Nearness of Objects:  
Extension of Approximation Space Model

James F. Peters*  
Department of Electrical and Computer Engineering,  
University of Manitoba, Winnipeg, Manitoba R3T 5V6 Canada  
jfpeters@ee.umanitoba.ca

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

Jaroslaw Stepaniuk  
Department of Computer Science, Białystok University of Technology  
Wiejska 45A, 15-351 Białystok, Poland  
jstepan@ii.pb.bialystok.pl

Near To
How near to the bark of a tree are the drifting snowflakes,  
swirling gently round, down from winter skies?  
How near to the ground are icicles,  
slowly forming on window ledges?  
...  
– Z. Pawlak and J.F. Peters,  
Spring, 2002.

Abstract. The problem considered in this paper is the extension of an approximation space to include a nearness relation. Approximation spaces were introduced by Zdzisław Pawlak during the early 1980s as frameworks for classifying objects by means of attributes. Pawlak introduced approximations as a means of approximating one set of objects with another set of objects using an indiscernibility relation that is based on a comparison between the feature values of objects. Until now, the focus has been on the overlap between sets. It is possible to introduce a nearness relation that can be used to determine the “nearness” of sets of objects that are possibly disjoint and, yet, qualitatively near to each other. Several members of a family of nearness relations are introduced in this article. The contribution of this article is the introduction of a nearness relation that makes it possible to extend Pawlak’s model for an approximation space and to consider the extension of generalized approximations spaces.

* Address for correspondence: Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, Manitoba R3T 5V6 Canada
Keywords: rough sets, approximation, approximation space, classification, feature, nearness relation, neighbourhood, proximity space

1. Introduction

Considerable work on approximation and approximation spaces and their applications has been carried in recent years [6, 7, 8, 9, 10, 20, 21, 22, 26, 29, 33, 35, 36, 37, 38, 41, 46, 45, 47, 49, 50, 52]. Zdzisław Pawlak introduced approximation spaces during the early 1980s as part of his research on classifying objects by means of attributes [20, 21, 22]. Approximation plays a fundamental role in rough set theory also introduced by Pawlak. There are close ties between approximation and general topology. For example, the approach to lower approximation and upper approximation introduced by Pawlak is closely related to the topological interior and closure operators, respectively [29, 39]. Topology is a rich source of constructs that can be used to enrich the original model of an approximation space as well as more recent models of generalized approximation spaces. For example, the nearness relation introduced in this article has mainly been inspired by Pawlak’s original work on the classification of objects [20] and by work on proximity spaces by Riesz [42], Efremović [2, 3] and others (see, e.g., [4, 16]). This has recently led to an approach to object recognition and extension of approximation spaces based on near sets (see, e.g., [12, 34] and [46, 47], respectively).

The problem considered in this paper is the extension of the Pawlak model for an approximation space to include a nearness relation [1, 5]. Until now, the focus has been on the overlap between sets. It is possible to introduce a nearness relation that can be used to discover the “nearness” of objects that are possibly disjoint and, yet, qualitatively near each other. The term qualitatively near is used here to mean closeness of descriptions or distinctive characteristics of objects. The term object denotes something perceptible. If we choose shading as a feature and let $B_{\text{shading}}(x) = \{y \mid \text{shading}(x) = \text{shading}(y)\}$, then the objects in Fig. 1 can be partitioned, where the objects in $B_{\text{shading}}(x)$, i.e., equivalence class containing objects that are descriptively indiscernible from $x$, are not adjacent to each other (see, e.g., $B_{\text{shading}}(g1) = \{g1, g2, g3\}$ in Fig. 1.1) or $B_{\text{shading}}(x1) = \{x1, x11, x15, x16\}$ in Fig. 1.2).

1.1: Hexagonal Objects

1.2: Pixel Neighbourhoods

Figure 1. Non-Adjacent Objects with Matching Descriptions

The approach to classifying objects such as those in Fig. 1.2 contrasts sharply with the approach to defining neighbourhoods with an adjacency relation in [5]. For example, the hexagons with mesh...
interiors in Fig. 1.1 are descriptively near each other but spatially non-adjacent. The refinement of approximation spaces in this article is close to what is known as a nearness space viewed from a topological perspective [13, 14, 52]. In addition, the proposed approach to nearness of objects is not restricted to the neighbourhood of a point \(x\) and \(x \in Cl(A)\) (closure of \(A\)) as in [52], since we consider the nearness of objects that are not points. It also has been pointed out in [52] that if we consider an approximation space \((X, E)\) with binary connection relation \(E\), then \((X, \delta_E)\) is a proximity space. This observation is based on [51]. Let \(x, y \in X\). Then

\[ x \delta_E y \iff \text{closed regions } x \text{ and } y \text{ share a common point}. \]

Then, for example, \(g_2 \delta_E h_1\) but not \((g_2 \delta_E g_3)\) in Fig. 1.1, since \(g_2\) and \(h_1\) share a common point, whereas \(g_2\) and \(g_3\) have no points in common. The feature-based nearness relation \(\delta\) introduced in this article contrasts with \(\delta_E\) in [52, 51] inasmuch as \(\delta\) is defined relative to objects with matching feature values. The contribution of this article is the introduction of a nearness relation that makes it possible to extend the Pawlak’s model for an approximation space and to consider the extension of generalized approximations spaces.

The article is organized as follows. Basic notions and notation for the classification of objects and approximation that form the basis for rough set theory are briefly introduced in Section 2. Nearness in proximity spaces is introduced in Section 3. A nearness form of generalized approximation space is presented in Section 4.

2. Rough Sets

If we classify objects by means of attributes, exact classification is often impossible.


A brief presentation of the foundations of rough set theory is given in this section. Rough set theory has its roots in Zdzisław Pawlak’s research on knowledge representation systems during the early 1970s [19]. Rather than attempt to classify objects exactly by means of attributes [20] or by means of features [32], Pawlak considered an approach to solving the object classification problem in a number of novel ways. First, in 1973, he formulated knowledge representation systems (see, e.g., [15, 19]). Then, in 1981, Pawlak introduced approximate descriptions of objects and considered knowledge representation systems in the context of upper and lower classification of objects relative to their attribute values [20, 21].

2.1. Information systems

We start with a system \(S = (X, A, V, \sigma)\), where \(X\) is a non-empty set of objects, \(A\) is a set of attributes, \(V\) is a union of sets \(V_a\) of values associated with each \(a \in A\), and \(\sigma\) is called a knowledge function defined as the mapping \(\sigma : X \times A \rightarrow V\), where \(\sigma(x, a) \in V_a\) for every \(x \in X\) and \(a \in A\). The function \(\sigma\) is referred to as a knowledge function about objects from \(X\). The set \(X\) is partitioned into elementary sets that later were called blocks, where each elementary set contains those elements of \(X\) which have matching attribute values. In effect, a block (elementary set) represents a granule of knowledge (see
2.1: Blocks of Objects

The universe of objects

2.2: Sample Set Approximation

Fig. 2. Rudiments of Rough Sets

Fig. 2.2). For example, for any $B \subseteq A$, the $B$-elementary set for an element $x \in X$ denoted by $B(x)$, is defined by

$$B(x) = \{ y \in X | \forall a \in B \sigma(x, a) = \sigma(y, a) \} \quad (1)$$

Consider, for example, Fig. 2.1 which represents a system $S$ containing a set $X$ of shaded circles and a feature set $A$ that contains only one feature, namely, shade. Assume that each circle in $X$ has only one colour. Then the set $X$ is partitioned into elementary sets or blocks, where each block contains circles with the same shade. In effect, elements of a set $B(x) \subseteq X$ in a system $S$ are classified as indiscernible if they are indistinguishable by means of their feature values for any $a \in B$. A set of indiscernible elements is called an elementary set [20]. Hence, any subset $B \subseteq A$ determines a partition $\{ B(x) : x \in X \}$ of $X$. This partition defines an equivalence relation $Ind(B)$ on $X$ called an indiscernibility relation such that $x \ Ind(B) y$ if and only if $y \in B(x)$ for every $x, y \in X$.

Assume that $Y \subseteq X$ and $B \subseteq A$, and consider an approximation of the set $Y$ by means of the attributes in $B$ and $B$-indiscernible blocks in the partition of $X$. The union of all blocks that constitute a subset of $Y$ is called the lower approximation of $Y$ (usually denoted by $B_*Y$), representing certain knowledge about $Y$. The union of all blocks that have non-empty intersection with the set $Y$ is called the upper approximation of $Y$ (usually denoted by $B^+Y$), representing uncertain knowledge about $Y$. The set $BN_B(Y) = B^+Y - B_*Y$ is called the $B$-boundary of the set $Y$. In the case where $BN_B(Y)$ is non-empty, the set $Y$ is a rough (imprecise) set. Otherwise, the set $Y$ is a crisp set. This approach to classification of objects in a set is represented graphically in Fig. 2.2, where the region bounded by the ellipse represents a set $Y$, the darkened blocks inside $Y$ represent $B_*Y$, the grey blocks represent the boundary region $BN_B(Y)$, and the grey and the darkened blocks taken together represent $B^+Y$.

Consequences of this approach to the classification of objects by means of their feature values have been remarkable and far-reaching. Detailed accounts of the current research in rough set theory and its applications are available, e.g., in [29, 30, 31]).
2.2. Approximation

Some categories (subsets of objects) cannot be expressed exactly by employing available knowledge. Hence, we arrive at the idea of approximation of a set by other sets.


One of the most profound, very important notions underlying rough set theory is approximation. In general, an approximation is defined as the replacement of objects by others that resemble the original objects in certain respects [11]. For example, consider a universe \( U \) containing objects representing behaviours of agents. In that case, we can consider blocks of behaviours in the partition \( U/R \), where the behaviours within a block resemble (are indiscernible from) one another by virtue of their feature values. Then any subset \( X \) of \( U \) can be approximated by blocks that are either subsets of \( X \) (lower approximation of the set \( X \) denoted \( R_* X \)) or by blocks having one or more elements in common with \( X \) (upper approximation of the set \( X \) denoted \( R^* X \)). In rough set theory, the focus is on approximating one set of objects by means of another set of objects based on the feature values of the objects [26]. The lower approximation operator \( R_* \) has properties that correspond closely to properties of what is known as the \( \Pi_0 \) topological interior operator [22, 39]. Similarly, the upper approximation operator \( R^* \) has properties that correspond closely to properties of the \( \Pi_0 \) topological closure operator [22, 39]. It was observed in [22] that the key to the rough set approach is provided by the exact mathematical formulation of the concept of approximative (rough) equality of sets in a given approximation space.

3. Nearness Relation in Proximity Spaces Defined by Partitions

This section introduces a proximity space \((U, \delta)\) defined relative to a nearness relation \( \delta \). It also considers an extension of the approximation space introduced by Zdzisław Pawlak [22] as well as an extension of generalized approximation spaces in Sect. 4.

3.1. Basic definition

The basic approach in this section is to define a relation is near so that the assertion \( X \) is near \( Y \) is explained relative to features of objects in non-empty subsets \( X, Y \subseteq U \), where is \( U \) a non-empty set of objects called the universe. In this section, a nearness relation \( \delta \) is defined in a natural way relative to the feature values of objects in \( U \) following the convention established by Pawlak for the classification of objects by means of their attributes [20]. The distinction between the terms attribute (i.e., property of an object represented by a partial function) and feature (appearance of an object represented by probe functions) is explained in [32]. To see this, let \( X, Y \subseteq U \) and let \( B \subseteq A \) (set of features of objects in \( U \)). In what follows, assume that all mappings associated with elements of \( A \) are real-valued (for symbolic features one can use, for example, value difference metric [48]). For simplicity, if \( a \in A, x \in U \), we write \( a(x) \) to represent a functional value \( a \) in abstraction from the model of \( a \) either as an attribute or as a feature. Then define \( D_B \) as shown in (2).

\[
D_B(x, y) = \sum_{a \in B} |a(x) - a(y)|. \tag{2}
\]
The pseudometric defined by Eqn. (2) can be extended on sets by (3).

\[ D_B(X,Y) = \inf_{x \in X, y \in Y} D_B(x,y). \]  

(3)

A requirement of nearness (i.e., only one object \( y \in Y \) has feature values that match the feature values of at least one object \( x \in X \)) and non-nearness can be formulated as shown in (4) and (5), respectively.

\[ X \delta Y \text{ if and only if } \exists y \in Y, \exists x \in X : D_B(x,y) = 0, \]  

(4)

where the “\( \cdot \)“ is used to separate the parts of a formula [43, 44]. Let \( X \overline{\delta} Y \) denote \( \text{not}(X \delta Y) \), i.e., negation of nearness as defined in (4). Then, one can observe

**Proposition 3.1.**

\[ X \overline{\delta} Y \text{ if and only if } \forall y \in Y, \forall x \in X : D_B(x,y) \neq 0. \]  

(5)

**Remark 3.1.** The intuition underlying (4) is that \( X \delta Y \) denotes the fact that the knowledge represented by all of the feature values of one of the objects in \( Y \) matches the feature values of one of the objects in the set \( X \). The assertion \( X \overline{\delta} Y \) in (5) says that if there does not exist an object \( y \in Y \) whose feature values match the feature values of an object \( x \in X \), then \( X \) is not near to \( Y \). No attempt has been made to give a measure of the degree of nearness, which is outside the scope of this article.

One can also observe that we have the following result.

**Proposition 3.2.** For any sets \( X, Y \subseteq U \)

\[ X \delta Y \text{ if, and only if } B^*X \cap B^*Y \neq \emptyset. \]

**Proof:** Observe\(^1\) from (4) that for any \( X, Y \subseteq U \), it is the case that \( X \delta Y \) if there exists \( y \in Y \) and \( x \in X \), where “\( x \) and \( y \) are indiscernible” if there exists \( x \in U \) such that \( B(x) \cap X \neq \emptyset \) and \( B(x) \cap Y \neq \emptyset \). Hence, \( X \delta Y \) if \( B^*X \cap B^*Y \neq \emptyset \).

\( \square \)

The upper approximations \( B^*X \), \( B^*Y \) are closures of \( X, Y \), respectively, in a topology defined by the \( B \)-indiscernibility relation. Hence, (4) defines the standard proximity (see [1]).

Let us recall from [16] a definition of \( \delta \)-neighbourhood in a proximity space \((U, \delta)\).

**Definition 3.1.** We say \( Y \) is a proximal or \( \delta \)-neighbourhood of \( X \), written \( X \ll Y \), if and only if \( X \delta Y \) and not \( X \delta (U - Y) \).

We have the following fact characterizing the \( B \)-upper approximation of \( X \subseteq U \) using the \( \delta \)-neighbourhood relation.

**Proposition 3.3.** Let \((U, \delta)\) be a proximity space where \( \delta \) is defined by (4). Then for any non-empty subset \( X \) of \( U \), we have (6).

\[ X \ll B^*X, \]  

i.e., the upper approximation \( B^*X \) of \( X \) is a \( \delta \)-neighbourhood of \( X \).

\(^1\)We follow the convention that iff is an abbreviation for if, and only if.
The property (6) characterizes the upper approximation of $X$ in terms of a proximity space. In fact, by we have $X \delta B^*X$ and, by definition of the upper approximation, there is no $y \in U$ such that $B(y) \cap X \neq \emptyset$ and $B(y) \cap (U - B^*X) \neq \emptyset$. Hence, $D_B(X, (U - B^*X)) \neq 0$, i.e., $non(X \delta (U - B^*X))$.

Let us recall the definition of a proximity space.

**Definition 3.2.** Proximity Space [16].

A binary relation $\delta$ on $\mathcal{P}(U)$ (powerset of $U$) is called a proximity of $U$ iff $\delta$ satisfies proximity axioms 1-5. The pair $(U, \delta)$ is called a proximity space.

1. $X \delta Y$ implies $Y \delta X$.
2. $(X \cup Y) \delta Z$ implies $X \delta Z$ or $Y \delta Z$.
3. $X \delta Y$ implies $X \neq \emptyset$, $Y \neq \emptyset$.
4. $X \overline{\delta} Y$ implies there exists $E \subseteq U$ so that $X \delta E$ and $(U - E) \overline{\delta} Y$.
5. $X \cap Y \neq \emptyset$ implies $X \delta Y$.

Let us observe that if the relation $\delta$ is defined by (4) and $U$ is finite, then $(U, \delta)$ is a proximity space. Under these assumptions it is easy to see that we have $D_B(x, y) = 0$ for some $x \in X$, $y \in Y$ if and only if $D_B(X, Y) = 0$. Hence, we have $X \delta Y$ if and only if $D_B(X, Y) = 0$, where $D_B$ is induced from the pseudometric $d_B$. In such a case, $(U, \delta)$ is a proximity space (see [16], page 8).

### 3.2. Relations between nearness and approximations

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.


In this section, we briefly consider an extension the approximation space introduced in [22] with the nearness relation $\delta$. In [26], an approximation space is represented by the pair $(U, R)$, where $U$ is a universe of objects, and $R \subseteq U \times U$ is an indiscernibility relation (denoted $Ind$ as in Sect. 2) defined by an attribute set (i.e., $R = Ind(A)$ for some attribute set $A$). In this case, $R$ is an equivalence relation. Let $[x]_R$ denote the equivalence class of an element $x \in U$ under the indiscernibility relation $R$, i.e., $[x]_R = \{y \in U : xRy\}$.

In this context, $R$-approximations of any set $X \subseteq U$ are based on the exact (crisp) containment of sets. Then set approximations are defined as follows:

- $x \in U$ belongs with certainty to $X \subseteq U$ (i.e., $x$ belongs to the $R$-lower approximation of $X$), if $[x]_R \subseteq X$.
- $x \in U$ possibly belongs $X \subseteq U$ (i.e., $x$ belongs to the $R$-upper approximation of $X$), if $[x]_R \cap X \neq \emptyset$. 

\[ x \in U \] belongs with certainty neither to the \( X \) nor to \( U - X \) (i.e., \( x \) belongs to the \( R \)-boundary region of \( X \)), if \([x]_R \cap (U - X) \neq \emptyset \) and \([x]_R \cap X \neq \emptyset \).

A nearness approximation space is represented by \((U, R, \delta)\), where \( \delta \) is the nearness relation defined in (4) in Sect. 3 and \( R \) is the indiscernibility relation defined by a set of attributes \( B \). What follows is a selection of properties of an approximation space relative to the nearness relation.

1. \( B(x) \delta X \) for any \( B(x) \subseteq B_*X \), since \( B_\delta(x) \cap X \neq \emptyset \).
2. \( B(x) \delta B^*X \) for any \( B(x) \subseteq B^*X \). Similarly, we have
3. \( B_\delta(x) \delta B_*X \) for any \( B(x) \subseteq B_*X \).
4. \( B_\delta(x) \delta B^*X \), since \( B_\delta(x) \subseteq B^*X \).
5. \( BN_\delta(X) \delta B_*X \), since \( BN_\delta(X) \cap B_*X = \emptyset \), and there is no object in \( B_*X \) that matches an object in \( BN_\delta(X) \).

### 3.3. Nearness and approximation extension

In this section, we illustrate an application of nearness when inducing of approximation extension.

Let us consider an extension \((U^\infty, R^\infty, \delta^\infty)\) of an approximation space \((U, R, \delta)\) where \( U \subseteq U^\infty \), \( R = R^\infty \cap (U \times U) \) and \( \delta = \delta^\infty \cap (U \times U) \). \( R^\infty \) is the indiscernibility relation in \( U^\infty \) defined by \( B \) and \( \delta^\infty \) is defined by (4).

**Remark 3.2.** \( \delta \) and \( \delta^\infty \) are binary relations on \( \mathcal{P}(U) \) and \( \mathcal{P}(U^\infty) \), respectively, where \( \mathcal{P}(X) \) denotes the powerset of the set \( X \).

Let us assume that \( X \subseteq U^\infty \) and only a partial information \( X \cap U \) about \( X \) is available. Then one can consider the lower approximation \( B_\delta(X \cap U) \) and the upper approximation \( B^*(X \cap U) \) of \( X \) in \( U \). These approximations can be extended to \( B_\delta(X) \) and \( B^*(X) \) in \( U^\infty \). For example, one can use the following inductive rules:

- for \( B_\delta \) : if \( x \in U^\infty - U \) and \( B(x) \delta B(y) \) for some \( y \in U \) such that \( y \in B_\delta(X \cap U) \) then \( x \in B_\delta X \);
- for \( B^* \) : if \( x \in U^\infty - U \) and \( B(x) \delta B(y) \) for some \( y \in U \) such that \( y \in B^*(X \cap U) \) then \( x \in B^*X \);
- for \( BN_\delta \) : if \( x \in U^\infty - U \) and \( B(x) \delta B(y) \) for some \( y \in U \) such that \( y \in BN_\delta(X \cap U) \) then \( x \in BN_\delta(X) \).

The above inductive rules are making it possible to classify to the approximation regions unseen objects, i.e., objects not belonging to the training sample \( U \). This can be done using the indiscernibility of these objects to some objects from the training sample. Certainly, one can extend this approach to the case of indiscernibility defined by similarity (tolerance) relation. Our idea is analogous to that used in nearest neighbour classifiers, i.e., unseen objects are classified on the basis of indiscernibility (similarity) to objects from the training sample. This idea was also used in the definition of the approximation space extension [46, 47].
3.4. Some remarks on non-standard proximities

Observe, that one can define another concept of nearness. That is, the assertions “$X$ is near $Y$” and “$X$ is not near $Y$” can be defined as shown in (7) and (8), respectively.

$$X \delta Y \text{ if f } \exists y \in Y \cdot \forall x \in X \cdot D_B(x, y) = 0. \quad (7)$$

$$X \overline{\delta} Y \text{ if f } \forall y \in Y \cdot \exists x \in X \cdot D_B(x, y) \neq 0. \quad (8)$$

The formulation of nearness in (7) asserts that if there exists an object $y$ in set $Y$ whose feature values match the feature values of every object $x$ in set $X$, then $X$ is near $Y$. The intuition underlying (7) is that $X \delta Y$ (i.e., $X$ is near $Y$) whenever the knowledge represented by all of the feature values of one of the objects in $Y$ matches the feature values of all of the objects in the set $X$ (for simplicity we will write $\delta$ ($\overline{\delta}$) instead of $\delta_B$ ($\overline{\delta_B}$)). In defining nearness of one set of objects to another set of objects, definition (7) prescribes a minimum condition for nearness, namely, only one of the objects in one set must have feature values that match those of all of the objects in the other set. The assertion $X \overline{\delta} Y$ in (8) says that if there does not exists an object $y \in Y$ whose feature values match the feature values of every object $x \in X$, then $X$ is not near to $Y$. In effect, $X \overline{\delta} Y$ occurs in the case where it is determined that no object in $Y$ matches our knowledge of all of the objects in $X$.

Example 3.1. Sets of Coloured Circles.

By way of illustration of the nearness relation, consider $\delta$ as defined in (7). Let $X, Y$ denote two sets of circles and assume $B \subseteq A$ contains one feature, namely, brightness (i.e., descriptor of visual perception). Further, let the brightness of each circle be represented by a number indicating the average brightness.
of the pixels contained in a circle in a greyscale image (0 for white, and 1 for black). Examples are given in Fig. 3. In Fig. 3, assume that each circle is made up of all white or all black pixels. Then, for example in Fig. 3.1, \( X \delta Y \) (i.e., \( X \) is near \( Y \) because there is a white circle in \( Y \) with an average brightness that matches the average brightness of all of the circles in \( X \)). That is, \( D(X1, Y1) = 0 \).

Similarly, in Fig. 3.4, \( X4 \delta Y4 \), since \( D(X4, Y4) = 0 \). This happens because either one of the black circles in \( Y4 \) has an average brightness matching the average brightness of all of the circles in \( X4 \). In Fig. 3.2, \( X2 \not\subseteq X2 \), since \( D(X2, Y2) \neq 0 \). That is, there is no circle in \( Y2 \) with an average brightness that matches the average brightness of all of the circles in \( X2 \). The presence of mixed circles in both sets prevents nearness. Again, for example, \( X3 \) and \( Y3 \) in Fig. 3.3 have no matching circles (i.e., the circles in the two sets have non-matching colours). Hence, \( D(X3, Y3) \neq 0 \), i.e., \( X3 \not\subseteq Y3 \). Notice that if \( \delta \) is defined as in (4), then \( X2 \not\subseteq X2 \).

Let us observe that \( \delta \) defined in (7), in general, fails to satisfy axiom 1 of the proximity space (see Def. 3.2). To see this, consider, for example, \( X1 \delta Y1 \), but \( Y1 \not\subseteq X1 \) (i.e., \( Y1 \) is not near \( X1 \)) in Fig. 3.1, since there is no object in \( X1 \) that matches all of the objects in \( Y1 \).

### 4. Nearness in Generalized Approximation Spaces

Several generalizations of the classical rough set approach based on approximation spaces defined as pairs of the form \((U, R)\), where \(R\) is the equivalence relation (called an indiscernibility relation) on a non-empty set \(U\), have been reported in the literature (see, e.g., [29, 30, 31, 45, 47, 49]). Let us mention two of them.

A generalized approximation space can be defined by a tuple \(GAS = (U, N, \nu)\) where \(N\) is a *neighbourhood function* defined on \(U\) with values in the powerset \(\mathcal{P}(U)\) of \(U\) (i.e., \(N(x)\) is the *neighbourhood* of \(x\)) and \(\nu\) is an *overlap function* defined on the Cartesian product \(\mathcal{P}(U) \times \mathcal{P}(U)\) with values in the interval \([0, 1]\) measuring the degree of overlap of sets. The lower \(GAS_{\ast}\) and upper \(GAS^{\ast}\) approximation operations can be defined in a \(GAS\) by (9) and (10).

\[
GAS_{\ast}(X) = \{ x \in U : \nu(N(x), X) = 1 \}, \quad \text{(9)}
\]
\[
GAS^{\ast}(X) = \{ x \in U : \nu(N(x), X) > 0 \}. \quad \text{(10)}
\]

In the standard case, \(N(x)\) is equal to the equivalence class \([x]_B\) of the indiscernibility relation \(Ind(B)\) for a set of features \(B\). In the case where \(R\) is a tolerance (similarity) relation\(^2\), \(\tau \subseteq U \times U\), we take \(N(x) = \{ y \in U : \tau xy \}\), i.e., \(N(x)\) is equal to the tolerance class of \(\tau\) defined by \(x\). The standard inclusion function \(\nu_{SRI}\) is defined for \(X, Y \subseteq U\) by (11).

\[
\nu_{SRI}(X, Y) = \begin{cases} \frac{|X \cap Y|}{|X|}, & \text{if } X \neq \emptyset, \\ 1, & \text{otherwise}. \end{cases} \quad \text{(11)}
\]

For applications, it is important to have some constructive definitions of \(N\) and \(\nu\).

\(^2\)Recall that a tolerance is a binary relation \(R \subseteq U \times U\) on a set \(U\) having the reflexivity and symmetry properties, i.e., \(xRx\) for all \(x \in U\) and \(xRy\) implies \(yRx\) for all \(x, y \in U\).
One can consider another way to define $N(x)$. Usually together with a GAS, we consider some set $F$ of formulas describing sets of objects in the universe $U$ of the GAS defined by semantics $\| \cdot \|_{GAS}$, i.e., $\| \alpha \|_{GAS} \subseteq U$ for any $\alpha \in F$. Now, one can take the set the neighbourhood function as shown in (12).

$$N_F(x) = \{ \alpha \in F : x \in \| \alpha \|_{GAS} \},$$  \hspace{1cm} (12)

and $N(x) = \{ \| \alpha \|_{GAS} : \alpha \in N_F(x) \}$. Hence, more general neighbourhood functions having values in $\mathcal{P}(U)$ can be defined and as a consequence different definitions of approximations are considered. For example, one can consider the following definitions of approximation operations in GAS defined in (13) and (14).

$$GAS_\circ (X) = \{ x \in U : \nu(Y, X) = 1 \text{ for some } Y \in N(x) \},$$  \hspace{1cm} (13)

$$GAS^\circ (X) = \{ x \in U : \nu(Y, X) > 0 \text{ for any } Y \in N(x) \}.$$  \hspace{1cm} (14)

Many similar functions can be defined. Let us consider two examples.

In the first example of a quasi-rough inclusion function, a threshold $t \in (0, 0.5)$ is used to relax the degree of inclusion of sets (see (9) and (10)). The function $\nu_t$ is defined by (15).

$$\nu_t(X, Y) = \begin{cases} 
1, & \text{if } \nu_{SRI}(X, Y) \geq 1 - t, \\
\frac{\nu_{SRI}(X, Y) - t}{1 - 2t}, & \text{if } t \leq \nu_{SRI}(X, Y) < 1 - t, \\
0, & \text{if } \nu_{SRI}(X, Y) \leq t.
\end{cases}$$  \hspace{1cm} (15)

**Remark 4.1.** The function $\nu_t$ defined in (15) is not a rough inclusion function, since it is possible to have $\nu_t(X, Y) = 1$ and $X \not\subseteq Y$.

One can obtain approximations considered in the variable precision rough set approach (VPRSM [53]) by replacing (9)-(10) with the function $\nu_t$ defined by (15) for $\nu$, assuming that $Y$ is a decision class and $N(x) = \{ B(x) \}$ for any object $x$, where $B$ is a given set of attributes.

Another example of application of the standard inclusion was obtained by using probabilistic decision functions (see, e.g., [54, 55]).

The rough inclusion relation can be also used for function approximation and relation approximation. In the case of function approximation the inclusion function $\nu^*$ for subsets $X, Y \subseteq U \times U$, where $X, Y \subseteq \mathbb{R}$ and $\mathbb{R}$ is the set of reals, is defined by (16).

$$\nu^*(X, Y) = \begin{cases} 
\frac{\text{card}(\pi_1(X \cap Y))}{\text{card}(\pi_1(X))}, & \text{if } \pi_1(X) \neq \emptyset, \\
1, & \text{if } \pi_1(X) = \emptyset,
\end{cases}$$  \hspace{1cm} (16)

where $\pi_1$ is the projection operation on the first coordinate. Assume now, that $X$ is a cube and $Y$ is the graph $G(f)$ of a function $f : \mathbb{R} \rightarrow \mathbb{R}$. Then, e.g., $X$ is in the lower approximation of $f$ if the projection on the first coordinate of the intersection $X \cap G(f)$ is equal to the projection of $X$ on the first coordinate. This means that the part of the graph $G(f)$ is “well” included in the box $X$, i.e., all arguments that belong to the projection of the box on the first coordinate the value of $f$ are included in the projection of box $X$ on the second coordinate.
Let us assume that the nearness relation $\delta$ is defined by (17).

\[
X \delta Y \iff \exists x \in X \cdot \exists y \in Y : y \in N(x).
\] (17)

Observe that, in general, the proximity relation defined by (17) does not satisfy the axiom (3.2) of a proximity space (see Def. 3.2). However, assuming that $x \in N(x)$ and $y \in N(y)$ for any $x, y \in U$, and additionally $X \delta Y$ for $X = Y = \emptyset$, we obtain that the relation $\delta \subseteq P(U) \times P(U)$ defined by (17) is a tolerance relation [5]. For example, for a given $\varepsilon \in [0, 1]$, a set of features $B$ with real values, and $x \in U$ one can define the neighbourhood of $x$ by

\[
N_B^\varepsilon(x) = \{ y \in U : \sum_{a \in B} \frac{|a(x) - a(y)|}{\max(a) - \min(a)} \leq \varepsilon \},
\]

where $\max(a)$, $\min(a)$ denote, respectively, the maximal and minimal number in the set $V_a$ of values of attribute $a$.

The following list contains a selection of properties of this form of an approximation space.

1. $\text{GAS}_0(X) \delta \{x\}$ for every $x \in \text{GAS}_0(X)$.
2. $\text{GAS}_0(X) \delta N(x)$, since $x \in N(x)$ will also be an object in $\text{GAS}_0(X)$.
3. $\text{GAS}_0(X) \delta \text{GAS}^\ast(X)$ if $\text{GAS}_0(X) \neq \emptyset$.
4. $N(x) \delta \{x\}$ for $x \in \text{GAS}_0(X)$.
5. $\text{GAS}_0(X) \delta \text{GAS}^\ast(X)$ if $\text{GAS}_0(X) \neq \emptyset$.

Usually families of approximation spaces labelled by some parameters are considered. By tuning such parameters according to chosen criteria (e.g., minimal description length), one can search for the optimal approximation space for a concept description.

Now, we would like to illustrate how the nearness relation can be used for approximation of concepts assuming that only partial information about concepts is available. Let us assume that $X \subseteq U^\infty$ where $U^\infty$ is the universe of all objects. However, only partial information about $X$ on a sample $U \subset U^\infty$ is given by a pair $(X \cap U, U \setminus X)$. Then the problem arises how to estimate the values $\nu(N(x), X)$ for $x \in U^\infty$ if $N(x) \cap U$ and $X \cap U$ are only available. Using the nearness relation $\delta$ on subsets of $U$, this can be done for $x \in U^\infty \setminus U$ as follows:

- $\nu(N(x), X) = 1$ iff $\nu(N(y) \cap U, X \cap U) = 1$ and $N(x) \delta N(y)$ for some $y \in U$,
- $\nu(N(x), X) > 0$ iff $\nu(N(y) \cap U, X \cap U) > 0$ and $N(x) \delta N(y)$ for some $y \in U$.

In this way, we have illustrated how to estimate the values of inclusion function for concept approximation if only partial information about concepts in a nearness generalized approximation space (NGAS)

\[
\text{NGAS} = (U^\infty, N, \nu, \delta),
\]

is given.
5. Nearness and Analogy-Making

The analogy-making is the central issue of cognition: Higher-level perception encompasses recognition, categorization, and analogy-making, and its central feature is the fluid application of one’s existing concept to new situation, often via conceptual slippage [17]. Analogy-making is defined in [18] as the perception of two or more non-identical objects situations as being the same at some abstract level. In making analogies elements of one situation are fluidly mapped to another situation [18].

In this section, we present a discussion about a possible application of nearness in analogy-making (see, e.g., [17, 18]).

Let us assume that there is a given hierarchy of languages of formulas defining sets of objects from a given universe. This hierarchy will be represented by a sequence $L_1, \ldots, L_i, \ldots$ of languages. The meaning of any formula $\alpha \in L_i$ is a subset of the universe $U$ and is denoted by $\|\alpha\|$. For example, one can consider the meaning of formulas from $L_i$ defined by some boolean functions. The meaning of formulas from $L_i$ are interpreted as concepts on the $i$th level of an abstract hierarchy of concepts. Now, let us recall the definition of nearness of sets given in (18).

\[ X \delta Y \text{ if and only if } \exists y \in Y \exists x \in X D_B(x, y) = 0. \]  

(18)

We assume that the considered sets belong to a given family $C$ of concepts (subsets of the universe $U$).

**Definition 5.1.** For positive integer $k$, two concepts $X, Y$ are $(B, k)$-near if $X \delta^{(k)} Y$, where $\delta^{(1)} = \delta$ and $\delta^{(k+1)}$ denotes composition of $\delta^{(k)}$ with $\delta$, for $k > 1$.

An important problem in analogy-making can be formulated as follows:

**Problem 1. Analogy Making** For given $(B, k)$-near concepts $X, Y \in C$, find a minimal $n$ such that for some $\alpha \in L_n$ the following condition is satisfied:

\[ B^*X \cup B^*Y \subseteq \|\alpha\|. \]

Certainly, this is only the first step in the formulation of the analogy-making problem. For example, there can be many formulas satisfying the above condition. Then some more advanced approach should be developed bearing in mind that it seems that humans have evolved in such a way as to make analogies in the real world that affect their survival and reproduction, and their analogy-making ability seems to carry over into abstract domain as well [18]. One can see here some analogies to adaptive clustering. Discovered clusters are used as relevant patterns in solving some tasks, such as classification or prediction. Discovering such clusters is closely related to conceptual slippage [17].

**Conclusion**

One of the outcomes of the approach to nearness of objects proposed in this paper is a view of proximity of objects to a degree. This paper has also presented an extension of approximation spaces introduced by Pawlak as well as generalized approximation spaces based on the introduction of a nearness relation. The significance of these extensions is that we now have the possibility of measuring our knowledge about objects based on the perception of the “nearness” of objects classified by means of attributes...
or features. The proposed extension is a direct outgrowth of the basic approach to classifying objects proposed by Zdzisław Pawlak during the early 1980s. The complement of the nearness relation (i.e., δ) suggests the possibility of extending the model for conflict also proposed by Pawlak, which is based on the complement of the indiscernibility relation (see, e.g., [24, 25, 27, 28]), especially in the context of approximation spaces [41]. In our further study, we plan to investigate computational complexity of different analogy-making problems and heuristics searching for solutions of such problems [17].

Acknowledgments. The research of James F. Peters is supported by NSERC grant 185986. The research of Andrzej Skowron is supported by the grant 3 T11C 002 26 from Ministry of Scientific Research and Information Technology of the Republic of Poland. The research of Jaroslaw Stepaniuk is supported by Bialystok University of Technology Rector’s grant.

References


