Rough Sets,
Rough Relations and Rough Functions

by

Zdzisław Pawlak

ICS Research Report 24/94

Warsaw, May 1994

Rough Sets, Rough Relations and Rough Functions

Zdzislaw Pawlak
Institute of Computer Science
Warsaw University of Technology
ul. Nowowiejska 15/19, 00 665 Warsaw, Poland

and

Institute of Theoretical and Applied Informatics
Polish Academy of Sciences
ul. Baltycka 5, 44 000 Gliwice, Poland

1   Introduction

It is well known that in classical set theory in order to define a set we have to specify its membership function $\mu_X : U \to \{0, 1\}$, such that $\mu_X(x) = 1$ if and only if x belongs to X and $\mu_X(x) = 0$ otherwise, i.e. $\mu_X(x) \in \{0, 1\}$ for every $x \in U$. For fuzzy sets (cf. [12]) and multisets $\mu_X(x) \in [0, 1]$ and $\mu_X(x) \in \{0, 1, 2, \ldots\}$ - respectively.

Another philosophy of defining sets is assumed in the rough set theory (cf. [3]). Basic information about rough sets can be found in [2] and [11]. The relationship of rough sets to other approaches dealing with similar problems is considered in [8], and [9]. Besides, several extensions of the rough set theory have been proposed, for example in [1] and [13].

The rough set theory is based on the assumption that we have initially some information (knowledge) about elements of the universe. Because with some elements the same information can be associated, hence two different elements can be indiscernible in view of the available information. Thus information associated with objects of the universe generates an indiscernibility relation on its elements. The indiscernibility relation is the starting point of the rough set theory and can be employed in two ways in order to define basic concepts of this theory - by defining either approximations or the rough membership function.

In this paper we are going to present some consequences of both approaches when defining rough sets, rough relations and rough functions.

2   Rough Sets and Approximations

Let U be a finite, nonempty set called the *universe*, and let I be a binary relation on U. By $I(x)$ we mean the set of all y such that xIy. If we assume that I is reflexive and symmetric, i.e.

$$xIx, \text{ for every } x \in U,$$

$$xIy \text{ implies } yIx, \text{ for every } x, y \in U,$$

then I is a tolerance relation.

If we assume additionally that the I is transitive, i.e.

$$xIy \text{ and } yIz \text{ implies } xIz, \text{ for every } x, y, z \in U,$$

then I is an equivalence relation and $I(x) = [x]_I$, i.e. - is an equivalence class of the relation I containing element x. I will be refered to as an *indiscernibility* relation.

We will define now two basic operations on sets in the rough set theory, called the *I-lower* and the *I-upper approximation*, and defined respectively by

$$I_*(X) = \{x \in U : I(x) \subseteq X\},$$

$$I^*(X) = \{x \in U : I(x) \cap X \neq \emptyset\}.$$

The difference between the upper and the lower approximation will be called the *I-boundary* of X and will be denoted by $BN_I(X)$, i.e.

$$BN_I(X) = I^*(X) - I_*(X).$$

This is to mean that if we "see" the set X through the information, which generates the indiscernibility I, only the above approximations of X can be "observed", but not the set X. The boundary region expresses how exactly the set X can be "seen" due to the indiscernibility I. If the boundary region is the empty set, X can be "observed" *exactly* trough the indiscernibility relation I, and in the opposite case the set X can be "observed" *roughly (approximately)* only - due to the indiscernibility I. The former sets are *crisp (exact)*, whereas the later - are *rough (inexact)*, with respect to indiscernibility I, or formally set X is *I-exact* iff $BN_I(X) = \emptyset$, i.e. $I^*(X) = I_*(X)$. otherwise the set X is *I-rough*.

Below some properties of approximations are given.

1) $I_*(X) \subseteq X \subseteq I^*(X)$,

2) $I_*(\emptyset) = I^*(\emptyset) = \emptyset, I_*(U) = I^*(U) = U$,

3) $I^*(X \cup Y) = I^*(X) \cup I^*(Y)$,

4) $I_*(X \cap Y) = I_*(X) \cap I_*(Y)$,

5) $X \subseteq Y$ implies $I_*(X) \subseteq I_*(Y)$ and $I^*(X) \subseteq I^*(Y)$,

6) $I_*(X \cup Y) \supseteq I_*(X) \cup I_*(Y)$,

7) $I^*(X \cap Y) \subseteq I^*(X) \cap I^*(Y)$,

8) $I_*(-X) = -I^*(X)$,

9) $I^*(-X) = -I_*(X)$,

10) $I_*(I_*(X)) = I^*(I_*(X)) = I_*(X)$,

11) $I^*(I^*(X)) = I_*(I^*(X)) = I^*(X)$.

It is easily seen that the lower and the upper approximation of a set are interior and closure operations in a topology generated by the indiscernibility relation.

3 Rough Sets and the Membership Function

Employing the concept of indiscernibility we can define the membership function for rough sets, the *rough membership*, as

$$\mu_X^I(x) = \frac{card \ (X \cap I(x))}{card \ I(x)}.$$

Obviously $\mu_X^I(x) \in [0, 1]$.
It can be shown (cf. [7]) that the rough membership function has the following properties:

a) $\mu_X^I(x) = 1$ iff $x \in I_*(X)$,

b) $\mu_X^I(x) = 0$ iff $x \in U - I^*(X)$,

c) $0 < \mu(x) < 1$ iff $x \in BN_I(X)$,

d) If $IND(I) = \{(x, x) : x \in U\}$, then $\mu_X^I(x)$ is the characteristic function of X,

e) If $xIND(I)y$, then $\mu_X^I(x) = \mu_X^I(y)$ provided $IND(I)$ is an equivalence relation,

f) $\mu_{U-X}^I(x) = 1 - \mu_X^I(x)$ for any $x \in U$,

g) $\mu_{X \cup Y}^I(x) \; \max(\mu_X^I(x), \mu_Y^I(x))$ for any $x \in U$,

h) $\mu_{X \cap Y}^I(x) \; \min(\mu_X^I(x), \mu_Y^I(x))$ for any $x \in U$,

i) If X is a family of pair wise disjoint sets of U, then $\mu_{\cup X}^I(x) = \sum_{x \in X} \mu_X^I(x)$ for any $x \in U$, provided that $IND(I)$ is an equivalence relation.

The membership function can be understood as a coefficient which expresses uncertainty of an element x being a member of the set X.

The above assumed membership function, can be used to define the two previously defined approximations of sets, as shown below

$$I_*(X) = \{x \in U : \mu_X^I(x) = 1\},$$

$$I^*(X) = \{x \in U : \mu_X^I(x) > 0\}.$$

Obviously the boundary region is defined now as

$$BN_I(X) = \{x \in U : 0 < \mu_X^I(x) < 1\}.$$

One can see that the both approaches to the definition of the rough set stresses various aspects of the rough set concept. The definition by approximations brings to light the topological structure of rough sets, whereas the membership approach - its numerical properties, which can be interpreted in probabilistic terms as conditional probability of y belonging to $I(x)$ under the condition that y belongs to X.

4 Rough Inclusion and Rough Equality of Sets

Having defined rough sets we can now proceed to define next important concept in the discussed approach, the inclusion of rough sets. We can employ both the approximations as well as the rough membership function to this end.

Suppose we are given two sets $X, Y \subseteq U$ and an indiscernibility relation I on U. We will say that

a) set X is *bottom I-included* in Y, $X \subseteq_{*I} Y$, if and only if $I_*(X) \subseteq I_*(Y)$,

b) set X is *top I-included* in Y, $X \subseteq_I^* Y$, if and only if $I^*(X) \subseteq I^*(Y)$,

c) set X is *roughly I-included* in Y, $X \subseteq_I Y$, if and only if $I_*(X) \subseteq I_*(Y)$ and $I^*(X) \subseteq I^*(Y)$.

If $X \subseteq_I Y$ (or $X \subseteq_{*I} Y, X \subseteq_I^* Y$) we will say that X is *rough I-subset* (*lower I-subset, upper I-subset*) of Y.

It is worthwhile to observe that if I is an equivalence relation on U and there are k equivalence classes in I, i.e. if $card(U/I) = k$, then there are 2^k, both the lower and upper, I-rough subsets of U, but there are

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

rough I-subsets of U (cf. [4])

We can also use the rough membership function to define inclusion of rough sets in the following way

$$X \subseteq_{(I)} Y \quad \text{iff} \quad \mu_X^I(x) \le \mu_Y^I(x) \quad \text{for any} \quad x \in U.$$

Obviously both definitions of inclusion are not equivalent.

Similarly, equality of sets can be defined also using both approaches like in the case of inclusion.

We will say that

a) sets X and Y are *bottom I-equal*, $Y =_{*I} Y$, if and only if $I_*(X) = I_*(Y)$ or $X \subseteq_{*I} Y$ and $Y \subseteq_{*I} X$,

b) sets X and Y are *top I-equal*, $Y =_I^* Y$, if and only if $I^*(X) = I^*(Y)$ or $X \subseteq_I^* Y$ and $Y \subseteq_I^* X$,

c) sets X and Y are *roughly I-equal*, $X =_I Y$, if and only if $I_*(X) = I_*(Y)$ and $I^*(X) = I^*(Y)$.

Using the rough membership function definition of equality of rough sets can be defined as

$$X =_I Y \text{ iff } \mu_X^I(x) = \mu_Y^I(x) \text{ or } X \subseteq_I Y \text{ and } X \supseteq_I Y.$$

Again both definitions of equality of rough sets are not equivalent.

5 Rough Relations

Suppose we are given sets X and Y and P and Q two indiscernibility relations on X and Y respectively. Let $R \subseteq X \times Y$ be any binary relation on $X \times Y$ and let $I = P \times Q$ be the product of indiscernibility relation. The *I-lower* and the *I-upper* approximation of R are defined below respectively

$$I_*(R) = \{(x, y) \in X \times Y : I(x, y) \subseteq R\},$$

$$I^*(R) = \{(x, y) \in X \times Y : I(x, y) \cap R \neq \emptyset\}.$$

The difference $BN_I(R) = I^*(R) - I_*(R)$ will be called the *I-boundary region of R*.

Similarly as in the case of rough sets relation R will be called *I-rough* if and only if $BN_I(R) \neq \emptyset$; otherwise the relation R is *I-exact*.

Of course one can define in a similar way relations with arbitrary number of arguments (cf. [5]).

Rough relations can be also defined also employing the rough membership function (cf. [10]), which will be defined as follows

$$\mu_R^I(x, y) = \frac{card(R \cap I(x, y))}{card I(x, y)}.$$

Obviously $0 \leq \mu_R^I(x, y) \leq 1$ and the relation R is *I-exact* if and only if $\mu_R^I(x, y) = 1$; otherwise the relation is *I-rough*.

Again it is easily seen that both definitions are not equivalent.

6 Rough Functions

It is obvious that the philosophy employed in the definition of rough sets and rough relations cannot be applied directly for definition of rough functions. Bearing in mind practical applications we will restrict our definition to rough real functions. To this end we have first to give definitions of rough sets on the real line, i.e. reformulate the concepts of approximations and the rough membership function referring to the set of reals. Another approach to rough functions has been proposed in [6] but we will not consider that approach in this paper.

Let R^+ be the set of nonnegative reals and let $S \subseteq R^+$ be the following sequence of reals $x_1, x_2, \ldots, x_i, \ldots$ such that $x_1 < x_2 < \ldots < x_i$. S will be called a *categorization of R^+* and the ordered pair $A = (R^+, S)$ will be referred to as an *approximation space*. Every categorization S of R^+ induces partition $\pi(S)$ on R^+ defined as $\pi(S) = \{0, (0, x_1), x_1, (x_1, x_2), x_2, (x_2, x_3), x_3, \ldots, x_i, (x_i, x_{i+1}), x_{i+1} \ldots\}$, where (x_i, x_{i+1}) denotes an open interval. By $S(x)$ we will denote block of the partition $\pi(S)$ containing x. In particular, if $x \in S$ then $S(x) = \{x\}$. Let $x \in (x_i, x_{i+1})$. By $\overline{S}(x)$ we denotes the closed interval $< x_i, x_{i+1} >$, called the *closure $S(x)$*. In what follows we will be interested in approximating closed intervals of the form $< 0, x >= Q(x)$ for any $x \in R^+$.

Suppose we are given an approximation space $A = (R^+, S)$. (Let us remark that the categorization S can be viewed as an indiscernibility relation defined on R^+).

By the the *S-lower* and the *S-upper* approximation of $Q(x)$, denoted by $S_*(Q(x))$ and $S^*(Q(x))$ respectively, we mean sets defined below:

$$S_*(Q(x)) = \{y \in R^+ : S(y) \subseteq Q(x)\}$$

$$S^*(Q(x)) = \{y \in R^+ : S(y) \cap Q(x) \neq \emptyset\}.$$

The above definitions of approximations of interval $< 0, x >$ can be understood as approximations of the real number x which are simple the ends of the interval $S(x)$. If $X \subseteq R$, then $\Delta(X) = Sup_{x,y \in X} |x - y|$. In particular $\Delta(S(x))$ will be denoted by $\Delta_S(x)$.

In other words given any real number x and a set of reals S, by the *S-lower* and the *S-upper* approximation of x we mean the numbers $S_*(x)$ and $S^*(x)$, which can be defined as

$$S_*(x) = Sup\{y \in S : y \leq x\}$$
$$S^*(x) = Inf\{y \in S : y \geq x\}.$$

We have

$$S(x) = (S_*(x), S^*(x)).$$

We will say that the number x is *exact* in $A = (R^+, S)$ iff $S_*(x) = S^*(x)$, otherwise the number x is *inexact (rough)* in $A = (R^+, S)$. Of course x is exact iff $x \in S$. Thus every inexact number x can be presented as pair of exact numbers $S_*(x)$ and $S^*(x)$ or as the interval $S(x)$. For example if N is the set of all non negative integers then every real number x such that non $x \in N$ is inexact in the approximation space $A = (R^+, N)$. In general if $A = (R^+, S)$ is an approximation space then the categorization S can be interpreted as a scale by means of which reals from R are measured with some approximation due to the scale S.

The introduced ideas of the rough set on the real line correspond exactly to those defined for arbitrary sets and can be seen as a special case of the general definition.

Now we give the definition of the next basic notion in the rough set approach - the rough membership function - referring to the real line.

The rough membership function for the set of reals will have the form

$$\mu_{Q(x)}(y) = \frac{\Delta(Q(x) \cap S(y))}{\Delta_S(y)}$$

The membership function $\mu_{Q(x)}(y)$ says to what degree any element y belongs to the interval $Q(x)$, or in other words it can be interpreted as the degree to which $x \leq y$.

Now we are ready to give the definition of a *rough real function*, in short *rough function*.

Suppose we are given a real function $f : X \to Y$, where both X and Y are sets of non negative reals and let $A = (X, S)$ and $B = (Y, P)$ be two approximation spaces.

By the (S, P) *lower approximation* of f we understand the function $f_* : X \to Y$ such that

$$f_*(x) = P_*(f(x)) \text{ for every } x \in X.$$

Similarly the $(S, P) - $ *upper approximation* of f is defined as

$$f^*(x) = P^*(f(x)) \text{ for every } x \in X.$$

We say that a function f is *exact* in x iff $f_*(x) = f^*(x)$; otherwise the function f is *inexact (rough)* in x. The number $f^*(x) - f_*(x)$ is the error of approximation of f in x.

Many basic concepts concerning functions can be expressed also in the rough function theory. For example the rough continuity of function can be defined as follows.

A function f is $(S, P) - $ *continuous (roughly continuous)* in x iff

$$f(\overline{S}(x)) \subseteq \overline{P}(f(x)).$$

If f is roughly continuous in x for every $x \in X$ we say that f is $(S, P) - continuous$ (*roughly continuous*).

The intuitive meaning of this definition is obvious. Whether the function is roughly continuous or not depends on the information we have about the function, i.e. it depends how exactly we "see" the function through the availably information (the indiscernibility relation).

7 Conlusions

We have tried in this paper to point out some problems occurring in the rough set theory, when defining basic concepts such as rough set, rough relation and rough function. There is no unique way to define these concepts, and either approximations or rough membership can be applied to this end. In general, both approaches are not equivalent. The first approach stresses the topological character of the concepts involved, whereas the second shows their numerical structure which can be sometimes interpreted in probabilistic terms. To understand better the relationship between both approaches further inquiry is necessary.

References

[1] Pawlak, Z., Wong S.K.M. and Ziarko, W. (1988). Rough Sets: Probabilistic Versus Deterministic Approach, *Int. J. Man-Machine Studies, 29*, pp. 81-95.

[2] Pawlak, Z. (1991). Rough Sets - Theoretical Aspects of Reasoning about Data. KLUWER ACADEMIC PUBLISHERS.

[3] Pawlak, Z. (1982a). Rough Sets, *Int. J. of Inf. and Comp. Sci.*, 11, 5, pp. 341-356.

[4] Pawlak, Z. (1982b). Rough Sets, Power Set Hierarchy, *ICS PAS Reports*, 470, 1982.

[5] Pawlak, Z. (1986). Rough Relations, *Bull. PAS, Tech. Ser.* Vol. 34, No.9-10, pp. 557-590.

[6] Pawlak, Z. (1987). Rough Functions, *Bull. PAS, Tech.* Ser.Vol. 35, No.5-6, pp. 249-251.

[7] Pawlak, Z. and Skowron, A. (1993). Rough Membership Functions. In: Yaeger, R.R., Fedrizzi, M., and Kacprzyk, J. (Eds.), *Advances in the Dempster Shafer Theory of Evidence*, John Wiley and Sons, pp. 251-271.

[8] Skowron, A. and Grzymala-Busse, J. (1992). From the Rough Set Theory to Evidence Theory. In: Yaeger, R.R., Fedrizzi, M., and Kacprzyk, J. (Eds.), *Advances in the Dempster Shafer Theory of Evidence*, John Wiley and Sons, pp. 193-235.

[9] Skowron, A. and Rauszer, C. (1992). The Discernibility Matrices and Functions in Information Systems. In : Slowinski, R. (Ed.) Intelligent Decision Support - Handbook of Advances and Applications of the Rough Set Theory. KLUWER ACADEMIC PUBLISHERS, pp. 311-362.

[10] Skowron, A. and Stepaniuk, J. (1994). Approximation of Relations, *ICS TUW Reports, 20/94*.

[11] Slowinski, R. (Ed.) (1992). Intelligent Decision Support - Handbook of Advances and Applications of the Rough Set Theory. KLUWER ACADEMIC PUBLISHERS.

[12] Zadeh, L. (1965). Fuzzy Sets. *Information and Control* 8, pp.338-353.

[13] Ziarko, W. (1993). Variable Precison Rough Set Model, *Journal of Computer and System Sciences*, vol. 40(1), pp. 39-59.