com(A))$, and $Com(A)$ 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 $Com(A)$ is both the set of all open and closed sets in T_A. Thus $\underline{Apr}_A(X)$ and $\overline{Apr}_A(X)$ can be interpreted as a interior and closure of the set X in the topological space T_A, respectively.

If $\overline{Apr}_A(X) = \underline{Apr}_A(X)$ for every $X \subset U$, then $A = (U,R)$ will be called a <u>discrete</u> approximation space.

One can easily check that if A is a discrete approximation space, then all atoms in A are one element sets.

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

1.3. Properties of approximations

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

(A1) $\overline{Apr}_A(X) \supset X \supset \underline{Apr}_A(X)$

(A2) $\underline{Apr}_A(1) = \overline{Apr}_A(1) = 1$

(A3) $\underline{Apr}_A(0) = \overline{Apr}_A(0) = 0$

(A4) $\overline{Apr}_A(\overline{Apr}_A(X)) = \underline{Apr}_A(\overline{Apr}_A(X)) = \overline{Apr}_A(X)$

(A5) $\underline{Apr}_A(\underline{Apr}_A(X)) = \overline{Apr}_A(\underline{Apr}_A(X)) = \underline{Apr}_A(X)$

(A6) $\overline{Apr}_A(X \cup Y) = \overline{Apr}_A(X) \cup \overline{Apr}_A(Y)$

(A7) $\underline{Apr}_A(X \cap Y) = \underline{Apr}_A(X) \cap \underline{Apr}_A(Y)$

(A8) $\overline{Apr}_A(X) = -\underline{Apr}_A(-X)$

(A9) $\underline{Apr}_A(X) = -\overline{Apr}_A(-X)$

where $-X$ is an abbreviation for $U - X$.

Moreover we have

(B1) $\overline{Apr}_A(X \cap Y) \subset \overline{Apr}_A(X) \cap \overline{Apr}_A(Y)$

(B2) $\underline{Apr}_A(X \cup Y) \supset \underline{Apr}_A(X) \cap \underline{Apr}_A(Y)$

(B3) $\overline{Apr}_A(X - Y) \supset \overline{Apr}_A(X) - \overline{Apr}_A(Y)$

(B4) $\underline{Apr}_A(X - Y) \subset \underline{Apr}_A(X) - \underline{Apr}_A(Y)$

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

(C1) $\overline{Apr}_A(X) \cup \underline{Apr}_A(-X) = 1$

(C2) $\overline{Apr}_A(X) \cup \overline{Apr}_A(-X) = 1$

(C3) $\underline{Apr}_A(X) \cup \overline{Apr}_A(-X) = 1$

(C4) $\underline{Apr}_A(X) \cup \underline{Apr}_A(-X) = -\, Bnd_A(X)$

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

(D1) $\overline{Apr}_A(X) \cap \underline{Apr}_A(-X) = 0$

(D2) $\overline{Apr}_A(X) \cap \overline{Apr}_A(-X) = Bnd_A(X)$

(D3) $\underline{Apr}_A(X) \cap \underline{Apr}_A(-X) = 0$

(D4) $\underline{Apr}_A(X) \cap \overline{Apr}_A(-X) = 0$

Moreover we have

(E1) If $X \subset Y$, then $\overline{Apr}_A(X) \subset \overline{Apr}_A(Y)$ and $\underline{Apr}_A(X) \subset \underline{Apr}_A(Y)$, for every 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 universum, 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 $R' \subset R$, we say that the space A' if <u>finer</u> than the space A or that the space A is <u>coarser</u> then the space A'.

If A' is finer than A, then

$$\underline{Apr}_{A'}(X) \supset \underline{Apr}_A(X)$$

$$\overline{Apr}_{A'}(X) \subset \overline{Apr}_A(X)$$

$$Bnd_{A'}(X) \subset Bnd_A(X)$$

for every $X \subset U$.

We say that an approximation space A is a <u>refinement</u> 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 a discrete approximation space is the finest possible approximation space.

1.5. Sample of a set

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

We say that Y $(Y \subset X)$ is a <u>lower</u> (<u>upper</u>) <u>sample</u> of X in A if $\overline{Apr}_A(Y) = \underline{Apr}_A(X)$ $(\overline{Apr}_A(Y) = \overline{Apr}_A(X))$.

A lower (upper) sample Y of X in A is <u>proper</u> if Y is the smallest sample of X in A. Thus each lower (upper) sample of X in A has some elements in common with every atom of $\underline{Apr}_A(X)$ $(\overline{Apr}_A(X))$; the proper lower (upper) sample of X in A has exactly one element in common with every atom in $\underline{Apr}_A(X)$ $(\overline{Apr}_A(X))$.

If A is a discrete approximation space, then for every $X \subset U$, X is identical with its sample, and X is proper sample of X in A.

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 $X \subset U$.

By $\underline{\mu}_A(X)$ $(\overline{\mu}_A(X))$ we denote the number of atoms in $\underline{Apr}_A(X)$ $(\overline{Apr}_A(X))$, and we call $\underline{\mu}_A(X)$ $(\overline{\mu}_A(X))$ the <u>internal</u> (<u>external</u>) <u>measure</u> of X in A.

If $\underline{\mu}_A(X) = \overline{\mu}_A(X)$ we say that X is <u>measurable</u> in A.

Thus the set X is measurable in A if and only if X is a composed set in A.

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

By the <u>accuracy</u> of approximation of X in A we mean the number

$$\eta_A(X) = \frac{\underline{\mu}_A(X)}{\overline{\mu}_A(X)} \quad ;$$

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

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

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 any $X \subset U$

$$\eta_A(X) \geqslant \eta_{A^*}(X).$$

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

$$\eta_A(X) = \frac{\overline{\mu}_A(Y')}{\overline{\mu}_A(Y'')}$$

1.7. Examples

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

<u>Example 1</u>. Let $\mathcal{R}^+$ be the set of non-negative real numbers, and let S be the indiscernibility relation on $\mathcal{R}^+$ defined by the following partition:

$$\langle 0,1\rangle, \langle 1,2\rangle, \langle 2,3\rangle, \ldots$$

where $\langle i,i+1\rangle$, $i = 0,1,2,\ldots$ denotes a half-opened interval. The corresponding approximation space will be denoted as $A = (\mathcal{R}^+,S)$.

In that approximation space we have for example,

$$\underline{Apr}\{1\} = \emptyset$$
$$\overline{Apr}\{1\} = \langle 0,1\rangle$$
$$\underline{Apr}(1\tfrac{1}{2}, 2\tfrac{1}{2}) = \emptyset$$
$$\overline{Apr}(1\tfrac{1}{2}, 2\tfrac{1}{2}) = \langle 1,3\rangle$$

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

$$\underline{Apr}(N) = \emptyset$$
$$\overline{Apr}(N) = \mathcal{R}^+$$

Thus the set of natural numbers is a disperse set in A, and N is the proper upper sample of $\mathcal{R}^+$.

<u>Example 2</u>. Let $A = (\mathcal{R}^+,S)$ 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 $n \geqslant 0$.

From the definition we have

$$\underline{Apr}(0,r) = \bigcup_{i=0}^{n-1} \langle i,i+1\rangle = \langle 0,n\rangle, \text{ for } n \geqslant 1 \text{ and } \emptyset \text{ for } n = 0.$$

$$\overline{Apr}(0,r) = \bigcup_{i=0}^{n} \langle i,i+1\rangle = \langle 0,n+1\rangle .$$

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

$$\underline{\mu}(0,r) = n$$
$$\overline{\mu}(0,r) = n+1,$$

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

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

Thus we can interprete the approximation space $A = (\mathcal{Z}^{+},S)$ as a measurement system, where

$\overline{\mu}_A\langle i,i+1\rangle = \underline{\mu}_A\langle i,i+1\rangle = 1$, $i = 0,1,\ldots$

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

Example 3. Let V be a finite set called a vocabulary and let V^* be the set of all finite sequences over V. Any subset of V^* will be called a language over V.

Let $R \subset V^* \times V^*$ be a indiscernibility relation, and let $A = (V^*,R)$ be an approximation space defined by V^* and R.

A language $L \subset V^*$ is recognizable in A if $\underline{Apr}_A(L) = \overline{Apr}_A(L)$.
The family of all recognizable languages in A, denoted as Rec(A), is the topology induced by $A = (V^*,R)$ and the base of the topology is V^*/R.

Example 4. Let $S = \langle X,A,V,\varsigma\rangle$ be an information system (see [10]), where

X - is the set of objects

A - is the set of attributes

$V = \bigcup V_a$, V_a - is the set of values of attribute $a \in A$

$\varsigma : X \times A \rightarrow V$ is a information function. $\varsigma_x : A \rightarrow V$, $x \in X$ is called an information about x in S, where

$\varsigma_x(a) = \varsigma(x,a)$

for every $x \in X$ and $a \in A$.

We define a binary relation $\widetilde{S}$ over X in the following way:

$x \underset{S}{\sim} y$ iff $\varsigma_x = \varsigma_y$.

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

Any subset $Y \subset X$ is called describable in S iff $\underline{Apr}_A(Y) = \overline{Apr}_A(Y)$. The set of all describable sets in S, denoted as Des(S), is a topology induced by S on X, and the base of the topology is $X/\widetilde{S}$.

Example 5. A family F of subsets of the set U is called an ideal if the following conditions are satisfied:

(i) if $X \in F$ and $Y \subset X$, then $Y \in F$

(ii) if $X \in F$ and $Y \in F$, then $X \cup Y \in F$.

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

(iii) if $X \in G$ and $X \subset Y$, then $Y \in G$.

(iv) if $X \in G$ and $Y \in G$ then $X \cap Y \in G$.

Let $A = (U,R)$ be an approximation space and let F be an ideal in U; then

(a) $\underline{Apr}(F)$ is an ideal

(b) $\overline{Apr}(F)$ is an ideal

Ad a). Because if $Y \subset X$, then $\underline{Apr}(Y) \subset \underline{Apr}(X)$, and because R is an ideal, condition (i) is satisfied, i.e. $\underline{Apr}(X) \in \underline{Apr}(F)$, $\underline{Apr}(Y) \subset \underline{Apr}(X)$ and $\underline{Apr}(Y) \in \underline{Apr}(F)$.

From the assumption that F is an ideal we have that $\underline{Apr}(X) \in \underline{Apr}(F)$, $\underline{Apr}(Y) \in \underline{Apr}(F)$ and $\underline{Apr}(X \cup Y) \in \underline{Apr}(F)$. Because $\underline{Apr}(X) \cup \underline{Apr}(Y) \in \underline{Apr}(F)$, then condition (ii) is satisfied and $\underline{Apr}(F)$ is an ideal.

Ad b) By similar reasoning as in a) we obtain that $\overline{Apr}(F)$ is an ideal.

If $A = (U,R)$ is an approximation space and G is a filter in U, then

(c) $\underline{Apr}(G)$ is a filter

(d) $\overline{Apr}(G)$ is not a filter.

Ad c) In a similar reasoning as before we obtain that $\underline{Apr}(G)$ is a filter.

Ad d) In the same way we can show that condition (iii) holds however condition (iv) in not satisfied, thus $\overline{Apr}(G)$ is not a filter.

Example 6. Let $A = (U,R)$ be an approximation space, and let Q be a topology on U, i.e. Q is a family of subsets of U, satisfying the conditions:

(i) $\emptyset \in Q$, $U \in Q$

(ii) if $X \in Q$ and $Y \in Q$, then $X \cap Y \in Q$

(iii) if $P \subset Q$, then $\bigcup P \in Q$

If Q is a topology on U and $A = (U,R)$ is an approximation space, then neither $\underline{Apr}_A(Q)$ nor $\overline{Apr}_A(Q)$ is a topology on U.

Suppose Q is a topology on U, then

(a) $\underline{Apr}(\emptyset) \in \underline{Apr}(Q)$; $\underline{Apr}(U) \in \underline{Apr}(Q)$

(b) $\underline{Apr}(X) \in \underline{Apr}(Q)$, $\underline{Apr}(Y) \in \underline{Apr}(Q)$ and $\underline{Apr}(X \cap Y) \in \underline{Apr}(Q)$ and consequently by (A7) $\underline{Apr}(X) \cap \underline{Apr}(Y) \in \underline{Apr}(Q)$,

(c) If $P \subset Q$ and $\bigcup P \in Q$, then $\underline{Apr}(P) \subset \underline{Apr}(Q)$ and $\underline{Apr}(\bigcup P) \in \underline{Apr}(Q)$; in virtue of (B2) $\bigcup \underline{Apr}(P) \subset \underline{Apr}(\bigcup P)$, and consequently $\underline{Apr}(Q)$ is not a topology on U.

$\overline{Apr}(Q)$ is not a topology because from (B1) follows that if $\overline{Apr}(X) \in \overline{Apr}(Q)$, and $\overline{Apr}(Y) \in \overline{Apr}(Q)$, then $\overline{Apr}(X) \cap \overline{Apr}(Y)$ may not belong to $\overline{Apr}(Q)$, i.e. condition (ii) may be not satisfied.

2. ROUGH EQUALITY OF SETS

2.1. Basic definitions

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

(a) The sets X,Y are <u>roughly bottom</u> - <u>equal</u> in A, in symbols $X \underset{A}{\eqsim} Y$, iff

$$\underline{Apr}_A(X) = \underline{Apr}_A(Y).$$

(b) The sets X,Y are <u>roughly top</u> - <u>equal</u> in A, in symbols $X \underset{A}{\simeq} Y$, iff

$$\overline{Apr}_A(X) = \overline{Apr}_A(Y).$$

(c) The sets X,Y are <u>roughly equal</u> in A, in symbols $X \underset{A}{\approx} Y$, iff

$$X \underset{A}{\eqsim} Y \text{ and } X \underset{A}{\simeq} Y.$$

It is easy to check that $\underset{A}{\eqsim}$, $\underset{A}{\simeq}$, $\underset{A}{\approx}$ are equivalence relations on $P(U)$.

In what follows we shall omit the subscript A if the approximation space A is understood - and write $\eqsim$, $\simeq$, $\approx$, instead $\underset{A}{\eqsim}$, $\underset{A}{\simeq}$, $\underset{A}{\approx}$.

2.2. Properties of rough equality

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

(A1) If $X \eqsim Y$, then $X \cap Y \eqsim X \eqsim Y$

(A2) If $X \simeq Y$, then $X \cup Y \simeq X \simeq Y$

(A3) If $X \simeq X'$ and $Y \simeq Y'$, then $X \cup Y \simeq X' \cup Y'$

(A4) If $X \eqsim X'$ and $Y \eqsim Y'$, then $X \cap Y \eqsim X' \cap Y'$

(A5) If $X \eqsim Y$, then $X - Y \eqsim 0$

(A6) $X - Y \simeq 0$ iff $X = Y$

(A7) If $X \eqsim Y$, then $-(-X) \eqsim Y$

(A8) If $X \simeq Y$, then $-(-X) \simeq Y$

(A9) If $X \approx Y$, then $-(-X) \approx Y$

(A10) If $X \eqsim Y$, then $X \cup -Y \eqsim 1$

(A11) If $X \simeq Y$, then $X \cup -Y \simeq 1$

(A12) If $X \simeq Y$, then $X \cap -Y \simeq 0$

(A13) If $X \eqsim Y$, then $X \cap -Y \eqsim 0$

Set X will be called roughly dense (r-dense) in A if $X \underset{A}{\simeq} 1$.

Set X will be called roughly co-dense (r-co-dense) in A if $X \underset{A}{\eqsim} 0$.

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 $X \subset Y$ and $Y \eqsim 0$, then $X \eqsim 0$

(B2) If $X \subset Y$ and $X \simeq 1$, then $Y \simeq 1$

(B3) If $X \simeq 1$, then $-X \eqsim 0$

(B4) If $X \eqsim 0$, then $-X \simeq 1$

(B5) If X is a r-dispersed set, then so is $-X$.

(B6) $Y \cap X \eqsim 0$ iff $X \eqsim 0$ or $Y \eqsim 0$

(B7) $Y \cup X \simeq 1$ iff $X \simeq 1$ or $Y \simeq 1$

(B8) If X,Y are both r-dense, then $X \simeq Y$

(B9) If X,Y are both r-co-dense then $X \eqsim Y$

(B10) If X,Y are both r-dispersed then $X \approx Y$.

(B11) $\overline{Apr}_A(X)$ is the union of all sets Y such that $X \underset{A}{\simeq} Y$

(B12) $\underline{Apr}_A(X)$ is the intersection of all sets Y, such that $X \underset{A}{\eqsim} Y$.

3. ROUGH INCLUSION OF SETS

3.1. Basic definitions

Let $A = (U,R)$ be an approximation space and let $X,Y \subset U$. We introduce the following definitions

(a) we say that X is roughly bottom-included in Y, in A, in symbols $X \underset{A}{\underset{\sim}{\subset}} Y$, if $\underline{Apr}_A(X) \subset \underline{Apr}_A(Y)$.

(b) We say that X is roughly top-included in Y, in A in symbols $X \underset{A}{\overset{\sim}{\subset}} Y$, if $\overline{Apr}_A(X) \subset \overline{Apr}_A(Y)$.

(c) We say that X is roughly included (r-included) in Y, in A in symbols $X \underset{A}{\underset{\sim}{\overset{\sim}{\subset}}} Y$ if $X \underset{A}{\underset{\sim}{\subset}} Y$ and $X \underset{A}{\overset{\sim}{\subset}} Y$.

If A is understood instead of $X \underset{A}{\underset{\sim}{\subset}} Y$, $X \underset{A}{\overset{\sim}{\subset}} Y$, $X \underset{A}{\underset{\sim}{\overset{\sim}{\subset}}} Y$, we shall write $X \underset{\sim}{\subset} Y$, $X \overset{\sim}{\subset} Y$, $X \underset{\sim}{\overset{\sim}{\subset}} Y$, respectively.

If $X \underset{A}{\overset{\sim}{\subset}} Y$, X is called rough upper-subset of Y in A;

If $X \underset{A}{\underset{\sim}{\subset}} Y$, X is called rough lower-subset of Y in A:

If $X \underset{A}{\underset{\sim}{\overset{\sim}{\subset}}} Y$, X is called rough subset of Y in A.

One can easily check that all rough inclusions $\underset{\sim}{\subset}$, $\overset{\sim}{\subset}$, $\underset{\sim}{\overset{\sim}{\subset}}$, are ordering relations.

3.2. Properties of rough inclusions

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

(A1) If $X \subset Y$, then $X \underset{\sim}{\subseteq} Y$, $X \overset{\sim}{\subset} Y$, $X \underset{\sim}{\overset{\sim}{\subseteq}} Y$

(A2) If $X \underset{\sim}{\subseteq} Y$ and $Y \underset{\sim}{\subseteq} X$, then $X \underset{\sim}{=} Y$

(A3) If $X \overset{\sim}{\subset} Y$ and $Y \overset{\sim}{\subset} X$, then $X \overset{\sim}{=} Y$

(A4) If $X \underset{\sim}{\overset{\sim}{\subseteq}} Y$ and $X \underset{\sim}{\overset{\sim}{\supseteq}} Y$, then $X \approx Y$

(A5) $X \overset{\sim}{\subset} Y$ iff $X \cup Y \overset{\sim}{=} Y$

(A6) $X \underset{\sim}{\subseteq} Y$ iff $X \cap Y \underset{\sim}{=} Y$

(B1) If $X \subset Y$ and $X \underset{\sim}{=} X'$, $Y \underset{\sim}{=} Y'$ then $X' \underset{\sim}{\subseteq} Y'$

(B2) If $X \subset Y$ and $X \overset{\sim}{=} X'$, $Y \overset{\sim}{=} Y'$ then $X' \overset{\sim}{\subset} Y'$

(B3) If $X \subset Y$, and $X \approx X'$, $Y \approx Y'$ then $X' \underset{\sim}{\overset{\sim}{\subseteq}} Y'$

(C1) If $X \overset{\sim}{\supset} X'$ and $Y \overset{\sim}{\supset} Y'$, then $X \cup Y \overset{\sim}{\supset} X' \cup Y'$

(C2) If $X \underset{\sim}{\supseteq} X'$ and $Y \underset{\sim}{\supseteq} Y'$, then $X \cap Y \underset{\sim}{\supseteq} X' \cap Y'$

(C3) $X \cap Y \underset{\sim}{\subseteq} X \overset{\sim}{\subset} X \cup Y$

(C4) If $X \underset{\sim}{\subseteq} Y$ and $X \underset{\sim}{=} Z$, then $Z \underset{\sim}{\subseteq} Y$

(C5) If $X \overset{\sim}{\subset} Y$ and $X \overset{\sim}{=} Z$, then $Z \overset{\sim}{\subset} Y$

(C6) If $X \underset{\sim}{\overset{\sim}{\subseteq}} Y$ and $X \approx Z$, then $Z \underset{\sim}{\overset{\sim}{\subseteq}} Y$.

3.3. Rough power sets

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

Thus

$$P_{\underline{A}}(X) = \{Y : Y \underset{\underline{A}}{\subseteq} X\}$$

$$P_{\bar{A}}(X) = \{Y : Y \underset{\bar{A}}{\overset{\sim}{\subset}} X\}$$

$$P_{A}(X) = \{Y : Y \underset{A}{\overset{\sim}{\subseteq}} X\}.$$

It is easy to see that

$$P(X) \subset P_{\underline{A}}(X)$$

$$P(X) \subset P_{\bar{A}}(X)$$

$$P(X) \subset P_{A}(X)$$

4. ROUGH SETS

4.1. Basic notions

Let $A = (U,R)$ be an approximation space, and let $\underset{\underline{A}}{\eqsim}$, $\underset{\bar{A}}{\simeq}$, $\underset{A}{\approx}$, be equivalence relations on $P(U)$.

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

$$\underline{A}^* = (P(U), \underset{\underline{A}}{\eqsim})$$

$$\bar{A}^* = (P(U), \underset{\bar{A}}{\simeq})$$

$$A^* = (P(U), \underset{A}{\approx})$$

in which objects are subsets of U and the relations $\underset{\underline{A}}{\eqsim}$, $\underset{\bar{A}}{\simeq}$, $\underset{A}{\approx}$ are the indiscernibility relations in the corresponding spaces $\underline{A}^*$, $\bar{A}^*$, A^*.

The approximation space $A^*(\underline{A}^*, \bar{A}^*)$ will be called the extension (lower, upper) of A.

Equivalence classes of the relation $\underset{A}{\approx}$ ($\underset{\underline{A}}{\eqsim}$, $\underset{\bar{A}}{\simeq}$) 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 $\underset{A}{\approx}$ ($\underset{\underline{A}}{\eqsim}$, $\underset{\bar{A}}{\simeq}$).

Every approximation space $\underline{A}^*$, $\bar{A}^*$, A^* induces a topology $Com(\underline{A}^*)$, $Com(\bar{A}^*)$, $Com(A^*)$ respectively, and consequently - the topological spaces

$T_{\underline{A}^*} = (P(U),\ Com(\underline{A}^*))$

$T_{\overline{A}^*} = (P(U),\ Com(\overline{A}^*))$

$T_{A^*} = (P(U),\ Com(A^*))$,

and $P(U)/\overline{\approx}_A$, $P(U)/\simeq_A$, $P(U)/\approx_A$ are the bases for the corresponding topological spaces.

Thus we are able to introduce the upper and lower approximations of a family of subsets of U in $\underline{A}^*$, $\overline{A}^*$, A^*, i.e., if $\mathfrak{X}$ is a certain family of subsets of U we can introduce the following approximations of $\mathfrak{X}$ in $\underline{A}^*$, $\overline{A}^*$, A^*

$\underline{Apr}_{\underline{A}^*}(\mathfrak{X})$, $\overline{Apr}_{\underline{A}^*}(\mathfrak{X})$

$\underline{Apr}_{\overline{A}^*}(\mathfrak{X})$, $\overline{Apr}_{\overline{A}^*}(\mathfrak{X})$

$\underline{Apr}_{A^*}(\mathfrak{X})$, $\overline{Apr}_{A^*}(\mathfrak{X})$.

In other words, if $X, Y \subset U$, and $X \approx_A Y$ ($X \overline{\approx}_A Y$, $X \simeq_A Y$) we say that X and Y are **close** (bottom, top) in A, otherwise the sets X, Y are remote in A.

Given any family $\mathfrak{X}$ of subsets of U we can classify members of the family $\mathfrak{X}$ according to the relation $\approx_A$ ($\overline{\approx}_A$, $\simeq_A$). The classification is possible if $\mathfrak{X}$ is composed family in A^* ($\overline{A}^*$, $\underline{A}^*$), otherwise the classification is impossible. In the letter case we can classify only the upper and lower approximation of the family $\mathfrak{X}$ in the approximation space A^* ($\underline{A}^*$, $\overline{A}^*$).

4.2. Extensions of higher order

In a similar way as before we can extend each of the approximation spaces A^*, $\overline{A}^*$, $\underline{A}^*$ 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:

$$A^{k+1} = (P^{k+1}(U),\ \approx_{A^k})$$

$$\underline{A}^{k+1} = (P^{k+1}(U),\ \overline{\approx}_{A^k})$$

$$\overline{A}^{k+1} = (P^{k+1}(U),\ \simeq_{A^k}),$$

where $P^k(U)$ is defines 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 uniquelly infinite sequence of approximation space of higher orders, allowing to cluster sets, families of sets etc., but we shall not discuss that problem here.

4.3. Ordering of rough sets

Every approximation space $A = (U,R)$ induces three relations $\leq_A$, $\ll_A$, $\lll_A$ (in short $\leq$, $\ll$, $\lll$) on the families $P(U)/\approx_A$, $P(U)/\simeq_A$, $P(U)/\overline{\approx}_A$ respectively.

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

(a) $[X]_{\approx} \leq [Y]_{\approx}$ iff $X \subseteq_{\approx} Y$

(b) $[X]_{\simeq} \ll [Y]_{\simeq}$ iff $X \tilde{\subset} Y$

(c) $[X]_{\overline{\approx}} \lll [Y]_{\overline{\approx}}$ iff $X \tilde{\subseteq} Y$

One can easy check that $\leq$, $\preccurlyeq$, $\preccurlyeq'$ are ordering relation on $P(U)/\approx$, $P(U)/\simeq$, $P(U)/\simeq'$ respectively.

Thus

$$\underline{P}^{*}(U) = (P(U)/\approx, \leq)$$
$$\overline{P}^{*}(U) = (P(U)/\simeq, \preccurlyeq)$$
$$P^{*}(U) = (P(U)/\simeq', \preccurlyeq')$$

are partially ordered families of sets.

The smallest element in $\underline{P}^{*}(U)$ is $[0]_{\approx}$, i.e. the family of all co-dense sets in the topological space T_A. The greatest element in $\underline{P}^{*}(U)$ is $[1]_{\approx} = 1$, i.e. the class consisting only of one set, the universum U.

The smallest element in $\overline{P}^{*}(U)$ is $[0]_{\simeq} = 0$, i.e. the class consisting of only one set, the empty set; the greatest element in $\overline{P}^{*}(U)$ is $[1]_{\simeq}$, i.e. the family of all dense sets in the topological space T_A.

The smallest element in $P^{*}(U)$ is $[0]_{\simeq'} = 0$, and the greatest element in $P^{*}(U)$ is $[1]_{\simeq'} = 1$.

Finally let us remark that if an approximation space $A = (U,R)$ has k atoms (elementary sets), then there are 2^k equivalence classes in a family $P(U)/\underset{A}{\approx}$, (and $P(U)/\underset{A}{\simeq}$), but there are

$$\sum_{i=1}^{k}\binom{k}{i} 2^{k-i} = 3^k$$

equivalence classes in a family $P(U)/\underset{A}{\simeq'}$.

References

[1] Engelking, R., General topology, PWN, 1977

[2] Konrad, E., Orłowska, E., Pawlak, Z., Knowledge representation systems, ICS PAS Reports (1981) No 433

[3] Konrad, E., Orłowska, E., Pawlak, Z., An approximate concept learning, Bericht 81-7 October 81, Berlin

[4] Marek, W., Pawlak, Z., Rough sets and information systems, ICS PAS Reports, (1981) No 441

[5] Orłowska, E., Pawlak, Z., Expressive power of knowledge representation systems, ICS PAS Reports (1981) No 432

[6] Pawlak, Z., Classification of objects by means of attributes, ICS PAS Reports (1981) No 429

[7] Pawlak, Z., Rough sets, ICS PAS Reports (1981) No 431

[8] Pawlak, Z., Rough relations, ICS PAS Reports (1981) No 435

[9] Pawlak, Z., Rough functions, ICS PAS Reports (to appear)

[10] Pawlak, Z., Information systems, theoretical foundations (in Polish) book (to appear)

[11] Pawlak, Z., Some remarks about rough sets, ICS PAS Reports (1982) No 456

[12] Robinson, A., Non-standard analysis, North-Holland Publishing Company, Amsterdam, 1966

[13] Vopenka, P., Mathematics in the Alternative Set Theory, Teubner-Texte zur Mathematik, Leipzig, 1979

[14] Żakowski, W., On a concept of rough sets, Demonstratio Mathematicae (to appear)

[15] Zadeh, L.A., Fuzzy sets, Information and Control, (1965), 8 pp. 338-353.

C O N T E N T S