Rough Real 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

In this paper we are going to present some ideas concernig rough functions outlined in Pawlak (1994). The concept of the rough function seem to be natural extension of the rough set theory and is needed in many applications, where experimental data are processes, in particular as a theoretical basis for rough controllers cf. Czogala et al. (1994), Mrozek et al. (1994) and Plonka et al. (1994).

The presented approach is somehow related to nonstandard analysis (Robinson, 1970) and measurement theory (Orlowska et al., 1984) but these aspects of rough function will be not considered here.

2 Basic of the Rough Set Concept

Basic ideas of the rough set theory can be found in Pawlak (1991). In this section we will give only those notions which are necessary to define concepts used in this paper.

Let U be a finite, nonempty set called the *universe*, and let I be an equivalence relation on U, called an *indiscernibility* relation. By $I(x)$ we mean the set of all y such that xIy, i.e. $I(x) = [x]_I$, i.e.- is an equivalence class of the relation I containing element x. The indiscernibility relation is meant to capture the fact that often we have limited information about elements of the universe and consequently are unable to discern them in view of the available information. Thus I represents our lack of knowledge about U.

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 as follows:

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

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

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.

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

If $I^*(X) = I_*(X)$ we say the set is *I-exact* otherwise the set X is *I-rough.* Thus rough sets are sets with unsharp boundaries.

*This work was supported by grant No. S503 021 06 from State Committee for Scientific Research (Komitet Badan Naukowych

Usually in order to define a set we use the membership function. The membership function for rough sets is defined employing the equivalence relation I as follows:

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

Obviously

$$\mu_X^I(x) \in [0,1].$$

The value of the membership function expresses the degree to which the element x belong to the set X in view of the indiscernibility relation I.

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

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

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

3 Rough Sets on the Real Line

In this section we reformulate the concepts of approximations and the rough membership function referring to the set of reals which are needed to formulate the control problem in the rough set setting.

Let $\mathbf{R}^+$ be the set of nonnegative reals and let $S \subseteq \mathbf{R}^+$ be a 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 $\mathbf{R}^+$ and the ordered pair $A = (\mathbf{R}^+, S)$ will be referred to as an *approximation space*. Every categorization S of $\mathbf{R}^+$ induces partition $\pi(S)$ on $\mathbf{R}^+$ 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 denote the closed interval $[x_i, x_{i+1}]$, called the *closure* of $S(x)$.

In what follows we will be interested in approximating closed intervals of the form $[0, x] = Q(x)$ for any $x \in \mathbf{R}^+$.

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

By 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:

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

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

The above definitions of approximations of the interval $[0, x]$ can be understood as approximations of the real number x which are simple the ends of the interval $S(x)$, therefore we will use the following abbreviations: $S_*(Q(x)) = S_*(x)$ and $S^*(Q(x)) = S^*(x)$. If $X \subseteq \mathbf{R}^+$, then $\Delta(X) = Sup|x - y|, x, y \in X$ and will be called the *length* X of In particular $\Delta(S(x))$ will be denoted by $\Delta_S(x)$ and $\Delta(X)$.

In other words given any real number x and a categorization 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

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

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

We will say that the number x is *exact* in $A = (\mathbf{R}^+, S)$ iff $S_*(x) = S^*(x)$, otherwise the number x is *inexact* (*rough*) in $A = (\mathbf{R}^+, S)$. 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 $\mathbf{N}$ is the set of all non negative integers then every real number x such that non $x \in \mathbf{N}$ is inexact in the approximation space $A = (\mathbf{R}^+, \mathbf{N})$.

In general if $A = (\mathbf{R}^+, S)$ is an approximation space then the categorization S can be interpreted as a scale by means of which reals from $\mathbf{R}^+$ are measured with some approximation due to the scale S.

The introduced ideas of the rough sets 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 $Q(x)$.

4 Rough Functions

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.

Let $A = (\mathbf{R}^+, S)$ be an approximation space and let $\{a_n\}$ be an infinite sequence of reals.

We will say that the sequence (a_n) is *roughly convergent* in A iff there exists i such for every $j > i, \overline{S}(a_j) = \overline{S}(a_i)$. $S_*(a_i)$ and $S^*(a_i)$ are refered to as the *lower* and the *upper limit* of the sequence (a_n).

We will say that the sequence (a_n) is *roughly periodic* in A iff there exists i and k such for every $j > i, \overline{S}(a_j) = \overline{S}(a_{j+k})$. The number k is the *period* of $\{a_n\}$.

Now we give definition of a very important concept, the rough continuity of real function.

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

$$f(S(x)) \subseteq P(f(x)).$$

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

Another exemplary definition concerns monotonicity of functions.

Particularly interesting is the relationship between dependency of attributes in information systems and the rough continuity of functions.

Let $S = (U, A)$, be an *information system*, where U is a finite set of *objects*, called the *universe* and A is a finite set of attributes. With every attribute $a \in A$ a set of *values* of attribute a, called *domain* of a is associated and is donetd by V_a. Every attribute $a \in A$ can be seen as a function $a : U \to V_a$, which to every object $x \in U$ assigns a value of the attribute a. Any subset of attributes $B \subseteq A$ determines the equivalence relation $IND(B) = \{x, y \in U : a(x) = a(y)$, for every $a \in A\}$. Let $B,\ C \subseteq A$. We will say that the set of attributes C *depends* on the set of attributes B, in symbols $B \to C$, iff $IND(B) \subseteq IND(C)$. If $B \to C$ then there exists a *dependency function* $f_{B,C} : V_{b_1} \times V_{b_2} \times \ldots \times V_{b_n} \to V_{c_1} \times V_{c_2} \times \ldots \times V_{c_m}$, such that $f_{B,C}(v_1, v_2, \ldots, v_n) = w_1, w_2, \ldots, w_m)$, iff $\sigma(v_1) \cap \sigma(v_2) \cap, \ldots, \cap \sigma(v_n) \subseteq \sigma(w_1) \cap \sigma(w_2) \cap, \ldots, \cap \sigma(w_m)$, where $v_1 \in V_{b_i}, w_j \in V_{c_j}, \sigma(v) = \{x \in U : a(v) = x\}$ and $v \in V_a$. The dependency function $B \to C$, where $B = \{b_1, b_2, \ldots, b_n\}$ and $C = \{c_1, c_2, \ldots, c_m\}$ assignes uniquelly to every n-tuple of values of attributes from B the m-tuple of values of attributes from C.

There exists the following importanat relationship. $B \to C$ iff $f_{B,C}$ is (B, C) roughly continuous.

Many other basic concepts concerning functions can be expressed also in the rough function setting.

5 Discretization of Rough Functions

The function $f_S : S \to Y$ such that $f_S(x) = f(x)$ for any $x \in S$ will be called a *S-discrete representation* of f or in short *S-discretization* of f.

Our main task is to give interpolation algorithms for discrete representation f_S giving the best approximation of f.

Let us first consider the *linear interpolation* formula. The linear interpolation of f will be dentoted by f_α and is defined as follows:

$$f_\alpha(x) = f(S_*(x)) + \mu_{Q(S_*(x))}(x).\Delta f(S(x)),$$

where $\Delta f(S(x)) = f(S^*(x)) - f(S_*(x))$.
The number

$$\frac{|f_a(x) - f(x)|}{f(x)}$$

will be called the realative error of the interpolation of f in x. The maximal error of interpolation will be called the error of interpolation of f.

If $f(S_*(x))$ and $f(S^*(x))$ are unknown we can use another interpolation formulas shown below.

1) Lower interpolation

$$f^1_\alpha(x) = P_*(f(S_*(x))) + \mu_{Q(S_*(x))}(x).\Delta_1 f(S(x)), \text{ where}$$

$$\Delta^1 f(S(x)) = P_*(f(S^*(x))) - P_*(f(S_*(x))$$

2) Upper interpolation

$$f^2_\alpha(x) = P^*(f(S_*(x))) + \mu_{Q(S_*(x))}(x).\Delta^2 f(S(x)), \text{ where}$$

$$\Delta^2 f(S(x)) = P^*(f(S^*(x))) - P^*(f(S_*(x))$$

3) Lower cross interpolation

$$f^3_\alpha(x) = P_*(f(S_*(x))) + \mu_{Q(S_*(x))}(x).\Delta^3 f(S(x)), \text{ where}$$

$$\Delta^3 f(S(x)) = P^*(f(S^*(x))) - P_*(f(S_*(x))$$

4) Upper cross interpolation

$$f^4_\alpha(x) = P^*(f(S*(x))) + \mu_{Q(S_*(x))}(x).\Delta^4 f(S(x)), \text{ where}$$

$$\Delta^4 f(S(x)) = P_*(f(S^*(x))) - P^*(f(S_*(x)));$$

The meaning of the above interpolation formlas is obvious.

We will be also interested in the following problem. Given a function $f : X \to Y$ and a number $0 \leq \epsilon \leq 1$. Find categorizations S and P such the error of interpolation of f is less than ϵ.

The following algorithm solves the problem.

$x_0 = z \in X$

$x_{i+1} = Sup\{x > x_i : |f(y) - p_{x_i,x}(y)| \leq \epsilon\}$ for any $y \in (x_i, x]$,

where $p_{x_i,x}(y)$ denotes the straight line determined by points x_i, x.

References

[1] **Czogala, E., Mrozek, A. and Pawlak, Z.(1993).** Fuzzy-rough Controllers. *ICS WUT Reports 32/94.*

[2] **Mrozek, A. and Plonka, L. (1994).** Knowledge Representation in Fuzzy and Rough Controllers. *ICS WUT Report 34/94.*

[3] **Plonka, L. and Mrozek, A. (1994)**. Rule-Based Stabilization of the Inverted Pendulum. *ICS WUT Reports 35/94.*

[4] **Orlowska, E. and Pawlak, Z. (1984)**. Measurement and Indiscernibility. *Bull. PAS, Math. Ser.* Vol. 32, No. 9-10, pp. 617-624.

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

[6] **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.

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

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

[9] **Pawlak, Z. (1994).** Rough Sets, Rough Relations and Rough Functions, *ICS WUT Reparts 24/94.*

[10] **Robinson, A. (1970).** Non-Standard Analysis. NORTH-HOLLNAD PUBLISHING COMPANY.