Calculi of Approximation Spaces in Intelligent Systems

Andrzej Skowron, Jarosław Stepaniuk, and Roman Swiniarski

Abstract. Solving complex real-life problems requires new approximate reasoning methods based on new computing paradigms. One such recently emerging computing paradigm is Rough–Granular Computing (Pedrycz et al. 2008, Stepaniuk 2008) (RGC, in short). The RGC methods have been successfully applied for solving complex problems in areas such as identification of behavioral patterns by autonomous systems, web mining, and sensor fusion. In RGC, an important role play special information granules (Zadeh 1979, Zadeh 2006) called as approximation spaces. These higher order granules are used for approximation of concepts or, in a more general sense, complex granules. We discuss some generalizations of the approximation space definition introduced in 1994 (Skowron and Stepaniuk 1994, Skowron and Stepaniuk 1996, Stepaniuk 2008). The generalizations are motivated by real-life applications of intelligent systems and are related to inductive extensions of approximation spaces.

Keywords: Rough Sets, Approximation Space, Approximation Space Extension, Granular Computing, Complex Granule Approximation, Intelligent Systems.

Andrzej Skowron
Institute of Mathematics, The University of Warsaw, Banacha 2, 02-097 Warsaw, Poland
e-mail: skowron@mimuw.edu.pl

Jarosław Stepaniuk
Department of Computer Science, Białystok University of Technology,
Wiejska 45A, 15-351 Białystok, Poland
e-mail: j.stepaniuk@pb.edu.pl

Roman Swiniarski
Department of Mathematical and Computer Sciences,
5500 Campanile Drive San Diego, CA 92182, USA and
Institute of Computer Science Polish Academy of Sciences,
Ordona 21, 01-237 Warsaw, Poland
e-mail: rswiniar@sciences.sdsu.edu

W. Pedrycz and S.-M. Chen (Eds.): Granular Computing and Intell. Sys., ISRL 13, pp. 35–55.
springerlink.com

1 Introduction

Rough sets, due to Zdzisław Pawlak, can be represented by pairs of sets which give the lower and the upper approximation of the original sets. In the standard version of rough set theory, an approximation space is based on the indiscernibility equivalence relation. Approximation spaces belong to the broad spectrum of basic issues investigated in rough set theory (see, *e.g.*, (Bazan et al. 2006, Greco et al. 2008, Jankowski and Skowron 2008, Skowron and Stepaniuk 1994, Skowron and Stepaniuk 1996, Słowiński and Vanderpooten 2000, Skowron et al. 2006, Stepaniuk 2008, Zhu 2009). Over the years different aspects of approximation spaces were investigated and many generalizations of the approach based on indiscernibility equivalence relation (Pawlak and Skowron 2007) were proposed. In this chapter, we discuss some aspects of generalizations of approximation spaces investigated in (Skowron and Stepaniuk 1994, Skowron and Stepaniuk 1996, Stepaniuk 2008) that are important for real-life applications, e.g., in searching for approximation of complex concepts (see, e.g., (Bazan et al. 2006, Bazan 2008). Rough set based strategies for extension of such approximation spaces from samples of objects onto their extensions are discussed. The extensions of approximation spaces can be treated as operations for inductive reasoning. The investigated approach enables us to present the uniform foundations for inducing approximations of different kinds of higher order granules such as concepts, classifications, or functions. In particular, we emphasize the fundamental role of approximation spaces for inducing diverse kinds of classifiers used in machine learning or data mining. The searching problem for relevant approximation spaces and their extensions is of high computational complexity. Hence, efficient heuristics should be used in searching for approximate solutions of this problem. Moreover, in hierarchical learning of complex concepts many different approximation spaces should be discovered. Learning of such concepts can be supported by domain knowledge and ontology approximation (see, e.g., (Bazan 2008, Bazan et al. 2006)).

The chapter is organized as follows. In Section 2 we discuss basic notions for our approach. In Section 3 we present a generalization of the approximation space definition from (Skowron and Stepaniuk 1994, Skowron and Stepaniuk 1996, Stepaniuk 2008). In particular, in Subsection 3.2 we present new rough set approach to function approximation.

2 Basic Notions

2.1 Attributes, Signatures of Objects and Two Semantics

In (Pawlak and Skowron 2007) any attribute a is defined as a function from the universe of objects U into the set of attribute values V_a. However, in applications we expect that the value of attribute should be also defined for objects from extensions of U, *i.e.*, for new objects which can be perceived in the future[1]. The universe U

[1] Objects from U are treated as labels of real perceived objects.

is only a sample of possible objects. This requires some modification of the basic definitions of attribute and signature of objects (Pawlak 1991, Pawlak and Skowron 2007).

One can give interpretation of attributes using a concept of interaction. In this case, we treat attributes as granules and we consider their interactions with environments. If a is a given attribute and e denotes a state of environment then the result of interaction between a and e is equal to a pair (l_e, v), where l_e is a label of e and $v \in V_a$. Analogously, if $IS = (U, A)$ is a given information system and e denotes a state of environment then by interaction of IS and e we obtain the information system $IS' = (U \cup \{l_e\}, A^*)$, where $A^* = \{a^* : a \in A\}$ and $a^*(u) = a(u)$ for $u \in U$ and $a^*(l_e) = v$ for some $v \in V_a$. Hence, information systems are dynamic objects created through interaction of already existing information systems with environments. Note that the initial information system can be empty, i.e. the set of objects of this information system is empty. Moreover, let us observe that elements of U are labels of the environment states rather than states.

One can represent any attribute by a family of formulas and interpret the attribute as the result of interaction of this set with the environment. In this case, we assume that for any attribute a under consideration there is given a relational structure R_a. Together with the simple structure $(V_a, =)$ (Pawlak and Skowron 2007), some other relational structures R_a with the carrier V_a for $a \in A$ and a signature τ are considered. We also assume that with any attribute a is identified a set of some generic formulas $\{\alpha_i\}_{i \in J}$ (where J is a set of indexes) interpreted over R_a as a subsets of V_a, i.e., $\|\alpha_i\|_{R_a} = \{v \in V_a : R_a, v \models \alpha_i\}$. Moreover, it is assumed that the set $\{\|\alpha_i\|_{R_a}\}_{i \in J}$ is a partition of V_a. Perception of an object u by a given attribute a is represented by selection of a formula α_i and a value $v \in V_a$ such that $v \in \|\alpha_i\|_{R_a}$. Using an intuitive interpretation one can say that such a pair (α_i, v) is selected from $\{\alpha_i\}_{i \in J}$ and V_a, respectively, as the result of sensory measurement. We assume that for a given set of attributes A and any object u the signature of u relative to A is given by $Inf_A(u) = \{(a, \alpha_u^a, v) : a \in A\}$, where (α_u^a, v) is the result of sensory measurement by a on u.

Let us observe that a triple (a, α_u^a, v) can be encoded by the atomic formula $a = v$ with interpretation

$$\|a = v\|_{U^*} = \{u \in U^* : (a, \alpha_u^a, v) \in Inf_a(u) \text{ for some } \alpha_u^a\}. \tag{1}$$

We also write (a, v) instead of (a, α_u^a, v) if this not lead to confusion.

One can also consider a soft version of the attribute definition. In this case, we assume that the semantics of the family $\{\alpha_i\}_{i \in J}$ is given by fuzzy membership functions for α_i and the set of these functions define a fuzzy partition (Klir 2007).

We construct granular formulas from atomic formulas corresponding to the considered attributes. In the consequence, the satisfiability of such formulas is defined if the satisfiability of atomic formulas is given as the result of sensor measurement. Hence, one can consider for any constructed formula α over atomic formulas its semantics $\|\alpha\|_U \subseteq U$ over U as well as the semantics $\|\alpha\|_U^* \subseteq U^*$ over U^*, where $U \subseteq U^*$ (see Figure 1). The difference between these two cases is the following. In the case of U, one can compute $\|\alpha\|_U \subseteq U$ but in the case $\|\alpha\|_{U^*} \subseteq U^*$, for objects from $U^* \setminus U$, there is no information about their membership relative to $\|\alpha\|_{U^* \setminus \|\alpha\|_U}$. One can estimate the satisfiability of α for objects $u \in U^* \setminus U$ only after the relevant sensory measurements on u are performed. In particular, one can use some methods for estimation of relationships among semantics of formulas over U^* using the relationships among semantics of these formulas over U. For example, one can apply statistical methods. This step is crucial in investigation of extensions of approximation spaces relevant for inducing classifiers from data (see, e.g., (Bazan et al 2006, Skowron et al. 2006, Pedrycz et al. 2008).

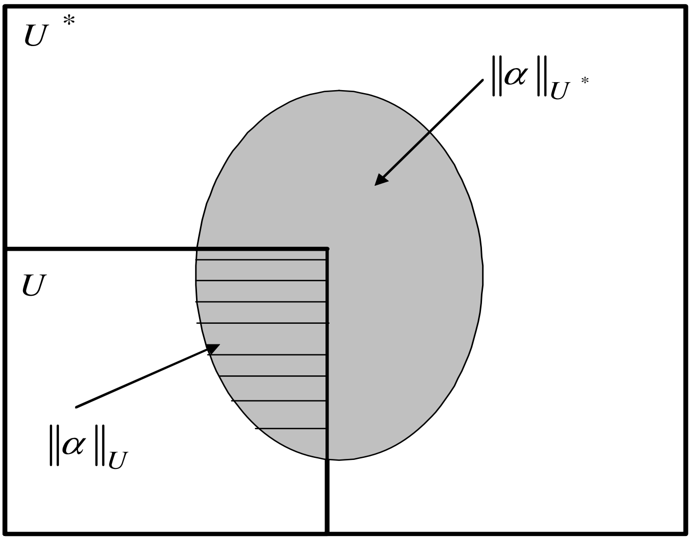

Fig. 1 Two semantics of α over U and U^*, respectively

2.2 Uncertainty Function

In (Skowron and Stepaniuk 1994, Skowron and Stepaniuk 1996, Stepaniuk 2008) the uncertainty function I defines for every object x from a given sample U of objects, a set of objects described similarly to x. The set $I(x)$ is called the neighborhood of x.

Let $P_\omega(U^*) = \bigcup_{i\geq 1} P^i(U^*)$, where $P^1(U^*) = P(U^*)$ and $P^{i+1}(U^*) = P(P^i(U^*)))$ for $i \geq 1$. For example, if $card(U^*) = 2$ and $U^* = \{x_1, x_2\}$, then we obtain $P^1(U^*) = \{\emptyset, \{x_1\}, \{x_2\}, \{x_1, x_2\}\}$, $P^2(U^*) = \{\emptyset, \{\emptyset\}, \{\{x_1\}\}, \{\{x_2\}\}, \{\{x_1, x_2\}\}, \{\emptyset, \{x_1\}\}, \{\emptyset, \{x_2\}\}, \{\emptyset, \{x_1, x_2\}\}, \ldots\}, \ldots$. If $card(U^*) = n$, where n is a positive natural number, then $card(P^1(U^*)) = 2^n$ and $card(P^{n+1}(U^*)) = 2^{card(P^n(U^*))}$, for $n \geq 1$. For example, $card(P^3(U^*)) = 2^{2^{2^n}}$.

In this chapter, we consider uncertainty functions of the form $I : U^* \longrightarrow P_\omega(U^*)$. The values of uncertainty functions are called granular neighborhoods. These granular neighborhoods are defined by the so called granular formulas. The values of such uncertainty functions are not necessarily from $P(U^*)$ but from $P_\omega(U^*)$. In the following sections, we will present more details on granular neighborhoods and granular formulas. Figure 2 presents an illustrative example of the uncertainty function with values in $P^2(U^*)$ rather than in $P(U^*)$. The discussed here generalization of neighborhoods are also motivated by the necessity of modeling or discovery of complex structural objects in solving problems of pattern recognition, machine learning, or data mining. These structural objects (granules) can be defined as sets on higher levels of the powerset hierarchy. Among examples of such granules are indiscernibility or similarity classes of patterns or relational structures discovered in images, clusters of time widows, indiscernibility or similarity classes of sequences of time windows representing processes, behavioral graphs (for more details see, e.g., (Skowron and Szczuka 2010, Bazan 2008)).

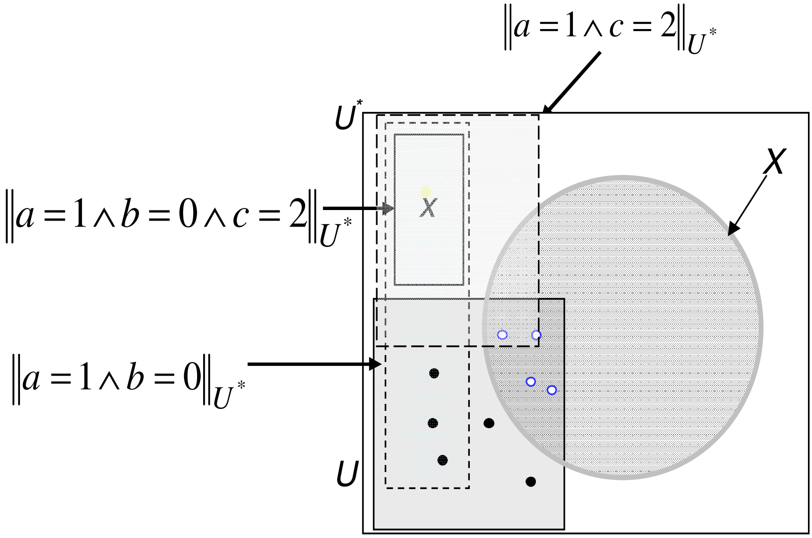

Fig. 2 Uncertainty function $I : U^* \to P^2(U^*)$. The neighborhood of $x \in U^* \setminus U$, where $Inf_A(x) = \{(a,1),(b,0),(c,2)\}$, does not contain training cases from U. The generalizations of this neighborhood described by formulas $\|a = 1 \wedge c = 2\|_{U^*}$ and $\|a = 1 \wedge b = 2\|_{U^*}$ have non empty intersections with U.

If $X \in P_\omega(U^*)$ and $U \subseteq U^*$ then by $X \upharpoonright U$ we denote the set defined as follows (i) if $X \in P(U^*)$ then $X \upharpoonright U = X \cap U$ and (ii) for any $i \geq 1$ if $X \in P^{i+1}(U^*)$ then $X \upharpoonright U = \{Y \upharpoonright U : Y \in X\}$. For example, if $U = \{x_1\}$, $U^* = \{x_1, x_2\}$ and $X = \{\{x_2\}, \{x_1, x_2\}\}$ $(X \in P^2(U^*))$, then $X \upharpoonright U = \{Y \upharpoonright U : Y \in X\} = \{Y \cap U : Y \in X\} = \{\emptyset, \{x_1\}\}$.

2.3 Rough Inclusion Function

The second component of any approximation space is the rough inclusion function (Skowron and Stepaniuk 1996, Stepaniuk 2008).

One can consider general constraints which the rough inclusion functions should satisfy. In this section, we present only some examples of rough inclusion functions.

The rough inclusion function $\nu : P(U) \times P(U) \to [0,1]$ defines the degree of inclusion of X in Y, where $X,Y \subseteq U$.[2]

In the simplest case the standard rough inclusion function can be defined by (see, e.g., (Skowron and Stepaniuk 1996, Pawlak and Skowron 2007)):

$$\nu_{SRI}(X,Y) = \begin{cases} \frac{card(X \cap Y)}{card(X)} & \text{if } X \neq \emptyset \\ 1 & \text{if } X = \emptyset. \end{cases} \tag{2}$$

Some illustrative example is given in Table 1.

Table 1 Illustration of Standard Rough Inclusion Function

X	Y	$\nu_{SRI}(X,Y)$
$\{x_1,x_3,x_7,x_8\}$	$\{x_2,x_4,x_5,x_6,x_9\}$	0
$\{x_1,x_3,x_7,x_8\}$	$\{x_1,x_2,x_4,x_5,x_6,x_9\}$	0.25
$\{x_1,x_3,x_7,x_8\}$	$\{x_1,x_2,x_3,x_7,x_8\}$	1

It is important to note that an inclusion measure expressed in terms of the confidence measure, widely used in data mining, was considered by Łukasiewicz (Łukasiewicz 1913) long time ago in studies on assigning fractional truth values to logical formulas.

The rough inclusion function was generalized in rough mereology (Polkowski and Skowron 1996). For definition of inclusion function for more general granules, e.g., partitions of objects one can use measure based on positive region (Pawlak and Skowron 2007), entropy (Hastie et al. 2008) or rough entropy (Pal et al. 2005, Małyszko and Stepaniuk 2010). Inclusion measures for more general granules were also investigated (Skowron 2001, Bianucci and Cattaneo 2009). However, more work in this direction should be done, especially on inclusion of granules with complex structures, in particular for granular neighborhoods.

3 Approximation Spaces

In this section, we present a generalization of the approximation space definition from (Skowron and Stepaniuk 1994, Skowron and Stepaniuk 1996, Stepaniuk 2008).

In applications, approximation spaces are constructed for a given concept or a family of concepts creating a partition (in the case of classification) rather than for the class of all concepts. Then the searching for components of approximation space relevant for the concept approximation becomes feasible. The concept is given only

[2] We assume that U is a finite sample of objects.

an a sample U of objects. We often restrict the definition of components of approximation space to objects from U^* and/or some patterns from $P_\omega(U^*)$ or from $P_\omega(U^\bullet)$, where $U^* \subseteq U^\bullet$ necessary for approximation of a given concept X only. The definitions of the uncertainty function and the inclusion function can be restricted to some subsets of U^* and $P_\omega(U^*)$ (or $U^* \subseteq U^\bullet$), respectively, which are relevant for approximated concept(s).

Moreover, the optimization requirements for the lower approximation and upper approximation are then also restricted to the given concept X. These requirements are expressing closeness of the induced concept approximations to the approximation of the concept(s) on the given sample of objects and are combined with the description length of constructed approximations for obtaining the relevant quality measures. Usually, the uncertainty functions and the rough inclusion functions are parameterized. Then searching (in the family of the these functions defined by possible values of parameters) for the (semi)optimal uncertainty function and the rough inclusion function relative to the selected quality measure becomes feasible.

Definition 1. An approximation space over a set of attributes A for a concept $X \subseteq U^*$ given on a sample $U \subseteq U^*$ of objects is a system

$$AS = (U, U^*, I, \nu, L) \tag{3}$$

where

- U is a sample of objects with known signatures relative to a given set of attributes A,
- L is a language of granular formulas defined over atomic formulas corresponding to generic formulas from signatures (see Section 2.2),
- the set U^* is such that for any object $u \in U^*$ the signature $Inf_A(u)$ of u relative to A can be obtained as the result of sensory measurements on u,
- $I : U^* \to P_\omega(U^\bullet)$ is an uncertainty function, where $U^* \subseteq U^\bullet$; we assume that the granular neighborhood $I(u)$ is computable from $Inf_A(u)$, i.e., from $Inf_A(u)$ it is possible to compute a formula $\alpha_{Inf_A(u)} \in L$ such that $I(x) = \|\alpha_{Inf_A(u)}\|_{U^*}$,
- $\nu : P_\omega(U^\bullet) \times P_\omega(U^\bullet) \to [0,1]$ is a partial rough inclusion function, such that form any $x \in U^*$ the value $\nu(I(x), X)$ is defined for the considered concept X.

In Section 3.2, we consider an uncertainty function with values in $P(P(U^*) \times P(R_+))$, where R_+ is the set of reals. Hence, we assume that the values of the uncertainty function I may belong to the space of possible patterns from $P_\omega(U^\bullet)$, where $U^* \subseteq U^\bullet = U^* \cup R_+$.

The partiality of the rough inclusion makes it possible to define the values of this functions on relevant patterns for approximation only.

We assume that the lower approximation operation $LOW(AS, X)$ and the upper approximation operation $UPP(AS, X)$ of the concept X in the approximation space AS satisfy the following condition:

$$\nu(LOW(AS, X), UPP(AS, X)) = 1. \tag{4}$$

Usually the uncertainty function and the rough inclusion function are parameterized. In this parameterized family of approximation spaces, one can search for an approximation space enabling us to approximate the concept X restricted to a given sample U with the satisfactory quality. The quality of approximation can be expressed by some quality measures. For example, one can use the following criterion:

1. $LOW(AS,X) \upharpoonright U$ is included in $X \upharpoonright U$ to a degree at least deg, i.e., $\nu(LOW(AS,X) \upharpoonright U, X \upharpoonright U)) \geq deg$,
2. $X \upharpoonright U$ is included in $UPP(AS,X) \upharpoonright U$ to a degree at least deg, this means that, $\nu(UPP(AS,X \upharpoonright U), X \upharpoonright U) \geq deg$,

where deg is a given threshold from the interval $[0,1]$.

The above condition expresses the degree to which at least the induced approximations in AS are close to the concept X on the sample U. One can also introduce the description length of the induced approximations. A combination of these two measures can be used as the quality measure for the induced approximation space. Then the searching problem for the relevant approximation space can be considered as the optimization problem relative to this quality measure. This approach may be interpreted as a form of the minimal description length principle (Rissanen 1985). The result of optimization can be checked against a testing sample. This enables us to estimate the quality of approximation. Note that further optimization can be performed relative to parameters of the selected quality measure.

3.1 Approximations and Decision Rules

In this section, we discuss generation of approximations on extensions of samples of objects.

In the example we illustrate how the approximations of sets (concepts) can be estimated using only partial information about these sets. Moreover, the example introduces uncertainty functions with values in $P^2(U)$ and rough inclusion functions defined for sets from $P^2(U)$.

Let us assume that $DT = (U, A \cup \{d\})$ is a decision table, where $U = \{x_1, \ldots, x_9\}$ is a set of objects and $A = \{a,b,c\}$ is a set of condition attributes (see Table 2).

In DT we compute two decision reducts: $\{a,b\}$ and $\{b,c\}$. We obtain the set $Rule\_set = \{r_1, \ldots, r_{12}\}$ of minimal (reduct based) decision rules (Pawlak and Skowron 2007).

From x_1 we obtain two rules:
r_1 : **if** $a = 1$ **and** $b = 1$ **then** $d = 1$, r_2 : **if** $b = 1$ **and** $c = 0$ **then** $d = 1$.
From x_2 and x_4 we obtain two rules:
r_3 : **if** $a = 0$ **and** $b = 2$ **then** $d = 1$, r_4 : **if** $b = 2$ **and** $c = 0$ **then** $d = 1$.
From x_5 we obtain one new rule:
r_5 : **if** $a = 0$ **and** $b = 1$ **then** $d = 1$.
From x_3 we obtain two rules:
r_6 : **if** $a = 1$ **and** $b = 0$ **then** $d = 0$, r_7 : **if** $b = 0$ **and** $c = 1$ **then** $d = 0$.
From x_6 we obtain two rules:
r_8 : **if** $a = 0$ **and** $b = 0$ **then** $d = 0$, r_9 : **if** $b = 0$ **and** $c = 0$ **then** $d = 0$.

Table 2 Decision table over the set of objects U

	a	b	c	d
x_1	1	1	0	1
x_2	0	2	0	1
x_3	1	0	1	0
x_4	0	2	0	1
x_5	0	1	0	1
x_6	0	0	0	0
x_7	1	0	2	0
x_8	1	2	1	0
x_9	0	0	1	0

From x_7 we obtain one new rule:
r_{10} : **if** $b = 0$ **and** $c = 2$ **then** $d = 0$.
From x_6 we obtain two rules:
r_{11} : **if** $a = 1$ **and** $b = 2$ **then** $d = 0$, r_{12} : **if** $b = 2$ **and** $c = 1$ **then** $d = 0$.
Let $U^* = U \cup \{x_{10}, x_{11}, x_{12}, x_{13}, x_{14}\}$ (see Table 3).

Table 3 Decision table over the set of objects $U^* - U$

	a	b	c	d	d_{class}
x_{10}	0	2	1	1	1 from r_3 or 0 from r_{12}
x_{11}	1	2	0	0	1 from r_4 or 0 from r_{11}
x_{12}	1	2	0	0	1 from r_4 or 0 from r_{11}
x_{13}	0	1	2	1	1 from r_5
x_{14}	1	1	2	1	1 from r_1

Let $h : [0,1] \to \{0, 1/2, 1\}$ be a function defined by

$$h(t) = \begin{cases} 1 & \text{if } t > 1/2 \\ 1/2 & \text{if } t = 1/2 \\ 0 & \text{if } t < 1/2. \end{cases} \quad (5)$$

Below we present an example of the uncertainty and rough inclusion functions:

$$I(x) = \{\|lh(r)\|_{U^*} : x \in \|lh(r)\|_{U^*} \text{ and } r \in Rule\_set\}, \quad (6)$$

where $x \in U^*$ and $lh(r)$ denotes the formula on the left hand side of the rule r, and

$$\nu_U(X, Z) = \begin{cases} h\left(\frac{card(\{Y \in X : Y \cap U \subseteq Z\})}{card(\{Y \in X : Y \cap U \subseteq Z\}) + card(\{Y \in X : Y \cap U \subseteq U^* \setminus Z\})}\right) & \text{if } X \neq \emptyset \\ 0 & \text{if } X = \emptyset, \end{cases} \quad (7)$$

where $X \subseteq P(U^*)$ and $Z \subseteq U^*$.

The uncertainty and rough inclusion functions can now be used to define the lower approximation $LOW(AS^*,Z)$, the upper approximation $UPP(AS^*,Z)$, and the boundary region $BN(AS^*,Z)$ of $Z \subseteq P(U^*)$ by:

$$LOW(AS^*,Z) = \{x \in U^* : \nu_U(I(x),Z) = 1\}, \tag{8}$$

and

$$UPP(AS^*,Z) = \{x \in U^* : \nu_U(I(x),Z) > 0\}, \tag{9}$$

$$BN(AS^*,Z) = UPP(AS^*,Z) \setminus LOW(AS^*,Z). \tag{10}$$

In the example, we classify objects from U^* to the lower approximation of Z if majority of rules matching this object are voting for Z and to the upper approximation of Z if at least half of the rules matching x are voting for Z. Certainly, one can follow many other voting schemes developed in machine learning or by introducing less crisp conditions in the boundary region definition. The defined approximations can be treated as estimations of the exact approximations of subsets of U^* because they are induced on the basis of samples of such sets restricted to U only. One can use some standard quality measures developed in machine learning to calculate the quality of such approximations assuming that after estimation of approximations on U^* full information about membership for objects relative to the approximated subsets of U^* is uncovered analogously to the testing sets in machine learning.

Let $C_1^* = \{x \in U^* : d(x) = 1\} = \{x_1, x_2, x_4, x_5, x_{10}, x_{13}, x_{14}\}$. We obtain the set $U^* \setminus C_1^* = C_0^* = \{x_3, x_6, x_7, x_8, x_9, x_{11}, x_{12}\}$. The uncertainty function and rough inclusion are presented in Table 4.

Table 4 Uncertainty function and rough inclusion over the set of objects U^*

	$I(\cdot)$	$\nu_U(I(\cdot),C_1^*)$
x_1	$\{\{x_1,x_{14}\},\{x_1,x_5\}\}$	$h(2/2)=1$
x_2	$\{\{x_2,x_4,x_{10}\},\{x_2,x_4,x_{11},x_{12}\}\}$	$h(2/2)=1$
x_3	$\{\{x_3,x_7\},\{x_3,x_9\}\}$	$h(0/2)=0$
x_4	$\{\{x_2,x_4,x_{10}\},\{x_2,x_4,x_{11},x_{12}\}\}$	$h(2/2)=1$
x_5	$\{\{x_5,x_{13}\},\{x_1,x_5\}\}$	$h(2/2)=1$
x_6	$\{\{x_6,x_9\},\{x_6\}\}$	$h(0/2)=0$
x_7	$\{\{x_3,x_7\},\{x_7\}\}$	$h(0/2)=0$
x_8	$\{\{x_8,x_{11},x_{12}\},\{x_8,x_{10}\}\}$	$h(0/2)=0$
x_9	$\{\{x_6,x_9\},\{x_3,x_9\}\}$	$h(0/2)=0$
x_{10}	$\{\{x_2,x_4,x_{10}\},\{x_8,x_{10}\}\}$	$h(1/2)=1/2$
x_{11}	$\{\{x_8,x_{11},x_{12}\},\{x_2,x_4,x_{11},x_{12}\}\}$	$h(1/2)=1/2$
x_{12}	$\{\{x_8,x_{11},x_{12}\},\{x_2,x_4,x_{11},x_{12}\}\}$	$h(1/2)=1/2$
x_{13}	$\{\{x_5,x_{13}\}\}$	$h(1/1)=1$
x_{14}	$\{\{x_1,x_{14}\}\}$	$h(1/1)=1$

Thus, in our example from Table 4 we obtain

$$LOW(AS^*, C_1^*) = \{x \in U^* : \nu_U(I(x), C_1^*) = 1\} = \{x_1, x_2, x_4, x_5, x_{13}, x_{14}\}, \quad (11)$$

$$UPP(AS^*, C_1^*) = \{x \in U^* : \nu_U(I(x), C_1^*) > 0\} =$$

$$\{x_1, x_2, x_4, x_5, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}\}, \quad (12)$$

$$BN(AS^*, C_1^*) = UPP(AS^*, C_1^*) \setminus LOW(AS^*, C_1^*) = \{x_{10}, x_{11}, x_{12}\}. \quad (13)$$

3.2 Function Approximations

In this subsection, we discuss the rough set approach to function approximation from available incomplete data. Our approach can be treated as a kind of rough clustering of functional data (Ramsay 2002).

Let us consider an example of function approximation. We assume that a partial information is only available about a function. This means that, some points from the graph of the function are known.

Before presenting a more formal description of function approximation we introduce some notation.

A function $f : U \longrightarrow R_+$ will be called a sample of a function $f^* : U^* \rightarrow R_+$, where R_+ is the set of non-negative reals and $U \subseteq U^*$ is a finite subset of U^*, if f^* is an extension of f.

By Gf (Gf^*) we denote the graph of f (f^*), respectively, i.e., the set $\{(x, f(x)) : x \in U\}$ ($\{(x, f^*(x)) : x \in U^*\}$). For any $Z \subseteq U^* \times R_+$ by $\pi_1(Z)$ and $\pi_2(Z)$ we denote the set $\{x \in U^* : \exists y \in R_+ \ (x,y) \in Z\}$ and $\{y \in R_+ : \exists x \in U^* \ (x,y) \in Z\}$, respectively.

First we define approximations of Gf given on a sample U of objects and next we show how to induce approximations of Gf^* over U^*, i.e., on extension of U.

Let Δ will be a partition of $f(U)$ into sets of reals of diameter less than $\delta > 0$, where δ is a given threshold. We also assume that $IS = (U, A)$ is a given information system. Let us also assume that for any object signature $Inf_A(x) = \{(a, a(x)) : a \in A\}$ (Pawlak and Skowron 2007) there is assigned an interval of non-negative reals with diameter less than δ. We denote this interval by $\Delta_{Inf_A(x)}$. Hence, $\Delta = \{\Delta_{Inf_A(x)} : x \in U\}$.

We consider an approximation space $AS_{IS,\Delta} = (U, I, \nu^*)$ (relative to given IS and Δ), where

$$I(x) = [x]_{IND(A)} \times \Delta_{Inf_A(x)}, \quad (14)$$

and

$$\nu^*(X,Y) = \begin{cases} \frac{card(\pi_1(X \cap Y))}{card(\pi_1(X))} & \text{if } X \neq \emptyset \\ 1 & \text{if } X = \emptyset, \end{cases} \quad (15)$$

for $X, Y \subseteq U \times R_+$.

The lower approximation and upper approximation of Gf in AS are defined by

$$LOW(AS_{IS,\Delta}, Gf) = \bigcup\{I(x) : \nu^*(I(x), Gf) = 1\}, \quad (16)$$

and

$$UPP(AS_{IS,\Delta}, Gf) = \bigcup\{I(x) : \ \nu^*(I(x), Gf) > 0\}, \tag{17}$$

respectively.

Observe that this definition is different from the standard definition of the lower approximation (Pawlak and Skowron 2007). The defined approximation space is a bit more general than in (Skowron and Stepaniuk 1996), e.g., the values of the uncertainty functions are subsets of $U \times R_+$ instead of U. Moreover, one can easily see that by applying the standard definition of relation approximation to f (Pawlak and Skowron 2007) (this is a special case of relation) the lower approximation of function is almost always equal to the empty set. The new definition is making it possible to express better the fact that a given neighborhood is "well" matching the graph of f (Skowron et al. 2006, Stepaniuk 2008). For expressing this a classical set theoretical inclusion of neighborhood into the graph of f is not satisfactory.

Example 1. We present the first illustrative example of a function approximation. Let $f : U \to R_+$ where $U = \{1,2,3,4,5,6\}$. Let $f(1) = 3,\ f(2) = 2,\ f(3) = 2,$ $f(4) = 5,\ f(5) = 5,\ f(6) = 2.$

Let $IS = (U, A)$ be an information system where $A = \{a\}$ and

$$a(x) = \begin{cases} 0 \text{ if } 0 \le x \le 2, \\ 1 \text{ if } 2 < x \le 4, \\ 2 \text{ if } 4 < x \le 6. \end{cases} \tag{18}$$

Thus the partition $U/IND(A) = \{\{1,2\},\{3,4\},\{5,6\}\}$. The graph of f is defined by $Gf = \{(x, f(x)) : x \in U\} = \{(1,3),(2,2),(3,2),(4,5),(5,5),(6,2)\}$.

We define approximations of Gf given on the sample U of objects.

We obtain $f(U) = \{2,3,5\}$ and let $\Delta = \{\{2,3\},\{5\}\}$ will be a partition of $f(U)$.

We consider an approximation space $AS_{IS,\Delta} = (U, I, \nu^*)$ (relative to given IS and Δ), where

$$I(x) = [x]_{IND(A)} \times \Delta_{Inf_A(x)}, \tag{19}$$

is defined by

$$I(x) = \begin{cases} \{1,2\} \times [1.5,4] & \text{if } x \in \{1,2\}, \\ \{3,4\} \times [1.7,4.5] & \text{if } x \in \{3,4\}, \\ \{5,6\} \times [3,4] & \text{if } x \in \{5,6\}. \end{cases} \tag{20}$$

We obtain the lower approximation and upper approximation of Gf in the approximation space $AS_{IS,\Delta}$:

$$\begin{gathered} LOW(AS_{IS,\Delta}, Gf) = \bigcup\{I(x) : \ \nu^*(I(x), Gf) = 1\} = \\ I(1) \cup I(2) = \{1,2\} \times [1.5,\ 4), \end{gathered} \tag{21}$$

and

$$\begin{gathered} UPP(AS_{IS,\Delta}, Gf) = \bigcup\{I(x) : \ \nu^*(I(x), Gf) > 0\} = \\ = I(1) \cup I(2) \cup I(3) \cup I(4) = \{1,2\} \times [1.5,\ 4) \cup \{4,5\} \times [1.7,\ 4.5), \end{gathered} \tag{22}$$

respectively.

Example 2. We present the second illustrative example of a function approximation. First, let us recall that an interval is a set of real numbers with the property that any number that lies between two numbers in the set is also included in the set. The closed interval of numbers between v and w ($v, w \in R_+$), including v and w, will be denoted by $[v, w]$.

Let us consider a function $f : R_+ \to R_+$. We have only a partial information about this function given by $G_i = \delta(x_i) \times \varepsilon(f(x_i))$, where $\delta(x_i)$ denotes a closed interval of reals to which belongs x_i, $\varepsilon(f(x_i))$ denotes a closed interval of reals to which belongs $f(x_i)$ and $i = 1, \ldots, n$. A family $\{G_1, \ldots, G_n\}$ is called a partial information about graph $Gf = \{(x, f(x)) : x \in R_+\}$.

Let Nh denotes a family of elements of $P(R_+) \times P(R_+)$ called neighborhoods.

In our example, we consider $Nh = \{X_1, X_2, X_3\}$, where $X_1 = [1,6] \times [0.1, 0.4]$, $X_2 = [7,12] \times [1.1, 1.4]$ and $X_3 = [13,18] \times [2.1, 2.4]$.

Let us recall the inclusion definition between closed intervals:

$$[v_1, w_1] \subseteq [v_2, w_2] \text{ iff } v_2 \leq v_1 \ \& \ w_1 \leq w_2 \tag{23}$$

We define the new rough inclusion function by

$$\nu(X, \{G_1, \ldots, G_n\}) = \begin{cases} 1 \quad \text{if} \quad \forall_{i \in \{1,\ldots,n\}} (\pi_1(G_i) \cap \pi_1(X) \neq \emptyset \to G_i \subseteq X) \\ 1/2 \text{ if } \exists_{i \in \{1,\ldots,n\}} (\pi_1(G_i) \cap \pi_1(X) \neq \emptyset \ \& \ G_i \cap X \neq \emptyset). \\ 0 \quad \text{if } \forall_{i \in \{1,\ldots,n\}} (\pi_1(G_i) \cap \pi_1(X) \neq \emptyset \to G_i \cap X = \emptyset) \end{cases} \tag{24}$$

Let sample values of a function $f : R_+ \to R_+$ are given in Table 5.

Let an approximation space $AS = (R_+, Nh, \nu)$ be given. We define the lower and upper approximation as follows:

$$LOW(AS, \{G_1, \ldots, G_n\}) = \bigcup \{X \in Nh : \ \nu(X, \{G_1, \ldots, G_n\}) = 1\}, \tag{25}$$

Table 5 Sample values of a function f, $\delta(x_i)$, $\varepsilon(f(x_i))$ and G_i

i	x_i	$f(x_i)$	$\delta(x_i)$	$\varepsilon(f(x_i))$	G_i	$\pi_1(G_i)$
1	1.5	0.55	$[1,2]$	$[0.5,0.6]$	$[1,2] \times [0.5,0.6]$	$[1,2]$
2	3.5	0.65	$[3,4]$	$[0.6,0.6]$	$[3,4] \times [0.6,0.6]$	$[3,4]$
3	5.5	0.025	$[5,6]$	$[0,0.5]$	$[5,6] \times [0,0.5]$	$[5,6]$
4	7.5	1.35	$[7,8]$	$[1.3,1.4]$	$[7,8] \times [1.3,1.4]$	$[7,8]$
5	9.5	1.15	$[9,10]$	$[1.1,1.2]$	$[9,10] \times [1.1,1.2]$	$[9,10]$
6	11.5	1.35	$[11,12]$	$[1.3,1.4]$	$[11,12] \times [1.3,1.4]$	$[11,12]$
7	13.5	2.25	$[13,14]$	$[2.2,2.3]$	$[13,14] \times [2.2,2.3]$	$[13,14]$
8	15.5	2.025	$[15,16]$	$[2,2.05]$	$[15,16] \times [2,2.05]$	$[15,16]$
9	17.5	2.475	$[17,18]$	$[2.45,2.5]$	$[17,18] \times [2.45,2.5]$	$[17,18]$

$$UPP(AS,\{G_1,\dots,G_n\}) = \bigcup\{X \in Nh: \ \nu(X,\{G_1,\dots,G_n\}) > 0\}. \quad (26)$$

In our example, we obtain

$$LOW(AS,\{G_1,\dots,G_9\}) = X_2 = [7,12] \times [1.1,1.4], \quad (27)$$

$$UPP(AS,\{G_1,\dots,G_9\}) = X_2 \cup X_3 = [7,12] \times [1.1,1.4] \cup [13,18] \times [2.1,2.4]. \quad (28)$$

The above defined approximations are approximations over the set of objects from sample $U \subseteq U^*$. Now, we present an approach for inducing of approximations of the graph Gf^* of function f^* on U^*, i.e., on extension of U. We use an illustrative example to present the approach.

It is worthwhile mentioning that by using boolean reasoning (Pawlak and Skowron 2007) one can generate patterns described by conjunctions of descriptors over IS such that the deviation of f on such patterns in U is less than a given threshold δ. This means that for any such a formula α the set $f(\|\alpha\|_U)$ has diameter less than a given threshold δ, i.e., the image of $\|\alpha\|_U$, i.e., the set $f(\|\alpha\|_U)$, is included into $[y-\delta/2, y+\delta/2)$ for some $y \in U$. Moreover, one can generate such minimal patterns, i.e., formulas α having the above property but no formula obtained by drooping some descriptors from α has that property (Pawlak and Skowron 2007, Bazan et al. 2002). By $PATTERN(A, f, \delta)$ we denote a set of induced patterns with the above properties. One can also assume[3] that $PATTERN(A, f, \delta)$ is extended by adding some shortenings of minimal patterns. For any pattern from $PATTERN(A, f, \delta)$ it is assigned an interval of reals Δ_α such that the deviation of f on $\|\alpha\|_U$ is in Δ_α, i.e., $f(\|\alpha\|_U) \subseteq \Delta_\alpha$.

Note that, for any boolean combination α of descriptors over A, it is also well defined its semantics $\|\alpha\|_{U^*}$ over U^*. However, there is only available information about a part of $\|\alpha\|_{U^*}$ equal to $\|\alpha\|_U = \|\alpha\|_{U^*} \cap U$. Assuming that the patterns from $PATTERN(A, f, \delta)$ are strong (i.e., their support is sufficiently large) one may induce that the following inclusion holds:

$$f^*(\|\alpha\|_{U^*}) \subseteq [y-\delta/2, y+\delta/2). \quad (29)$$

We can now define a generalized approximation space making it possible to extend the approximation of $Gf = \{(x, f(x) : \ x \in U\}$ over the defined previously approximation space AS to approximation of $Gf^* = \{(x, f(x) : \ x \in U^*\}$, where $U \subseteq U^*$.

Let us consider a generalized approximation space

$$AS^* = (U, U^*, I^*, \nu^*_{tr}, L^*), \quad (30)$$

where

- tr is a given threshold from the interval [0, 0.5),
- L^* is a language of boolean combinations of descriptors over the information system IS (Pawlak and Skowron 2007) used for construction of patterns from the set $PATTERN(A, f, \delta)$,

[3] Analogously to shortening of decision rules (Pawlak and Skowron 2007).

- $I^*(x) = \{\|\alpha\|_{U^*} \times \Delta_\alpha : \alpha \in PATTERN(A, f, \delta)\ \&\ x \in \|\alpha\|)_{U^*}\}$ for $x \in U^*$, where U^* is an extension of the sample U, i.e., $U \subseteq U^*$,
- for any finite family $\mathscr{X} \subseteq P(U^*) \times \mathscr{I}$, where $P(U^*)$ is the powerset of U^*, $\mathscr{I}$ is a family of intervals of reals of diameter less than δ and for any Y from $U^* \times R_+$ representing the graph of a function from U^* into R_+

$$\nu^*_{tr}(\mathscr{X}, Y) = \begin{cases} 1 \text{ if } & Max < tr \\ 1/2 \text{ if } & tr \leq Max < 1 - tr \\ 0 \text{ if } & Max \geq 1 - tr, \end{cases} \tag{31}$$

where

1. $Max = max\{\frac{|y^* - mid(\pi_2(Z)|}{max\{y^*, mid(\pi_2(Z)\}} : Z \in \mathscr{X}\ \&\ \nu^*_U(Z, Y) > 0\}$,
2. $\nu^*_U(Z, Y) = \nu^*\left((\pi_1(Z) \cap U) \times \pi_2(Z), Y \cap (U \times R_+)\right)$, where ν^* is defined by the equation (15),
3. $mid(\Delta) = \frac{a+b}{2}$, where $\Delta = [a, b)$,
4. $$y^* = \frac{1}{c} \sum_{Z \in \mathscr{X}:\ \nu^*_U(Z,Y) > 0} mid(\pi_2(Z)) \cdot card(\pi_1(Z \cap Y) \cap U), \tag{32}$$

   where

   $$c = card\left(\bigcup_{Z \in \mathscr{X}:\ \nu^*_U(Z,Y) > 0} \pi_1(Z) \cap U\right). \tag{33}$$

The lower approximation of Gf^* is defined by

$$\begin{gathered} LOW^*(AS^*, Gf^*) = \\ \{(x, y) : \nu^*_{tr}(I^*(x), Gf^*) = 1\ \&\ x \in U^*\ \&\ y \in [y^* - \delta/2, y^* + \delta/2)\}, \end{gathered} \tag{34}$$

where y^* is obtained from equation (32) in which $\mathscr{X}$ is substituted by $I(x)$ and Y by Gf^*, respectively.

The upper approximation of Gf^* is defined by

$$\begin{gathered} UPP^*(AS^*, Gf^*) = \\ \{(x, y) : \nu^*_{tr}(I^*(x), Gf^*) > 0\ \&\ x \in U^*\ \&\ y \in [y^* - \delta/2, y^* + \delta/2)\}, \end{gathered} \tag{35}$$

where y^* is obtained from equation (32) in which $\mathscr{X}$ is substituted by $I(x)$ and Y by Gf^*, respectively.

Let us observe that for $x \in U^*$ the condition $(x, y) \notin UPP^*(AS^*, Gf^*)$ means that $\nu^*_{tr}(I^*(x), Gf^*) = 0\ \&\ y \in [y^* - \delta/2, y^* + \delta/2)$ or $y \notin [y^* - \delta/2, y^* + \delta/2)$. The first condition describes the set of all pairs (x, y), where the deviation of y from y^* is small (relative to δ) but the prediction of y^* on the set of patterns $I^*(x)$ is very risky.

The values of f^* can be induced by

$$\widehat{f^*}(x) = \begin{cases} [y^* - \delta/2, y^* + \delta/2) & \text{if} \quad \nu^*_{tr}(I^*(x), Gf^*) > 0 \\ \quad undefined & \text{otherwise,} \end{cases} \tag{36}$$

where $x \in U^* \setminus U$ and y^* is obtained from equation (32) in which $\mathscr{X}$ is substituted by $I(x)$ and Y by Gf^*, respectively.

Let us now explain the formal definitions presented above. The value of uncertainty function $I^*(x)$ for a given object x consists all patterns of the form $\|\alpha\|_{U^*} \times \Delta_\alpha$ such that $\|\alpha\|_{U^*}$ is matched by the object x. The condition $x \in \|\alpha\|_{U^*}$ can be verified by checking if the A-signature of x, i.e., $Inf_A(x)$ is matching α (to a satisfactory degree). The deviation on $\|\alpha\|_{U^*}$ is bounded by the interval Δ_α of reals. The degree to which Z is included to Y is estimated by $\nu^*_U(Z,Y)$, i.e., by degree to which the restricted to U projection of the pattern Z is included into Y projected on U. The estimated value for $f^*(x)$ belongs to the interval $[y^* - \delta/2, y^* + \delta/2)$ obtained by fusion of centers of intervals assigned to patterns from $\mathscr{X}$. In this fusion, the weights of these centers are reflecting the strength on U of patterns matching Y to a positive degree. The result of fusion is normalized by c. The degree to which a family of patterns $\mathscr{X}$ is included into Y is measured by the deviation of the value y^* from centers of intervals of patterns Z matching Y to a positive degree (i.e., $\nu^*_U(Z,Y) > 0$). In Figure 3 we illustrate the idea of the presented definition of y^*, where

- $Z_i = \|\alpha_i\|_{U^*} \times \Delta_{\alpha_i}$ for $i = 1,2,3$,
- $I^*(x) = \{Z_1, Z_2, Z_3\}$,
- the horizontal bold lines illustrate projections of sets Z_i $(i = 1,2,3)$ on U^*,
- the vertical bold lines illustrate projections of sets Z_i $(i = 1,2,3)$ on R_+,
- $y^* = \frac{1}{c}\sum_{t=1}^{3} mid(\Delta_{\alpha_t})card(\|\alpha_t\|_U)$ and c is defined by equation (33),
- $\nu^*_U(Z_i, Gf^*) > 0$ for $i = 1,2,3$ because $(x_1, f(x_1)) \in Z_1$ and $(x_2, f(x_2)) \in Z_2 \cap Z_3$ for $x_1, x_2 \in U$,
- $\nu^*_{tr}(I^*(x), Gf^*) = 1$ means that deviations $|y^* - mid(\Delta_{\alpha_i})|$ are sufficiently small (the exact formula is given by (34)).

The quality of approximations of Gf^* can be estimated using some selected measure defined by a combination of

- closeness between projections on U of approximations of Gf^* and approximations of Gf on sample U (see formulas (21-22)),
- the description lengths of approximations.

The considered approximation space is parameterized by the set of patterns $PATTERN(A, f, \delta)$. The optimization problem for function approximation is defined as the searching problem for a set of patterns optimizing the mentioned above measure based on a version of the minimal length principle (Rissanen 1985). Next, the closeness between the result of optimization and a testing sample of Gf^* can be used for estimation of the approximation quality. Note that one can also tune some parameters of the selected measure for improving the approximation quality. A more detailed discussion on optimization of function approximation will be presented in our next paper.

The presented illustrative method for function approximation based on the rough set approach can be treated as one of many possible ways for inducing function

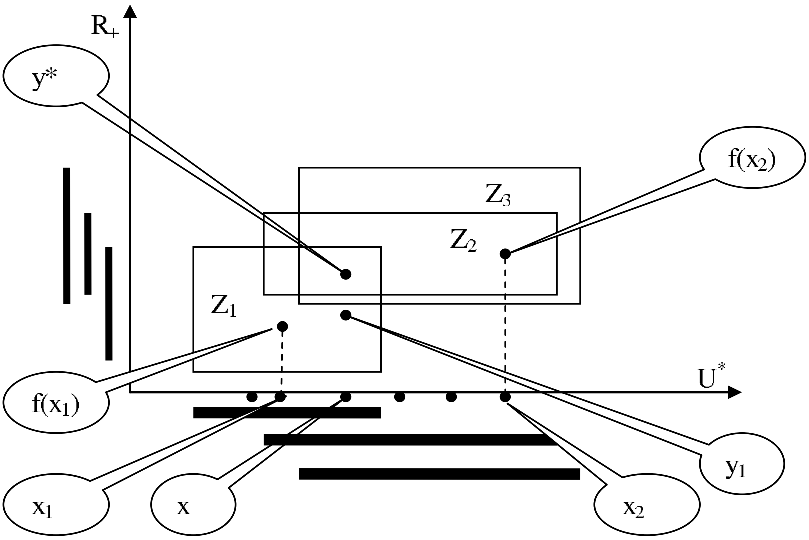

Fig. 3 Inducing the value y^*

approximations from data. For example, some of these methods may use distance functions between objects from U^* or more advanced rules for fusion of votes of matched by objects patterns voting for different approximations of function values.

Example 3. We present a method based on Boolean reasoning for extraction of patterns from data on which the deviation of values of the approximated function is bounded by a given threshold. These patterns are used as left hand sides of decision rules of the following form

if *pattern* **then** *the decision deviation is bounded by a given threshold.*

We consider a decision table $DT = (U, A \cup \{d\})$ such that $U = \{x_1, \ldots, x_5\}$, $A = \{a, b, c\}$ and $V_a = V_b = V_c = \{0, 1\}$ see Table 6 (the last column of the table (labeled by support) means the number of objects described exactly in the same way).

We define the square matrix $[m^{\varepsilon}_{x_i,x_j}]_{x_i,x_j \in U}$ (which is an analogy to discernibility matrix in the standard rough set model) by

$$m^{\varepsilon}_{x_i,x_j} = \{e \in A : e(x_i) \neq e(x_j) \& \mid d(x_i) - d(x_j) \mid > \varepsilon\}. \qquad (37)$$

Table 6 Decision Table with Real Valued Decision

U	a	b	c	d	*support*
x_1	0	0	0	160	10
x_2	1	0	0	165	20
x_3	1	0	1	170	30
x_4	0	1	0	175	20
x_5	0	1	1	180	20

Let us observe that for any $x_i, x_j \in U$ $m^{\varepsilon}_{x_i,x_j} \subseteq P(A)$.

Table 7 Square Matrix $[m^5_{x_i,x_j}]_{x_i,x_j \in U}$

	x_1	x_2	x_3	x_4	x_5
x_1	$\emptyset$	$\emptyset$	$\{a,c\}$	$\{b\}$	$\{b,c\}$
x_2	$\emptyset$	$\emptyset$	$\emptyset$	$\{a,b\}$	$\{a,b,c\}$
x_3	$\{a,c\}$	$\emptyset$	$\emptyset$	$\emptyset$	$\{a,b\}$
x_4	$\{b\}$	$\{a,b\}$	$\emptyset$	$\emptyset$	$\emptyset$
x_5	$\{b,c\}$	$\{a,b,c\}$	$\{a,b\}$	$\emptyset$	$\emptyset$

We assume that $\varepsilon = 5$ and the square matrix $[m^5_{x_i,x_j}]_{x_i,x_j \in U}$ is presented in Table 7.

From Table 7 we obtain the boolean function $g(a,b,c)$ with three boolean variables a,b,c corresponding to attributes:

$$g(a,b,c) = (a \vee c) \wedge b \wedge (b \vee c) \wedge (a \vee b) \wedge (a \vee b \vee c) \wedge (a \vee b). \qquad (38)$$

This function is an analogy to discernibility function in the standard rough set model. There are two prime implicants of this boolean function g: $a \wedge b$ and $b \wedge c$. Hence we obtain two reducts: $\{a,b\}$ and $\{b,c\}$. Using reducts and values of an object $x \in U$ we construct decision rules of the form

$$\textbf{if} \ \ldots \ \textbf{then} \ d \in (d(x) - \varepsilon, d(x) + \varepsilon). \qquad (39)$$

The set of decision rules based on reducts $\{a,b\}$ and $\{b,c\}$ is presented in Table 8.

Let us consider new object $x_{new} \notin U$ described by $a(x_{new}) = 0, b(x_{new}) = 0$ and $c(x_{new}) = 1$. In classification of x_{new} there are applied two decision rules r^1_1 and r^2_3. Using the quality of the rules $Quality(r^1_1) = 10$, $Quality(r^2_3) = 30$ and that $mid(155,165) = 160$, $mid(165,175) = 170$ we obtain that

$$d(x_{new}) = \frac{Quality(r^1_1)mid(155,165) + Quality(r^2_3)mid(165,175)}{Quality(r^1_1) + Quality(r^2_3)} = 167.5. \qquad (40)$$

Table 8 Decision rules generated from reducts $\{a,b\}$ and $\{b,c\}$

Object	Decision rule	Quality of the rule
x_1	r^1_1 : **if** $a = 0$ **and** $b = 0$ **then** $d \in (155,165)$	10
x_1	r^2_1 : **if** $b = 0$ **and** $c = 0$ **then** $d \in (155,165)$	10
x_2	r^1_2 : **if** $a = 1$ **and** $b = 0$ **then** $d \in (160,170)$	20
x_2	r^2_2 : **if** $b = 0$ **and** $c = 0$ **then** $d \in (160,170)$	20
x_3	r^1_3 : **if** $a = 1$ **and** $b = 0$ **then** $d \in (165,175)$	30
x_3	r^2_3 : **if** $b = 0$ **and** $c = 1$ **then** $d \in (165,175)$	30
x_4	r^1_4 : **if** $a = 0$ **and** $b = 1$ **then** $d \in (170,180)$	20
x_4	r^2_4 : **if** $b = 1$ **and** $c = 0$ **then** $d \in (170,180)$	20
x_5	r^1_5 : **if** $a = 0$ **and** $b = 1$ **then** $d \in (175,185)$	20
x_5	r^2_5 : **if** $b = 1$ **and** $c = 1$ **then** $d \in (175,185)$	20

4 Conclusions

We discussed a generalization of approximation spaces based on granular formulas and neighborhoods. We emphasized the fundamental role of approximation spaces in inducing classifiers such as rule based classifiers or function approximations. The approach can be extended for other kinds of classifiers, e.g., knn classifiers, neural networks, or ensembles of classifiers. Efficient strategies searching for relevant approximation spaces for approximation of higher order granules are crucial for application (e.g., in searching for approximation of complex concepts). In our current projects, we are developing such strategies based on some versions of the minimal length principle (Rissanen 1985). The presented approach provides the uniform foundations for implementing diverse strategies searching for (semi)optimal classifiers of different kinds. All of these strategies are based on searching for relevant approximation spaces. In particular, searching for relevant neighborhoods can be supported by feature construction methods based on Boolean reasoning. Searching for relevant approximation spaces can be realized in a network of cooperating strategies searching for classifiers for a given data set. The uniform foundations of our approach can facilitate the cooperation among strategies in such network. We also plan to compare the proposed method for function approximation with the traditional ones such as regression based methods (Hastie et al. 2008) or methods based on the functional data analysis approach (Ramsay 2002).

Acknowledgements

The research has been partially supported by the grants N N516 069235, N N516 368334, and N N516 077837 from Ministry of Science and Higher Education of the Republic of Poland.

References

Bazan, J.G.: Hierarchical classifiers for complex spatio-temporal concepts. In: Peters, J.F., Skowron, A., Rybiński, H. (eds.) Transactions on Rough Sets IX. LNCS, vol. 5390, pp. 474–750. Springer, Heidelberg (2008)

Bazan, J.G., Osmólski, A., Skowron, A., Ślęzak, D., Szczuka, M.S., Wróblewski, J.: Rough Set Approach to the Survival Analysis. In: Alpigini, J.J., Peters, J.F., Skowron, A., Zhong, N. (eds.) RSCTC 2002. LNCS (LNAI), vol. 2475, pp. 522–529. Springer, Heidelberg (2002)

Bazan, J., Skowron, A., Swiniarski, R.: Rough sets and vague concept approximation: From sample approximation to adaptive learning. In: Peters, J.F., Skowron, A. (eds.) Transactions on Rough Sets V. LNCS, vol. 4100, pp. 39–62. Springer, Heidelberg (2006)

Bianucci, D., Cattaneo, G.: Information entropy and granulation co–entropy of partitions and coverings: A summary. In: Peters, J.F., Skowron, A., Wolski, M., Chakraborty, M.K., Wu, W.-Z. (eds.) Transactions on Rough Sets X. LNCS, vol. 5656, pp. 15–66. Springer, Heidelberg (2009)

Greco, S., Matarazzo, B., Słowiński, R.: Dominance-based rough set approach and bipolar abstract rough approximation spaces. In: Chan, C.-C., Grzymala-Busse, J.W., Ziarko, W.P. (eds.) RSCTC 2008. LNCS (LNAI), vol. 5306, pp. 31–40. Springer, Heidelberg (2008)

Grünwald, P.D.: The minimum description length principle. The MIT Press, Cambridge (2007)

Hastie, T., Tibshirani, R., Friedman, J.H.: The Elements of statistical learning: Data mining, inference, and prediction, 2nd edn. Springer, Heidelberg (2008)

Jankowski, A., Skowron, A.: A wistech paradigm for intelligent systems. In: Peters, J.F., Skowron, A., Düntsch, I., Grzymała-Busse, J.W., Orłowska, E., Polkowski, L. (eds.) Transactions on Rough Sets VI. LNCS, vol. 4374, pp. 94–132. Springer, Heidelberg (2007)

Jankowski, A., Skowron, A.: Logic for artificial intelligence: The Rasiowa-Pawlak school perspective. In: Ehrenfeucht, A., Marek, V., Srebrny, M. (eds.) Andrzej Mostowski and Foundational Studies, pp. 106–143. IOS Press, Amsterdam (2008)

Klir, G.J.: Uncertainty and information: Foundations of generalized information theory. John Wiley & Sons, Hoboken (2007)

Łukasiewicz, J.: Die logischen Grundlagen der Wahrscheinlichkeitsrechnung. In: Borkowski, L. (ed.) Jan Łukasiewicz - Selected works, pp. 16–63. North Holland Publishing Company/Polish Scientific Publishers, Amsterdam/Warsaw (1913)

Malyszko, D., Stepaniuk, J.: Adaptive multilevel rough entropy evolutionary thresholding. Inf. Sci. 180(7), 1138–1158 (2010)

Ng, K.S., Lloyd, J.W., Uther, W.T.B.: Probabilistic modelling, inference and learning using logical theories. Ann. of Math. and Artif. Intell. 54(1-3), 159–205 (2008)

Nguyen, H.S.: Approximate Boolean Reasoning: Foundations and Applications in Data Mining. In: Peters, J.F., Skowron, A. (eds.) Transactions on Rough Sets V. LNCS, vol. 4100, pp. 344–523. Springer, Heidelberg (2006)

Pal, S.K., Shankar, B.U., Mitra, P.: Granular computing, rough entropy and object extraction. Pattern Recognition Letters 26(16), 2509–2517 (2005)

Pawlak, Z.: Rough sets: Theoretical aspects of reasoning about data. In: System Theory, Knowledge Engineering and Problem Solving, vol. 9. Kluwer Academic Publishers, Dordrecht (1991)

Pawlak, Z., Skowron, A.: Rudiments of rough sets; Rough sets: Some extensions; Rough sets and boolean reasoning. Inf. Sci. 177(1), 3–27, 28–40, 41–73 (2007)

Pedrycz, W., Skowron, A., Kreinovich, V. (eds.): Handbook of granular computing. John Wiley & Sons, New York (2008)

Peters, J., Skowron, A., Stepaniuk, J.: Nearness of objects: extension of the approximation space model. Fundam. Inform. 79(3,4), 497–512 (2007)

Polkowski, L., Skowron, A.: Rough mereology: A new paradigm for approximate reasoning. Int. J. Approx. Reason. 51, 333–365 (1996)

Ramsay, J.O., Silverman, B.W.: Applied functional data analysis. Springer, Berlin (2002)

Rissanen, J.: Minimum-description-length principle. In: Kotz, S., Johnson, N. (eds.) Encyclopedia of statistical sciences, pp. 523–527. John Wiley & Sons, New York (1985)

Skowron, A., Stepaniuk, J.: Generalized approximation spaces. In: Lin, T.Y., Wildberger, A.M. (eds.) The Third Int. Workshop on Rough Sets and Soft Computing Proceedings (RSSC 1994), pp. 156–163. San Jose State University, San Jose (1994)

Skowron, A., Stepaniuk, J.: Tolerance approximation spaces. Fundam. Inform. 27, 245–253 (1996)

Skowron, A.: Toward intelligent systems: Calculi of information granules. In: Hirano, S., Inuiguchi, M., Tsumoto, S. (eds.) Proc. Int. Workshop On Rough Set Theory And Granular Computing (RSTGC 2001),Matsue, Japan, May 20-22 (2001); Bull. Int. Rough Set Society 5(1-2), 9–30; see also: Terano T., Nishida T., Namatame A., Tsumoto S., Ohsawa Y., Washio T. (eds) New frontiers in Artificial Intelligence, Joint JSAI Workshop Post Proceedings.LNCS(LNAI), vol. 2253, pp. 251–260. Springer, Heidelberg

Skowron, A., Stepaniuk, J., Peters, J., Swiniarski, R.: Calculi of approximation spaces. Fundam. Inform. 72(1-3), 363–378 (2006)
Skowron, A., Szczuka, M.: Toward interactive computations: A rough-granular approach. In: Koronacki, J., Raś, Z.W., Wierzchoń, S.T., Kacprzyk, J. (eds.) Advances in Machine Learning II. SCI, vol. 263, pp. 23–42. Springer, Heidelberg (2010)
Słowiński, R., Vanderpooten, D.: A generalized definition of rough approximations based on similarity. IEEE Trans. Knowl. and Data Eng. 12, 331–336 (2000)
Stepaniuk, J.: Rough–granular computing in knowledge discovery and data mining. Springer, Heidelberg (2008)
Zadeh, L.A.: Fuzzy sets and information granularity. In: Gupta, M., Ragade, R., Yager, R. (eds.) Advances in Fuzzy Set Theory and Applications, pp. 3–18. North-Holland, Amsterdam (1979)
Zadeh, L.A.: Generalized theory of uncertainty (GTU)-principal concepts and ideas. Comput. Stat. and Data Anal. 51, 15–46 (2006)
Zhu, W.: Relationship between generalized rough sets based on binary relation and covering. Inf. Sci. 179(3), 210–225 (2009)
Ziarko, W.: Variable precision rough set model. J. Comput. and Sys. Sci. 46, 39–59 (1993)